



HHS Public Access

Author manuscript

Prog Neuropsychopharmacol Biol Psychiatry. Author manuscript; available in PMC 2019 December 20.

Published in final edited form as:

Prog Neuropsychopharmacol Biol Psychiatry. 2018 December 20; 87(Pt A): 22–32. doi:10.1016/j.pnpbp.2017.06.029.

Defining the place of habit in substance use disorders

Youna Vandaele¹ and Patricia H. Janak^{1,2}

¹Department of Psychological and Brain Sciences, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore MD 21218, USA

²The Solomon H. Snyder Department of Neuroscience, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore MD 21205 USA

Abstract

It has long been suggested that alcohol or substance use disorders could emerge from the progressive development and dominance of drug habits. Like habits, drug-related behaviors are often triggered by drug-associated cues. Like habits, addictive behaviors are strong, rigid and “hard to break”. Like habits, these behaviors are insensitive to their outcome and persist despite negative consequences. “Pathological habit” thus appears as a good candidate to explain the transition to compulsive drug use. However, drug use could also be considered as a goal-directed choice, driven by the expectation of drug outcomes. For example, drug addicts may engage in drug-seeking behaviors because they view the drug as more valuable than available alternatives. Substance use disorders therefore may not be all about habit, nor fully intentional, and could be considered as resulting from an imbalance between goal-directed and habitual control. The main objective of this review is to disentangle the relative contribution of habit formation and impairment of goal-directed behavior in this unbalanced control of addictive behaviors. Although deficits in goal-directed behavior have been demonstrated in alcohol and substance use disorders, reliable demonstration of abnormal habit formation has been curtailed by the paucity of paradigms designed to assess habit as a positive result. Refining our animal and human model of habit is therefore required to precisely define the place of habit in substance use disorders and develop appropriate and adapted neurobehavioral treatments.

Keywords

habit; goal-directed; drugs; addiction; learning; stimulus-response

Introduction

Substance use disorders (SUD) are commonly defined as chronic, relapsing neurobiological diseases characterized by compulsive drug use despite negative consequences. The notion of compulsion has been used to explain (1) the dissonance between an individual’s thoughts and actions – the addict wants to stop but continues drug use – and, (2) the puzzling paradox

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

of addictive behaviors – the pursuit of drug use despite knowledge that this behavior has strong negative consequences. To explain compulsive drug use, it has long been suggested that addictive behavior could emerge from the progressive development and dominance of drug habits (Everitt and Robbins, 2005; Pierce and Vanderschuren, 2010; Robbins and Everitt, 1999; Tiffany, 1990). In contrast to flexible goal-directed behaviors which are performed because they are expected to produce some desirable outcomes, habitual behaviors are characterized by rigid, inflexible and retrospective actions, triggered by associated stimuli, and performed in spite of their immediate consequence (Balleine and Dickinson, 1998; Dickinson et al., 1983; Ostlund and Balleine, 2008). In the context of SUD, if drug-seeking behaviors are habitual, then these behaviors should be insensitive to their associated negative consequences on the addict's life. Therefore, the concept of "pathological" drug-seeking habit allows the resolution of the puzzling paradox of maintained drug use despite negative consequences.

An alternative approach to explain SUD is to consider drug use as an intentional though suboptimal choice (Ahmed, 2010; Heyman, 2010). In this framework, addictive behavior would arise from a higher estimated value for drug than for alternative rewards. For example, drug addicts may view the cost of abstinence to be higher than its benefits (Pickard and Ahmed, 2015). This may be particularly true given the uncertainty and delay associated with benefits of abstinence, while drug intake can produce immediate and certain satisfaction or relief from withdrawal symptoms. According to this account, drug use would be a goal-directed choice, driven by the expectation of drug outcomes.

These two approaches of SUD considering drug-related behavior as an intentional choice or as a pathological habit appear to be in direct opposition since drug seeking cannot be concurrently goal-directed and intentional on one side, and habitual, automatic and unconscious on the other. Dual-process theories help us resolve this contradiction by suggesting that drug seeking can be mediated by goal-directed and habitual controllers in parallel, with one system predominantly engaged at any given time depending on several factors. For example, repetition of a given behavior in the same situation through practice or overtraining facilitates the formation of habit (Adams, 1982; Dickinson, 1985; Dickinson et al., 1995; Holland, 2004). The first cigarette after waking up, or the glass of wine with the dinner are good illustrations of habitual and automatic drug-taking responses, established through repetition and closely linked to contextual and conditioned stimuli (stimulus-response associations). The goal-directed system remains however available and if one makes the decision to reduce cigarette or wine consumption, an intentional goal-directed control can be exerted over habitual systems to adjust the behavior and comply with one's New Year's resolutions. Non-addicted individuals can do that with varying success and can control their drug intake because there is a balance between goal-directed and habitual systems, and transitions are possible from one system to another. A growing body of research suggests that this balance is disrupted in SUD (Ersche et al., 2016; Everitt and Robbins, 2016, 2005; Ostlund and Balleine, 2008; Sjoerds et al., 2013; Voon et al., 2015). Persistence in drug use could then be explained by: 1) abnormal formation of habits, 2) impairment in goal-directed behaviors, with biases toward drug-related goals, 3) deficits in goal-directed control over habits, or a combination of these processes.

The objective of this review is to consider the relative contribution of these three cognitive processes in SUD, with a special emphasis on the concept of habits. Indeed, this concept remains unclear, with multiple definitions used, along with different approaches toward operationalization of the concept in distinct subfields of human and animal psychology. In this review, we first discuss the main definitions of habit and the models classically used to study them. In the second part, we review the literature using these models to consider how drugs of abuse and addictive processes may affect the balance between habitual and goal-directed control to promote compulsive behaviors.

Defining and modeling the concept of habit

Habit has several meanings and definitions. The informal conception of habit is often related to the notion of acting without thinking: it may be a behavioral routine, for example when driving to and from work each day. It may also be related to motor skills: individual actions following in a behavioral sequence without requiring attention or explicit cognitive control. But we also often think about habit after making a mistake, for example, when trying to switch on the light after an electric outage. In this context, habit is considered as the expression of a stimulus-response association. Consider a person who opens the refrigerator every time he or she enters the kitchen. This is not a routine, nor is it a skill. The stimulus of the refrigerator, in the context of the kitchen, triggers the automatic response of opening the refrigerator, even though the person may not be planning to cook or to eat. These examples illustrate the existence of several definitions of habit. Since these variants are often studied in different procedures, it is unclear whether they are correlated and part of one single habit concept or are independent and reflect distinct neural and cognitive processes. In the following section, we will focus on the definition of habit as a stimulus-response association, which has had the most influence on the study of habit in SUD.

Habit as a stimulus-response association

In the field of animal psychology, a goal-directed behavior is considered dependent upon a Response-Outcome (R-O) association, the response being mediated by the expectation of the outcome, whereas a habit instead relies on a Stimulus-Response (S-R) association (Dickinson, 1994, 1985). During habitual learning, the outcome strengthens the S-R association through repeated experience of the R-O contingency in presence of stimuli, but is not directly encoded in the associative structure driving the performance (Adams and Dickinson, 1981). Since expression of habitual responding is independent of the outcome, one way to discriminate between goal-directed and habitual behavior is to question the reliance on the R-O association. More specifically, goal-directed behavior relies on (1) the knowledge of a causal relationship between the response and the outcome, and (2) the current motivational value of that outcome (Balleine and Dickinson, 1998; Dickinson, 1994, 1989). Therefore, reducing the contingency between the response and the outcome, or reducing the value of the outcome, should lead to a reduction in goal-directed responding. In contrast, habitual behavior should not be affected by these manipulations.

Reducing the attribution of a causal relationship between the response and the outcome can be achieved with R-O contingency degradation procedures (Balleine and Dickinson, 1998)

(figure 1A). In instrumental procedures, rodents can be trained to perform one action (i.e. lever pressing) to obtain a reward, and another action (i.e. chain pulling) to obtain an alternative reward. During the test, one of the outcomes is delivered non-contingently such that its probability of delivery in the session is equally likely if the rats respond appropriately or not. The contingency of this R-O association is thus degraded. The alternative outcome remains contingent upon the corresponding response. When responding is goal-directed, the performance of the degraded response is reduced compared to the non-degraded alternative. If the behavior is habitual, then performance is insensitive to the contingency degradation.

The second way to discriminate between goal-directed and habitual behavior is to devalue the outcome (Adams and Dickinson, 1981; Balleine and Dickinson, 1998; Colwill, 1993; Dickinson, 1994; Rescorla, 1987)(figure 1A). This can be achieved using sensory-specific satiety or conditioned-taste aversion. In sensory-specific satiety, animals are given about 1 hour of unlimited access to the training outcome immediately before a test session conducted in extinction (in the absence of reward). During the test, animals that reduce their performance of the response that, in training, had earned the now devalued outcome are goal-directed, because they rely on the representation of the current motivational value of the outcome. The reward can also be devalued by pairing its consumption with a systemic injection of lithium chloride, which creates a gastro-intestinal distress and results in the development of a conditioned taste aversion. While devaluation by specific satiety affects the motivational state of the animal, conditioned taste aversion directly affects the taste perception of the outcome. Both procedures lead to a reduction in performance when that performance is goal-directed.

Numerous experimental conditions including the schedule of reinforcement (Dickinson et al., 1983), number of R-O associations (Colwill and Rescorla, 1985; Colwill and Triola, 2002; Holland, 2004), stress (Dias-ferreira et al., 2009), context (Ostlund et al., 2010; Thrailkill and Bouton, 2015), length of training (Adams, 1982; Dickinson, 1985; Holland, 2004) and exposure to drugs (Corbit et al., 2014, 2012; Nelson and Killcross, 2006; Nordquist et al., 2007) are known to specifically promote goal-directed or habitual control. Among these factors, the schedule of reinforcement is the most convenient way to bias rodent behavior toward goal-directed or habitual responding; random ratio (RR) schedules of reinforcement have been shown to promote goal -directed behavior while random interval (RI) schedules and overtraining promote the formation of habit (Dickinson et al., 1983). Dickinson et al. (1983) suggested that this differential effect of RR and RI on the control of instrumental performance is related to differences in contingency between the response rate and the reinforcement rate. After overtraining or under RI schedules, the uncoupling between response rate and reinforcement rate would facilitate habitual learning. More recently, it was proposed instead that uncertainty (determined by the probability of reward delivery) is the main factor driving the formation of habit in RI schedules (Daw et al., 2005; Derusso et al., 2010). Indeed, training rats under fixed- (and therefore predictable) interval schedule does not promote habit even though the training contingency between response rate and reinforcement rate is low. In contrast, increasing the uncertainty of reward delivery by reducing the probability of reinforcement in random interval schedule does not affect contingency, but decreases temporal contiguity (mean time between lever pressing and reward delivery) and promotes habitual response strategies (Derusso et al., 2010).

The distinction between RR and RI schedules of reinforcement illustrates the main limit of this habit model; expression of habit in this model is defined as a negative result – an absence of goal-directed behavior. Devaluation tests and degradation contingency procedures only probe the existence and expression of R-O associations. Thus, it makes sense that training procedures like RI schedules, that weaken the R-O association by reducing the contingency or the contiguity between the response and the outcome, lead to an expression of behavior less sensitive to the representation of the outcome, thus qualifying as habit. RI schedules create habits that are artificially promoted by weakening R-O associations, and may not accurately map onto long-term S-R-O experiences that shift from R-O to S-R control of performance with extended training.

Interaction between habitual and goal-directed control

Devaluation procedures have been successfully adapted to human research, using sensory-specific satiety (Hogarth and Chase, 2011; Tricomi et al., 2009; Valentin et al., 2007) or an explicit instruction that a given outcome no longer earns points (de Wit et al., 2009). Interestingly, a new outcome devaluation paradigm in which S-R and R-O systems compete for the control of performance, was developed to assess the balance between goal-directed and habitual control in human subjects (de Wit et al., 2012)(figure 1B). In the training phase of the task, participants are presented with pictures (the stimuli) and with repeated practice learn which associated response (a left or right key press) earns points signaled by a subsequent picture (the Outcome). Some outcomes are then devalued (no longer earn points) and the test assesses whether participants can suppress previously-learned responses that no longer yield valuable outcomes, while continuing to respond for still-valuable outcomes. A failure to inhibit prepotent responding for devalued outcomes is considered as “slips of action” and is interpreted as relative reliance on S-R habit over R-O goal-directed control. Using this task, an index can be calculated to measure the relative involvement of habitual versus goal-directed control of action. While this task allows assessment of how S-R habit and R-O goal-directed compete for the control of performance, a second approach emerging from neuro-computational psychology suggests that habit and goal-directed control could operate in parallel.

In this neuro-computational framework, goal-directed and habitual systems have been conceptualized as relying on model-based and model-free control, respectively (Daw et al., 2005; Dolan and Dayan, 2013; Doya et al., 2002). Model-based control is characterized by computationally-demanding prospective planning in which a decision tree of all the possible consequences of each action is built on the basis of an internal model of the world. Thus, like goal-directed behavior, actions under model-based control are flexible and slower to execute (because all the possible paths have to be analyzed). In contrast, model-free control is defined as a retrospective decision-making process relying on iterative updating of expectation, based on reward prediction error (Sutton, 1988). Under model-free control, a given situation, i.e., a “state”, or set of stimuli, acquires a cached value (stored scalar value) through repeated experience of the outcome, and comes to elicit a specific response. These cached values depend on the reward history and are thus unrelated to predicted outcomes. Unlike model-based control, model-free control does not require the simulation of future possible outcomes or a representation of logical transitions between states. It is therefore

computationally efficient, with a faster execution, but also less flexible. Like habit, an instrumental action under model-free control has no immediate sensitivity to outcome devaluation because the update of new cached value, resulting from a reduction in outcome utility, can only be acquired iteratively through direct experience (Daw et al., 2005; Dolan and Dayan, 2013).

One procedure designed to study model-based and model-free control is the sequential 2-stage Markov decision task (Daw et al., 2011; Gläscher et al., 2010) (figure 1C). In the first stage, 2 options are presented (A vs. B): selection of one option (option A) commonly leads to a 2nd choice situation between a specific pair of stimuli (C vs. D; probability 70%). More rarely, this choice in stage 1 leads to a different 2nd choice situation between another pair of stimuli (E vs. F; probability 30%). The probabilities of the 2nd stage set of stimuli are reversed following the choice of the alternative option in stage 1 (option B). The selection of options in the second stage is rewarded or not, according to variable and unpredictable probabilities. Under model-free control, subjects will repeat the actions previously rewarded, irrespective of the likelihood of the first transition. In contrast, subjects using model-based control are proposed to build an internal representation of task structure to inform their choice behavior, thus taking into account the likelihood of transitions. For example, after rewarded choices following rare transitions (ex: A → F → reward), model-based subjects will tend to choose the alternative option at step 1, because they know that it will more likely lead to the presentation of the previously rewarded stimulus during the second step. In this example, they would choose option B because this is more likely to gain access to option F. Such decision-making demonstrates an understanding of the causal structure of the task.

Using this model, it was shown that model-based and model-free control are engaged in parallel during the task and that subjects rely on both systems to make decisions (Daw et al., 2011; Otto et al., 2013). In addition a large inter-individual variability in the degree of dependence on each system was found (Otto et al., 2013). This task thus allows computing of indexes of model-free and model-based control in individuals that can be then correlated with neural activity of specific brain regions, with particular personality traits, psychiatric symptoms, or with a history of drug dependence, for example. In this approach, model-free control is not defined as an absence of model-based control but can be computed and estimated independently, which constitutes an important advantage for delineating the relative contributions of goal-directed and habitual control. To our knowledge, neither the slips of action task nor the sequential two-stage task have been adapted to rodents or monkeys, which limits its benefits thus far to human correlative studies.

Interestingly, the definition of a model-free system seems to match well with the S-R account of habit (Dezfouli and Balleine, 2012; Dolan and Dayan, 2013; Sutton, 1988). For instance, in the model-free actor-critic algorithm, action selection involves two interacting components: one learning the prediction of reward value and computing prediction errors (the critic), and another forming a policy to select a given action in a particular state (the actor). The reward prediction error, implemented by phasic dopaminergic signal (Daw and Doya, 2006; Keiflin and Janak, 2015; Schultz et al., 1997) is then used to update (1) the predictive value for each state and (2) the policy in that state. The policy and value signal thus contribute to the formation of a state-action association with an error correction

feedback signal to acquire and shape actions. Even though this state-action (or stimulus-response) association is a core component of the model-free system, the dopaminergic prediction error signal has also been implicated in model-based control (Dolan and Dayan, 2013; Gläscher et al., 2010; Sharpe et al., 2017) suggesting that the computations described above may not be specific to the habitual, model-free system.

It was recently suggested that the propagation of dopaminergic prediction error signals in spiraling striatal circuits could constitute the neural basis for a hierarchical decision-making model involving different levels of abstraction, from high-level cognitive plans to low-level motor execution (Keramati and Gutkin, 2013). In this computational model, an abstract cognitive plan (i.e. desire for a glass of wine) can be broken into a sequence of lower level actions (i.e. reach for bottle – reach for glass – pour the wine into the glass...etc). The dopaminergic prediction error signal early in the spiraling striatal circuits (i.e. nucleus accumbens) would reinforce the abstract motivational level, while its propagation in dorsolateral parts of striatum would reinforce lower-level automatic actions (Keiflin and Janak, 2015; Keramati and Gutkin, 2013). Although this model was implemented with model-free computations (TD model of reinforcement), it is tempting to draw a parallel between higher level cognitive plans and model-based goal-directed decision-making. With the same logic, propagation of the dopaminergic prediction error signal within spiraling striatal circuits would underlie the transition from model-based to model-free control (Keiflin and Janak, 2015). This hierarchical approach of decision making is closely related to a series of studies suggesting that in real-life situations, every choice is not either goal-directed or habitual but instead, high-level goal-directed choices can be implemented through habitual action sequences, to attain desired goals (Dezfouli et al., 2014; Dezfouli and Balleine, 2013, 2012). More recently, another computational approach suggested an alternative hierarchical organization of human decision-making: habitual control is exerted over goal selection, and the goal selected is then attained through planning and deliberation (Cushman and Morris, 2015). While these two hierarchical models suggest an opposite relationship between habit and goal-directed behavior, they share the common assumption that goal-directed and habit systems are integrated and complementary, rather than competitive.

While the above analysis indicates a need for future experimental work for determining reliable measures to specifically assess the development and expression of S-R habit, the models used to date have proved highly valuable for understanding the neural basis and cognitive processes underlying changes in the control of instrumental performance. Notably, findings from the use of these models point to goal-directed and habitual systems that are engaged in parallel during learning, and are in agreement with the notion that repetition and practice promote habitual responding, while goal-directed control can be exerted over habit in situations requiring deliberation and flexibility. Thus, in normal conditions, there is a balance between goal-directed and habitual system.

Habit and Substance use disorders

Several lines of evidence suggest that the balance between habitual and goal-directed control of behavior is disrupted in SUD. Indeed, addictive drugs have been proposed to promote the

formation of habit (Everitt and Robbins, 2005), impair goal-directed control over habit (Baler and Volkow, 2006), and bias goal-directed choice toward drug-related activities (Lucantonio et al., 2014a). Although our current models make it difficult to determine whether drugs alter the balance by promoting habit or by impairing goal-directed control, we will review the literature about control of action and SUD with an eye towards disentangling the relative contribution of these two systems in addictive behaviors. More specifically, we will describe how drugs of abuse disrupt the balance between habit and goal-directed behavior, and how these changes can lead to the development of addictive behaviors. Following that, we will see that this unbalance could more specifically result from impairments in the goal-directed system.

Unbalanced goal-directed and habitual control; bias toward habit

Since the seminal work of Dickinson and colleagues in 2002 (Dickinson et al., 2002), it has been repeatedly shown that drugs of abuse can promote the formation of habit, as defined using devaluation and contingency degradation procedures. For example, using devaluation procedures, several studies in rodent suggest that responding for cocaine (Miles et al., 2003), ethanol (Corbit et al., 2012; Dickinson et al., 2002; Lopez et al., 2014; Mangieri et al., 2012) and nicotine (Clemens et al., 2014; Loughlin et al., 2017) become habitual with extended training or under RI schedule of reinforcement (but see Loughlin et al., 2017 for limited FR1 training). These results have to be interpreted with caution because devaluing a drug can be challenging, particularly when this drug is self-administered intravenously. Some of these studies suffer from unspecific devaluation effect (Dickinson et al., 2002), lack of devaluation effects (as measured on consumption) (Lopez et al., 2014), or the use of sucrose substitution procedures (Dickinson et al., 2002; Mangieri et al., 2012; Miles et al., 2003), which promote the association of the instrumental response with both the drug and the sucrose. However, a more recent ethanol self-administration study that did not use sucrose substitution, demonstrated a faster reliance on habit system for rats trained with ethanol compared to rats trained with natural reward, with a progressive loss of sensitivity to devaluation across training (Corbit et al., 2012). Interestingly, chronic ethanol intoxication itself was able to promote habitual responding for sucrose seeking in this study. Facilitation of habit learning with natural reward has also been observed after chronic exposure to cocaine (Corbit et al., 2014; LeBlanc et al., 2013) and amphetamine (Nelson and Killcross, 2006; Nordquist et al., 2007), presumably through dopaminergic sensitization. Post-training administration of cocaine immediately after instrumental responding for a natural reward was sufficient to induce or accelerate habitual control over performance (Schmitzer-Torbert et al., 2015). Using contingency degradation procedure, it was also shown that ethanol (Mangieri et al., 2014) and cocaine exposure (Gourley et al., 2013) promote the expression of habit. Finally, early life cocaine exposure in rat can strengthen habit expression during adulthood, presumably by increasing BDNF expression in the medial prefrontal cortex (Hinton et al., 2014). Acute or chronic intoxication to various drugs of abuse induces critical changes in cortico-striatal loops, notably in the dorsal striatum, that may participate to this shift from goal-directed to habitual drug seeking behavior (Barker et al., 2015; Belin et al., 2009; Corbit and Janak, 2016a; Gremel and Lovinger, 2016; Ron and Barak, 2016).

These data from animal models have been confirmed in human, in experiments showing that alcohol dependent patients are less sensitive to outcome devaluation (Sjoerds et al., 2013), and that performance of methamphetamine (Voon et al., 2015) and alcohol (Sebold et al., 2014) dependent subjects is biased toward model-free control in the 2-stage decision task. A recent study replicated these results in cocaine addicted patients in an outcome devaluation paradigm (Ersche et al., 2016). Interestingly, cocaine addicts were also less efficient in suppressing S-R responding for devalued outcomes in the slips of action task, providing additional evidence for an imbalance between goal-directed and habitual behavior in addicts (Ersche et al., 2016). Although the formation of new S-R associations in drug addicts is not altered, their ability to overcome familiar S-R associations is disrupted, as indicated by more perseverative responding after a change in contingency (McKim et al., 2016). By using distinct computations to estimate model-based and model-free components of decision-making in the sequential 2-stage task, Sebold and colleagues' study suggests that the imbalance between goal-directed and habitual control in alcoholic patients results from impaired goal-directed (model-based) function, rather than promotion of habitual (model-free) responding (Sebold et al., 2014).

Interestingly, the shift from goal-directed behavior to habit induced by drugs of abuse does not require high exposure or extended drug self-administration. Indeed, moderate doses or limited exposure are sometimes sufficient to promote habitual responding (Gourley et al., 2013; Loughlin et al., 2017; Schmitzer-Torbert et al., 2015). Moreover, drug-induced habits are not specific to drug-seeking behaviors, but also applies to natural or monetary rewards (Corbit et al., 2014, 2012; Nelson and Killcross, 2006; Nordquist et al., 2007; Sebold et al., 2014; Voon et al., 2015). If drug-induced habits do not require heavy drug use and are not specific to drug-related behaviors, how could it lead to compulsive drug seeking specifically? One line of thought is to consider that behavior becomes in general more rigid. Drug addicts become more prone to develop routines, or S-R associations and this tendency interacts with every day choices in real-life situations. Repeated exposure to drug-associated contexts (places in which the drug is consumed, social circle) or drug-associated stimuli in addicts may facilitate habitual behavior in response to these drug contexts and conditioned stimuli (CS), more specifically.

A number of studies highlight the significant role of drug CS and contexts in the maintenance and persistence of drug seeking behavior (Berridge and Robinson, 2016; Caggiula et al., 2001; Corbit and Janak, 2016a; Crombag et al., 2008; Jentsch and Taylor, 1999; Robinson and Berridge, 1993). Nicotine for example, can magnify the incentive salience and reinforcing properties of drug CS (Caggiula et al., 2001; Chaudhri et al., 2006; Chiamulera, 2005). In absence of nicotine, these stimuli can sustain instrumental responding in rats (Donny et al., 2003; Palmatier et al., 2006), and elicit satisfaction or reduce subjective craving in human smokers (Brauer et al., 2001). It was shown that contextual cues are involved in the associative structure of the habit system, and can account for residual responding after devaluation of the reinforcer (Thrailkill and Bouton, 2015). It is therefore noteworthy that exposure to an alcohol-paired context during the devaluation test was sufficient to bias rats' behavior to the expression of habit (Ostlund et al., 2010). Drug CSs are sometimes reported to be stronger than natural reward CSs (Kearns et al., 2011); for example, a single experience of pairing a cue with cocaine self-administration is sufficient to

Author Manuscript

Author Manuscript

Author Manuscript

elicited cocaine seeking up to 9 month after extinction and forced abstinence, an effect not observed with sucrose self-administration (Ciccocioppo et al., 2004). It was also shown that the influence of drug CS on drug-seeking behavior grows over time, a phenomenon called “incubation of craving”(Grimm et al., 2001). This incubation of craving is stronger and more persistent for drugs (peak of CS-elicited drug seeking after up to 3 month of abstinence) compared to natural rewards (peak of sucrose seeking after 1 month of abstinence)(Lu et al., 2004). Finally, Pavlovian to instrumental transfer (PIT) has been demonstrated with ethanol self-administration (Corbit and Janak, 2007), and it was recently shown that the strength of this effect increases over extended ethanol self-administration (Corbit and Janak, 2016b). These results suggest that instrumental performance becomes more sensitive to Pavlovian influence after extended drug use (Corbit and Janak, 2016a). It is therefore possible that the overreliance on the habit system in SUD combined with stronger sensitivity to the influence of drug CSs, may bias habitual behavior toward drug seeking behavior and away from alternative activities.

But drug CSs can also elicit craving for drug through the conscious representation of the outcome including the satisfaction produced by the drug or the alleviation of withdrawal symptoms. Behavior is then goal-directed. This goal-directed choice could however trigger an automatic sequence of behavior that would be habitual and insensitive to the outcome, as suggested by the hierarchical decision-making theory (Dezfouli et al., 2014; Keramati and Gutkin, 2013). For example, the goal-directed choice of going to the bar to get a drink, may lead to an automatic sequence of drinking glasses of alcohol, one after the other, until the dawn. The smell of the smoke can elicit a smoking desire (conscious or unconscious) which can then trigger a motor program leading to the action of smoking (Keiflin and Janak, 2015). Once the decision of smoking is made, breaking the sequence is very hard because every action naturally and logically follow the previous one, without further deliberation. Accelerated propagation of dopaminergic prediction error signal in striatal spiraling loops following drug exposure could participate to bias hierarchical decision-making to “low level” automatic sequences of drug-related actions (Keiflin and Janak, 2015; Keramati and Gutkin, 2013).

Drug effects on sequence learning have received little attention in the field of habit and SUD, but one study from Zapata and colleagues provides some insights (Zapata et al., 2010). In this study, rats were trained to self-administer cocaine according to a seeking/taking chained schedule. To obtain an injection of cocaine, rats had to make one lever press on the “drug-seeking” lever followed by another lever press on the second “drug-taking” lever. To investigate the sensitivity to outcome devaluation, responding on the drug-taking lever was extinguished after short and extended training. During the test, performance on the drug seeking lever was measured in absence of drug-taking lever. It was shown that responding on the seeking lever was reduced after short training, but maintained after extended training. Although the results can be interpreted as a transition to habitual drug seeking with extended drug access, alternative interpretations were proposed (Holmes and Clemens, 2011). Chunking of the drug-seeking and drug-taking responses may have occurred during extended training, and since the outcome of the full sequence remains valued (cocaine is not devalued), the seeking response persists despite the devaluation of the middle link in the chain. Thus, insensitivity to devaluation of the taking response with extended cocaine access

could result from sequence learning, which provides some support for the hierarchical decision-making theory. Although this theory is appealing to explain the intricacy of goal-directed and habitual control in real-life decisions, further studies are needed to determine the effect of drug exposure on hierarchical decision-making, and the place of imbalanced decision hierarchy in the development of SUD.

Impairment of goal-directed control

The multiple studies discussed above converge to demonstrate an imbalance between habit and goal-directed control in SUD, with a bias toward more habitual drug-seeking behaviors. Does this imbalance result from an abnormal formation of habit, an impairment in goal-directed decision-making, or both? Since habit is defined in most models as a negative result, it is difficult to directly answer this question. Positive evidence that exposure to drugs of abuse increases the formation of S-R association are still missing because models used to investigate habit typically do not assess this specific question. Notably, two studies in humans have addressed this question and suggest that substance use disorders are associated with impairment of goal-directed control rather than accelerated formation of habit (McKim et al., 2016; Sebold et al., 2014). But whether or not abnormal habits are involved, habitual learning is probably not the only process at stake in this imbalance. Indeed, numerous studies suggest that goal-directed control can be impaired during SUD, particularly when decisions involve the pros and cons of drug abstinence. This cognitive impairment likely participates to tilt the balance toward habitual control of drug seeking behavior.

SUDs are associated with a wide range of changes in cortico-striatal circuits which can lead to impairment of several executive functions (Baler and Volkow, 2006; Barker et al., 2015; Everitt and Robbins, 2016; Gremel and Lovinger, 2016). In this framework, drug-induced dysfunction in the orbitofrontal cortex (OFC) has received growing interest over the last decade (Lucantonio et al., 2012; Schoenbaum et al., 2016). Indeed, numerous neuroadaptations have been reported in the OFC after chronic exposure to drugs of abuse. At a structural level, decrease in gray matter concentration has been reported among cocaine users (Ersche et al., 2011; Franklin et al., 2002). The degree of gray matter reduction was correlated with compulsivity, as measured by the compulsive drug use scale, and the duration of cocaine dependence (Ersche et al., 2011). Abnormalities in frontal white matter have also been described for alcohol, cocaine, opioid and methamphetamine dependent patients (Bae et al., 2006; Bühler and Mann, 2011; Lyoo et al., 2004). In human, binge drinking is associated with disturbed dendritic complexity in higher-order prefrontal regions (Morris et al., 2017), and in rodent, chronic exposure to psychostimulants (DePoy and Gourley, 2015) and withdrawal from ethanol (McGuier et al., 2015) decreases and increases dendritic spine density of OFC neurons, respectively. After chronic alcohol exposure, electrophysiological and molecular changes have been found in primates and rodents (Laguesse et al., 2016; Nimitvilai et al., 2017, 2016). At a functional level, OFC is hypoactive during cocaine use as indicated by a general decrease in glucose metabolism (London et al., 1990; Volkow et al., 1990). In contrast, this same region is hyperactive during early period of drug abstinence and this increased in activity is correlated with the intensity of drug craving (Volkow et al., 1991). Finally, protracted withdrawal from cocaine or

methamphetamine is associated with hypoactive OFC, correlated with a decrease in D2 receptors in the striatum (Volkow et al., 1993).

A series of elegant experiments in rodents demonstrated the critical role of OFC in model-based control (Gremel et al., 2016; Gremel and Costa, 2013; Schoenbaum et al., 2016, 2011). These studies show that OFC is required for guiding behavior on the basis of information about expected outcome, specifically when such information must be derived from the structure of the environment. The Pavlovian over-expectation task provides a good demonstration of OFC function, and cocaine associated deficits. In this task, rats first learn that several cues predict reward delivery. Then, in a second phase, 2 of the cues are presented together in compound, followed by the same reward. This results in an increase in responding reflecting higher reward expectation due to a process defined as summation. In the last probe test phase, cues are presented separately and without reward. Responding to the isolated cue is now lower compared to a control cue not involved in compound training. This task involves predictions about outcomes and operations on these simulated predictions (summation). Interestingly, inactivation of OFC during the compound phase suppressed the summation effect, and the reduction of responding during probe test trials, demonstrating the role of this region in predictions of future outcomes (Takahashi et al., 2009). It was shown that chronic cocaine exposure disrupted these same behaviors in the over-expectation task, and these cocaine-induced deficits were restored with an optogenetic excitation of OFC neurons during the compound phase (Lucantonio et al., 2014). These results strongly suggest that OFC-mediated model-based behavior can be impaired in SUD. This impairment could explain the poor performance of drug addicts in a gambling task involving real-life decisions based on uncertain reward and punishment (Balconi et al., 2014; Bechara et al., 2002, 2001; Bechara and Martin, 2004).

Additional evidence of impaired goal-directed behavior in SUDs comes from a study investigating the relation between deficits in model-based control and the occurrence of symptoms from psychiatric disorders involving or not a compulsive component (Gillan et al., 2016a). Using a dimensional approach, this study shows that indexes of eating disorders, impulsivity, obsessive compulsive disorder (OCD) and alcohol use disorders are associated with lower reliance on model-based learning. A factor analysis reveals that deficit in model-based control is associated with a symptom dimension involving “compulsive behavior and intrusive thought”, suggesting that impaired goal-directed behavior would be a good trans-diagnostic symptom of psychiatric disorders involving clinical compulsions (Gillan et al., 2016b).

Impaired model-based goal-directed control may explain drug addicts' difficulty in considering future negative consequences of drug-related behaviors and could bias choices toward more immediate drug gratifications (Bickel et al., 2012; Camchong et al., 2011; Kirby et al., 1999; Madden et al., 1997). Furthermore, certain situations or contexts can prompt individuals to express habit, and failures in response inhibition could then interfere with individual's ability to exert goal-directed control over that habitual responding. The relation between SUD and deficits in inhibitory control is well documented in the literature (Baler and Volkow, 2006; Morein-Zamir and Robbins, 2015). For example, alcoholics (Noël et al., 2007) and stimulant-dependent individuals (Ersche et al., 2012; Monterosso et al.,

2005; Morein-Zamir et al., 2013) exhibit higher motor impulsivity, defined as poor inhibitory control associated with deficits in the ability to withhold prepotent responding. In rodents, abstinence from chronic alcohol or 9-THC exposure is associated with increases in motor impulsivity, as indicated by higher premature responding in the 5-choice serial reaction time task (Irimia et al., 2015a, 2015b; Walker et al., 2011). In addition, several studies demonstrated that motor impulsivity can precede and predict subsequent drug self-administration (Dalley et al., 2007; Diergaarde et al., 2008; Economidou et al., 2009), suggesting that it could constitute a vulnerability factor for SUD (Dalley et al., 2011; Morein-Zamir and Robbins, 2015). Impaired response inhibition in substance use disorder is generally thought to reflect reduced self-control resulting from prefrontal cortex dysfunctions (Baler and Volkow, 2006; Goldstein and Volkow, 2012). This reduction in top-down cognitive control can dampen goal-directed control and create opportunity for impulsive habit expression. Accordingly, motor impulsivity measured with the Barret's impulsivity scale has been selectively associated with a deficit in goal-directed control in an outcome devaluation paradigm (Hogarth et al., 2012), and in the sequential 2-stage task (Gillan et al., 2016a).

All the studies described above highlight the strong relation between SUD and poor decision-making that could result from deficits in the goal-directed system associated with impaired inhibitory control. Importantly, motor impulsivity has been directly associated with deficits in goal-directed control, a correlation fitting well with the idea of an imbalance between goal-directed and habitual behavior; impulsive "habitual" actions are prompted by the situation rather than selected based on an evaluation of their predicted consequences, and deficits in response inhibition can disrupt an individual's ability to exert goal-directed control over this habitual responding.

Conclusion

Under non-pathological conditions, goal-directed and habitual processes operate in parallel and complement each other to control decision-making and performance (Figure 2A). Repetition and practice in stable situations promote the formation of habit, but in situations that warrant deliberation, such as changes in environmental contingency, individuals can re-exert goal-directed control over habit to flexibly adjust their responding. Thus, transitions between habit and goal-directed control are bidirectional and the control of behavior by these two systems can be balanced. This balance is disrupted in SUD as a result of abnormal formation of habit (Figure 2B), deficits in goal-directed behavior (Figure 2C), or a combination of both (Figure 2D). The main objective of this review was to tease apart these processes to determine their relative contribution to SUD. In this last section, we will consider each of these 3 possibilities.

Overreliance on the habit system in SUD (Figure 2B) is often suggested to explain the power of drug-associated cues to trigger drug seeking behaviors, as a consequence of S-R associations (Everitt and Robbins, 2016, 2005; Robinson and Berridge, 1993). However, it is important to highlight that evidence that habits rely on S-R associations are in most cases indirect. In addition, from a theoretical perspective, drug habits alone are not sufficient to explain compulsive drug seeking. Indeed, for an abnormal habit to persist in face of negative

drug consequences, drug addicts must also have difficulty re-exerting goal-directed control over their habitual drug seeking behavior (Ostlund and Balleine, 2008). In other words, drug addicts should be unable to overcome or inhibit habit in situations that warrant deliberation. This implies a deficit in goal-directed control and makes the first possibility unlikely (figure 2B).

The idea of a deficit in goal-directed control is appealing to explain the puzzling paradox of drug pursuit despite disastrous negative consequences, a hallmark of SUD. It would be interesting to demonstrate that drug addiction is associated with persistent habit under conditions that should encourage a transition to goal-directed performance, for example by showing, in reinforced sessions, persistent responding for a drug that became aversive after CTA-induced devaluation (Ostlund and Balleine, 2008). To our knowledge, this effect has not yet been demonstrated but persistent reward seeking after reversal or quinine adulteration in animals chronically exposed to cocaine (Calu et al., 2007; Stalnaker et al., 2009) or alcohol (Hopf et al., 2010; Lesscher et al., 2010) respectively, suggests a deficit in goal-directed control over habit and provides some support to this hypothesis.

In this conception, the imbalance between habit and goal-directed systems in SUD cannot only be explained by abnormal habit, but also results from deficits in goal-directed systems. Indeed, the evidence reviewed above suggests that goal-directed control is impaired after chronic drug exposure and in SUD, leaving us with the second and third possibilities (Figure 2C, and D). According to the “disrupted self-control” account, habit would not be abnormally strong but would be expressed more frequently as a result of impaired goal-directed and inhibitory control. Drug seeking responses are prompted by the situation rather than rejected based on an evaluation of their predicted consequences, because the goal-directed system is disrupted. According to this view, drug addicts would persist in habitual drug seeking despite negative consequences because they fail to predict and consider these negative outcomes, even though they know about their existence. In addition, when engaged in habitual drug seeking behavior, drug addicts would have difficulty re-exerting goal-directed control over their behavior because of failures in inhibitory control.

Whether or not the habit system is promoted during substance use disorder is a question difficult to answer. Since habit is defined as an S-R association but mainly operationalized as an absence of outcome representation, current models used to tackle this question cannot demonstrate the reliance on habit per se. For example, in the devaluation paradigm, insensitivity to outcome devaluation after drug exposure can result from the formation of habit, but may also reflect a deficit in goal-directed behavior. One study focused on formation and expression of S-R associations demonstrated that drug addicts learn new S-R associations as well as control subjects but are specifically impaired in overcoming well-learned S-R associations (McKim et al., 2016). These results suggest that formation of habits is not promoted in SUD but their expression is more resistant to changes in environmental contingencies. This inflexibility in habitual responding can result from a deficit in response inhibition (i.e. the ability to exert goal -directed control over existing habit). Another experiment in alcohol-dependent patients used a computational approach to separate model-based and model-free components of decision-making, and obtained findings

suggesting that chronic alcohol use selectively impairs goal-directed function with little effect on habit (Sebold et al., 2014).

Of note, exposure to drugs of abuse induces long lasting neurobiological changes in the neural circuitry proposed to underlie goal-directed and habitual responding. Exposure to drugs of abuse induces changes in several brain regions proposed to underlie habit such as the dorsolateral striatum, Substantia Nigra pars compacta, and the central amygdala (Gremel and Lovinger, 2016; Jedynak et al., 2007; Koob and Volkow, 2016; Lingawi and Balleine, 2012; Ron and Barak, 2016) and habitual responding in SUD is associated with hyperactivity of putamen, the human homologue of rodent dorsolateral striatum (de Wit et al., 2012; Dolan and Dayan, 2013). In addition, regions involved in goal-directed behavior, such as dorsomedial striatum, basolateral amygdala, OFC and prelimbic cortex (Balleine and O'Doherty, 2010; Hart et al., 2014) are also altered after chronic exposure to drugs of abuse (Chen et al., 2013; Jedynak et al., 2007; Koob and Volkow, 2016; Laguesse et al., 2016; Lucantonio et al., 2012; Nimitvilai et al., 2017; Ron and Barak, 2016; Wang et al., 2012; Wassum and Izquierdo, 2015). Converging evidence suggests that development of addictive behaviors is associated with a shift at the neural level from prefrontal cortical to striatal control as well as a progression from ventral to more dorsal parts of the striatum (Belin et al., 2009; Corbit et al., 2012; Corbit and Janak, 2016a; Everitt and Robbins, 2016, 2005). Although these neural transitions and the concomitant alterations in goal-directed and habit circuits support the 3rd possibility (figure 2D), further behavioral evidence is needed to dissociate between the 2nd and 3rd possibilities and determine whether formation of habits is promoted or not in SUD.

The development of additional animal models could facilitate tests for the involvement of habit in SUD. In an attempt to develop new models that probe the S-R association basis of habits, one suggestion is to place more emphasis on the influence of stimuli on instrumental responding (Corbit and Janak, 2016a). Although the role of context in the associative structure of habit has been recently demonstrated (Thrailkill and Bouton, 2015), the stimuli in instrumental procedures used to study habit in rodent remain poorly defined; demonstrating that behavior relies on S-R associations is therefore challenging. Providing salient, explicit and predictive stimuli in instrumental procedures could influence performance (Corbit and Janak, 2016b), facilitate the formation of habit, as assessed by outcome devaluation or contingency degradation, and could provide for additional measures to estimate the formation and expression of S-R associations, for example, the response latency from stimulus presentation (Keramati et al., 2011). Another related approach to demonstrate S-R habit would be to use reversal tasks to probe the S-R association, by changing the S-R contingency while preserving the R-O contingency and outcome value. Defining the measures specifically assessing the formation or expression of S-R association in reversal tasks will require further research, as reversal tasks also depend upon goal-directed control to adjust behavior to changing conditions (Stalnaker et al., 2009). Finally, adaptation to rodents of procedures used in humans to measure S-R responding and/or model-free vs model-based behavior, as reviewed earlier, could be a valuable avenue for exploration.

In summary, thus far, extensive research focused on the involvement of habit in SUD has demonstrated an imbalance between goal-directed and habitual control of drug seeking behavior. A growing body of research indicates impairments in goal-directed behavior system associated with impulsivity, compulsivity, and SUD, which presumably participate to tilt the balance toward habitual drug seeking. Whether abnormal formation of drug habits themselves also contribute to addictive behavior remains an open question that perhaps can be better addressed with new paradigms designed to assess habit as a positive result, i.e., the formation of S-R associations, rather than as an absence of outcome representation. While some efforts have been made to directly assess the formation and expression of S-R association in humans (de Wit et al., 2012; McKim et al., 2016), a refinement of our animal and human models of habit is needed to precisely define the place of habit in SUD and develop appropriate neurobehavioral treatments.

Acknowledgments

We thank Dr Ronald Keiflin for critical discussions and for his comments on the manuscript.

Funding

This work was supported by the National Institutes of Health grants DA035943, AA018025, AA014925.

References

- Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B.* 1982; 34:77–98. DOI: 10.1080/14640748208400878
- Adams CD, Dickinson A. Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B Comp. Physiol. Psychol.* 1981; 33:109–121. DOI: 10.1080/14640748108400816
- Ahmed SH. Validation crisis in animal models of drug addiction: Beyond non-disordered drug use toward drug addiction. *Neurosci. Biobehav. Rev.* 2010; 35:172–184. DOI: 10.1016/j.neubiorev.2010.04.005 [PubMed: 20417231]
- Bae SC, Lyoo IK, Sung YH, Yoo J, Chung A, Yoon SJ, Kim DJ, Hwang J, Kim SJ, Renshaw PF. Increased white matter hyperintensities in male methamphetamine abusers. *Drug Alcohol Depend.* 2006; 81:83–88. DOI: 10.1016/j.drugalcdep.2005.05.016 [PubMed: 16005161]
- Balconi M, Finocchiaro R, Campanella S. Reward sensitivity, decisional bias, and metacognitive deficits in cocaine drug addiction. *J Addict Med.* 2014; 8:399–406. DOI: 10.1097/adm.000000000000065 [PubMed: 25303980]
- Baler RD, Volkow ND. Drug addiction: the neurobiology of disrupted self-control. *Trends Mol Med.* 2006; 12:559–566. [PubMed: 17070107]
- Balleine BW, Dickinson A. Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology.* 1998; 37:407–419. DOI: 10.1016/S0028-3908(98)00033-1 [PubMed: 9704982]
- Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology.* 2010; 35:48–69. DOI: 10.1038/npp.2009.131 [PubMed: 19776734]
- Barker JM, Corbit LH, Robinson DL, Gremel CM, Gonzales RA, Chandler LJ. Corticostriatal circuitry and habitual ethanol seeking. *Alcohol.* 2015; 49:817–824. [PubMed: 26059221]
- Bechara A, Dolan S, Denburg N, Hinds A, Anderson SW, Nathan PE. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia.* 2001; 39:376–389. DOI: 10.1016/S0028-3932(00)00136-6 [PubMed: 11164876]
- Bechara A, Dolan S, Hinds A. Decision-making and addiction (part I): Impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future

- consequences. *Neuropsychologia*. 2002; 40:1675–1689. DOI: 10.1016/S0028-3932(02)00015-5 [PubMed: 11992656]
- Bechara A, Martin EM. Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology*. 2004; 18:152–62. DOI: 10.1037/0894-4105.18.1.152 [PubMed: 14744198]
- Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. *Behav Brain Res*. 2009; 199:89–102. [PubMed: 18950658]
- Berridge KC, Robinson TE. Liking, Wanting, and the Incentive-Sensitization Theory of Addiction. *Am. Psychol*. 2016; 71:670–679. DOI: 10.1037/amp0000059 [PubMed: 27977239]
- Bickel WK, Jarmolowicz DP, Mueller ET, Koffarnus MN, Gatchalian KM. Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: Emerging evidence. *Pharmacol. Ther*. 2012; 134:287–297. DOI: 10.1016/j.pharmthera.2012.02.004 [PubMed: 22387232]
- Brauer LH, Behm FM, Lane JD, Westman EC, Perkins C, Rose JE. Individual differences in smoking reward from de-nicotinized cigarettes. *Nicotine Tob Res*. 2001; 3:101–109. [PubMed: 11403723]
- Bühler M, Mann K. Alcohol and the human brain: A systematic review of different neuroimaging methods. *Alcohol. Clin. Exp. Res*. 2011; 35:1771–1793. DOI: 10.1111/j.1530-0277.2011.01540.x [PubMed: 21777260]
- Caggiula AR, Donny EC, White AR, Chaudhri N, Booth S, Gharib MA, Hoffman A, Perkins KA, Sved AF. Cue dependency of nicotine self-administration and smoking. *Pharmacol Biochem Behav*. 2001; 70:515–530. [PubMed: 11796151]
- Calu DJ, Stalnaker TA, Franz TM, Singh T, Shaham Y, Schoenbaum G. Withdrawal from cocaine self-administration produces long-lasting deficits in orbitofrontal-dependent reversal learning in rats. *Learn. Mem*. 2007; 14:325–328. DOI: 10.1101/lm.534807 [PubMed: 17522022]
- Camchong J, MacDonald AW, Nelson B, Bell C, Mueller BA, Specker S, Lim KO. Frontal hyperconnectivity related to discounting and reversal learning in cocaine subjects. *Biol. Psychiatry*. 2011; 69:1117–1123. DOI: 10.1016/j.biopsych.2011.01.008 [PubMed: 21371689]
- Chaudhri N, Caggiula AR, Donny EC, Booth S, Gharib M, Craven L, Palmatier MI, Liu X, Sved AF. Operant responding for conditioned and unconditioned reinforcers in rats is differentially enhanced by the primary reinforcing and reinforcement-enhancing effects of nicotine. *Psychopharmacol*. 2006; 189:27–36.
- Chen BT, Yau HJ, Hatch C, Kusumoto-Yoshida I, Cho SL, Hopf FW, Bonci A. Rescuing cocaine-induced prefrontal cortex hypoactivity prevents compulsive cocaine seeking. *Nature*. 2013; 496:359–362. DOI: 10.1038/nature12024 [PubMed: 23552889]
- Chiamulera C. Cue reactivity in nicotine and tobacco dependence: a “multiple-action” model of nicotine as a primary reinforcement and as an enhancer of the effects of smoking-associated stimuli. *Brain Res Brain Res Rev*. 2005; 48:74–97. [PubMed: 15708629]
- Ciccocioppo R, Martin-Fardon R, Weiss F. Stimuli associated with a single cocaine experience elicit long-lasting cocaine-seeking. *Nat Neurosci*. 2004; 7:495–496. DOI: 10.1038/nn1219 [PubMed: 15048121]
- Clemens KJ, Castino MR, Cornish JL, Goodchild AK, Holmes NM. Behavioral and neural substrates of habit formation in rats intravenously self-administering nicotine. *Neuropsychopharmacology*. 2014; 39:2584–2593. DOI: 10.1038/npp.2014.111 [PubMed: 24823947]
- Colwill RM. An associative analysis of instrumental learning. *Curr. Dir. Psychol. Sci*. 1993; 2:111–116.
- Colwill RM, Rescorla RA. Instrumental responding remains sensitive to reinforcer devaluation after extensive training. *J. Exp. Psychol. Anim. Behav. Process*. 1985; 11:520–536. DOI: 10.1037/0097-7403.11.4.520
- Colwill RM, Triola SM. Instrumental responding remains under the control of the consequent outcome after extended training. *Behav. Processes*. 2002; 57:51–64. DOI: 10.1016/S0376-6357(01)00204-2 [PubMed: 11864775]

- Corbit LH, Chieng BC, Balleine BW. Effects of repeated cocaine exposure on habit learning and reversal by N-acetylcysteine. *Neuropsychopharmacology*. 2014; 39:1893–1901. [PubMed: 24531561]
- Corbit LH, Janak PH. Habitual Alcohol Seeking: Neural Bases and Possible Relations to Alcohol Use Disorders. *Alcohol Clin Exp Res*. 2016a; 40:1380–1389. [PubMed: 27223341]
- Corbit LH, Janak PH. Changes in the Influence of Alcohol-Paired Stimuli on Alcohol Seeking across Extended Training. *Front Psychiatry*. 2016b; 7:169. [PubMed: 27777560]
- Corbit LH, Janak PH. Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcohol Clin Exp Res*. 2007; 31:766–774. [PubMed: 17378919]
- Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. *Biol Psychiatry*. 2012; 72:389–395. [PubMed: 22440617]
- Crombag HS, Bossert JM, Koya E, Shaham Y. Context-induced relapse to drug seeking: a review. *Philos. Trans. R. Soc. B Biol. Sci.* 2008; 363:3233–3243. DOI: 10.1098/rstb.2008.0090
- Cushman F, Morris A. Habitual control of goal selection in humans. *Proc. Natl. Acad. Sci.* 2015; 112:13817–13822. DOI: 10.1073/pnas.1506367112 [PubMed: 26460050]
- Dalley JW, Everitt BJ, Robbins TW. Impulsivity, compulsivity, and top-down cognitive control. *Neuron*. 2011; 69:680–694. [PubMed: 21338879]
- Dalley JW, Fryer TD, Brichard L, Robinson ESJ, Theobald DEH, Laane K, Pena Y, Murphy ER, Shah Y, Probst K, Abakumova I, Aigbirhio FI, Richards HK, Hong Y, Baron J-C, Everitt BJ, Robbins TW. Nucleus Accumbens D2/3 Receptors Predict Trait Impulsivity and Cocaine Reinforcement. *Science (80-.)*. 2007; 315:1267–1270. DOI: 10.1126/science.1137073
- Daw ND, Doya K. The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 2006; 16:199–204. DOI: 10.1016/j.conb.2006.03.006 [PubMed: 16563737]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011; 69:1204–1215. DOI: 10.1016/j.neuron.2011.02.027 [PubMed: 21435563]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 2005; 8:1704–11. DOI: 10.1038/nn1560 [PubMed: 16286932]
- de Wit S, Corlett PR, Aitken MR, Dickinson A, Fletcher PC. Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 2009; 29:11330–8. DOI: 10.1523/JNEUROSCI.1639-09.2009 [PubMed: 19741139]
- de Wit S, Watson P, Harsay Ha, Cohen MX, van de Vijver I, Ridderinkhof KR. Corticostriatal Connectivity Underlies Individual Differences in the Balance between Habitual and Goal-Directed Action Control. *J. Neurosci.* 2012; 32:12066–75. DOI: 10.1523/JNEUROSCI.1088-12.2012 [PubMed: 22933790]
- DePoy LM, Gourley SL. Synaptic Cytoskeletal Plasticity in the Prefrontal Cortex Following Psychostimulant Exposure. *Traffic*. 2015; 16:919–940. [PubMed: 25951902]
- Derusso AL, Fan D, Gupta J, Shelest O, Costa RM, Yin HH. Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Front. Integr. Neurosci.* 2010; 4:1–8. DOI: 10.3389/fnint.2010.00017 [PubMed: 20161992]
- Dezfouli A, Balleine BW. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput Biol*. 2013; 9:e1003364. [PubMed: 24339762]
- Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. *Eur J Neurosci*. 2012 Apr; 35(7):1036–51. 2012. DOI: 10.1111/j.1460-9568.2012.08050.x [PubMed: 22487034]
- Dezfouli A, Lingawi NW, Balleine BW. Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc L. B Biol Sci.* 2014; 369
- Dias-ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, Sousa N. Chronic Stress Causes Frontostriatal Reorganization and Affects Decision-Making. 2009; 325:621–625.
- Dickinson A. Instrumental conditioning. In: Mackintosh NJ, editor *Animal Learning and Cognition*. Academic Press; San Diego CA: 1994. 45–79.

- Dickinson A. Expectancy theory in animal conditioning. In: Klein SB, Mowrer RR, editors *Contemporary Learning Theories*. Psychology Press; 1989.
- Dickinson A. Actions and Habits: The Development of Behavioural Autonomy. *Philos. Trans. R. Soc. B Biol. Sci.* 1985; doi: 10.1098/rstb.1985.0010
- Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. *Anim. Learn. Behav.* 1995; 23:197–206. DOI: 10.3758/BF03199935
- Dickinson A, Nicholas DJ, Adams CD. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B.* 1983; 35:35–51. DOI: 10.1080/14640748308400912
- Dickinson A, Wood N, Smith JW. Alcohol seeking by rats: action or habit? *Q. J. Exp. Psychol. B.* 2002; 55:331–48. DOI: 10.1080/0272499024400016 [PubMed: 12350285]
- Diergaarde L, Pattij T, Poortvliet I, Hogenboom F, de Vries W, Schoffelmeer AN, De Vries TJ. Impulsive choice and impulsive action predict vulnerability to distinct stages of nicotine seeking in rats. *Biol Psychiatry.* 2008; 63:301–308. [PubMed: 17884016]
- Dolan RJ, Dayan P. Goals and habits in the brain. *Neuron.* 2013; 80:312–325. DOI: 10.1016/j.neuron.2013.09.007 [PubMed: 24139036]
- Donny EC, Chaudhri N, Caggiula AR, Evans-Martin FF, Booth S, Gharib MA, Clements LA, Sved AF. Operant responding for a visual reinforcer in rats is enhanced by noncontingent nicotine: implications for nicotine self-administration and reinforcement. *Psychopharmacol.* 2003; 169:68–76.
- Doya K, Samejima K, Katagiri K, Kawato M. Multiple Model-Based Reinforcement Learning. *Neural Comput.* 2002; 14:1347–1369. DOI: 10.1162/089976602753712972 [PubMed: 12020450]
- Economidou D, Pelloux Y, Robbins TW, Dalley JW, Everitt BJ. High impulsivity predicts relapse to cocaine-seeking after punishment-induced abstinence. *Biol Psychiatry.* 2009; 65:851–856. [PubMed: 19181308]
- Ersche KD, Barnes A, Jones PS, Morein-Zamir S, Robbins TW, Bullmore ET. Abnormal structure of frontostriatal brain systems is associated with aspects of impulsivity and compulsivity in cocaine dependence. *Brain.* 2011; 134:2013–2024. [PubMed: 21690575]
- Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LH, Luijten M, de Wit S, Sahakian BJ, Bullmore ET, Robbins TW. Carrots and sticks fail to change behavior in cocaine addiction. *Science (80-.)*. 2016; 352:1468–1471.
- Ersche KD, Jones PS, Williams GB, Turton AJ, Robbins TW, Bullmore ET. Abnormal brain structure implicated in stimulant drug addiction. *Science (80-.)*. 2012; 335:601–604.
- Everitt BJ, Robbins TW. Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. *Annu Rev Psychol.* 2016; 67:23–50. [PubMed: 26253543]
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci.* 2005; 8:1481–1489. [PubMed: 16251991]
- Franklin TR, Acton PD, Maldjian JA, Gray JD, Croft JR, Dackis CA, O'Brien CP, Childress AR. Decreased gray matter concentration in the insular, orbitofrontal, cingulate, and temporal cortices of cocaine patients. *Biol. Psychiatry.* 2002; 51:134–142. [PubMed: 11822992]
- Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife.* 2016a; 5
- Gillan CM, Robbins TW, Sahakian BJ, van den Heuvel OA, van Wingen G. The role of habit in compulsivity. *Eur. Neuropsychopharmacol.* 2016b; 26:828–840. DOI: 10.1016/j.euroneuro.2015.12.033 [PubMed: 26774661]
- Gläscher J, Daw N, Dayan P, O'doherty JP. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron.* 2010; 66:585–595. DOI: 10.1016/j.neuron.2010.04.016 [PubMed: 20510862]
- Goldstein RZ, Volkow ND. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nat. Rev. Neurosci.* 2012; 12:652–669. DOI: 10.1038/nrn3119.Dysfunction
- Gourley SL, Olevska A, Gordon J, Taylor JR. Cytoskeletal Determinants of Stimulus-Response Habits. *J. Neurosci.* 2013; 33:11811–11816. DOI: 10.1523/JNEUROSCI.1034-13.2013 [PubMed: 23864670]

- Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, Costa RM. Endocannabinoid Modulation of Orbitofrontal Circuits Gates Habit Formation. *Neuron*. 2016; 90:1312–1324. [PubMed: 27238866]
- Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun*. 2013; 4:2264. [PubMed: 23921250]
- Gremel CM, Lovinger DM. Associative and sensorimotor cortico-basal ganglia circuit roles in effects of abused drugs. *Genes, Brain Behav*. 2016; :71–85. DOI: 10.1111/gbb.12309 [PubMed: 27457495]
- Grimm JW, Hope BT, Wise RA, Shaham Y. Neuroadaptation. Incubation of cocaine craving after withdrawal. *Nature*. 2001; 412:141–2. DOI: 10.1038/35084134 [PubMed: 11449260]
- Hart G, Leung BK, Balleine BW. Dorsal and ventral streams: The distinct role of striatal subregions in the acquisition and performance of goal-directed actions. *Neurobiol. Learn. Mem*. 2014; 108:104–118. DOI: 10.1016/j.nlm.2013.11.003 [PubMed: 24231424]
- Heyman GM. *Addiction: A Disorder of Choice*. Harvard University Press; 2010.
- Hinton EA, Wheeler MG, Gourley SL. Early-life cocaine interferes with BDNF-mediated behavioral plasticity. *Learn Mem*. 2014; 21:253–257. DOI: 10.1101/lm.033290.113 [PubMed: 24737916]
- Hogarth L, Chase HW. Parallel goal-directed and habitual control of human drug-seeking: implications for dependence vulnerability. *J Exp Psychol Anim Behav Process*. 2011; 37:261–276. [PubMed: 21500933]
- Hogarth L, Chase HW, Baess K. Impaired goal-directed behavioural control in human impulsivity. *Q. J. Exp. Psychol*. 2012; 65:305–316.
- Holland PC. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process*. 2004; 30:104–17. DOI: 10.1037/0097-7403.30.2.104 [PubMed: 15078120]
- Holmes NM, Clemens KJ. Multiple Interpretations of Cocaine-Seeking Behavior after Prolonged Self-Administration Training. *J. Neurosci*. 2011; 31:3935–3936. DOI: 10.1523/JNEUROSCI.6354-10.2011 [PubMed: 21411635]
- Hopf FW, Chang SJ, Sparta DR, Bowers MS, Bonci A. Motivation for alcohol becomes resistant to quinine adulteration after 3 to 4 months of intermittent alcohol self-administration. *Alcohol. Clin. Exp. Res*. 2010; 34:1565–1573. DOI: 10.1111/j.1530-0277.2010.01241.x [PubMed: 20586757]
- Irimia C, Polis IY, Stouffer D, Parsons LH. Persistent effects of chronic 9-THC exposure on motor impulsivity in rats. *Psychopharmacology (Berl)*. 2015a; 232:3033–3043. DOI: 10.1007/s00213-015-3942-x [PubMed: 25925779]
- Irimia C, Wiskerke J, Natividad LA, Polis IY, De Vries TJ, Pattij T, Parsons LH. Increased impulsivity in rats as a result of repeated cycles of alcohol intoxication and abstinence. *Addict. Biol*. 2015b; 20:263–274. DOI: 10.1111/adb.12119 [PubMed: 24341858]
- Jedynak JP, Uslander JM, Esteban JA, Robinson TE. Methamphetamine-induced structural plasticity in the dorsal striatum. *Eur. J. Neurosci*. 2007; 25:847–853. DOI: 10.1111/j.1460-9568.2007.05316.x [PubMed: 17328779]
- Jentsch JD, Taylor JR. Impulsivity resulting from frontostriatal dysfunction in drug abuse: Implications for the control of reward-related stimuli. *Psychopharmacology (Berl)*. 1999; 146:373–390. [PubMed: 10550488]
- Kearns DN, Gomez-Serrano MA, Tunstall BJ. A Review of Preclinical Research Demonstrating that Drug and Non-Drug Reinforcers Differentially Affect Behavior. *Curr. Drug Abus. Rev*. 2011; 4:261–269. DOI: 10.2174/1874473711104040261
- Keiflin R, Janak PH. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*. 2015; 88:247–263. DOI: 10.1016/j.neuron.2015.08.037 [PubMed: 26494275]
- Keramati M, Dezfouli A, Piray P. Speed/Accuracy Trade-Off between the Habitual and the Goal-Directed Processes. *PLoS Comput. Biol*. 2011; 7:e1002055. doi: 10.1371/journal.pcbi.1002055 [PubMed: 21637741]
- Keramati M, Gutkin B. Imbalanced Decision Hierarchy in Addicts Emerging from Drug-Hijacked Dopamine Spiraling Circuit. *PLoS One*. 2013; :8. doi: 10.1371/journal.pone.0061489

- Kirby KN, Petry NM, Bickel WK. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J. Exp. Psychol. Gen.* 1999; 128:78–87. DOI: 10.1037/0096-3445.128.1.78 [PubMed: 10100392]
- Koob GF, Volkow ND. Neurobiology of addiction: a neurocircuitry analysis. *The Lancet Psychiatry.* 2016; 3:760–773. DOI: 10.1016/S2215-0366(16)00104-8 [PubMed: 27475769]
- Laguesse S, Morisot N, Phamluong K, Ron D. Region specific activation of the AKT and mTORC1 pathway in response to excessive alcohol intake in rodents. *Addict. Biol.* 2016; doi: 10.1111/adb.12464
- LeBlanc KH, Maidment NT, Ostlund SB. Repeated Cocaine Exposure Facilitates the Expression of Incentive Motivation and Induces Habitual Control in Rats. *PLoS One.* 2013; 8doi: 10.1371/journal.pone.0061355
- Lesscher HMB, Van Kerkhof LWM, Vanderschuren LJMJ. Inflexible and indifferent alcohol drinking in male mice. *Alcohol. Clin. Exp. Res.* 2010; 34:1219–1225. DOI: 10.1111/j.1530-0277.2010.01199.x [PubMed: 20477770]
- Lingawi NW, Balleine BW. Amygdala Central Nucleus Interacts with Dorsolateral Striatum to Regulate the Acquisition of Habits. *J. Neurosci.* 2012; 32:1073–1081. DOI: 10.1523/JNEUROSCI.4806-11.2012.Amygdala [PubMed: 22262905]
- London ED, Cascella NG, Wong DF, Phillips RL, Dannals RF, Links JM, Herning R, Grayson R, Jaffe JH, Wagner HN. Cocaine-induced reduction of glucose utilization in human brain. A study using positron emission tomography and [fluorine 18]-fluorodeoxyglucose. *Arch. Gen. Psychiatry.* 1990; 47:567–74. [PubMed: 2350209]
- Lopez MF, Becker HC, Chandler LJ. Repeated episodes of chronic intermittent ethanol promote insensitivity to devaluation of the reinforcing effect of ethanol. *Alcohol.* 2014; 48:639–645. DOI: 10.1016/j.alcohol.2014.09.002 [PubMed: 25266936]
- Loughlin A, Funk D, Coen KLAD. Habitual nicotine-seeking in rats following limited training. *Psychopharmacology (Berl).* 2017; :1–11. DOI: 10.1007/s00213-017-4655-0
- Lu L, Grimm JW, Hope BT, Shaham Y. Incubation of cocaine craving after withdrawal: a review of preclinical data. *Neuropharmacology.* 2004; 47:214–226. DOI: 10.1016/j.neuropharm.2004.06.027 [PubMed: 15464139]
- Lucantonio F, Stalnaker TA, Shaham Y, Niv Y, Schoenbaum G. The impact of orbitofrontal dysfunction on cocaine addiction. *Nat Neurosci.* 2012; 15:358–366. [PubMed: 22267164]
- Lucantonio F, Takahashi YK, Hoffman AF, Chang CY, Bali-Chaudhary S, Shaham Y, Lupica CR, Schoenbaum G. Orbitofrontal activation restores insight lost after cocaine use. *Nat. Neurosci.* 2014; 17:1092–1099. DOI: 10.1038/nn.3763 [PubMed: 25042581]
- Lyoo IK, Streeter CC, Ahn KH, Lee HK, Pollack MH, Silveri MM, Nassar L, Levin JM, Sarid-Segal O, Ciraulo DA, Renshaw PF, Kaufman MJ. White matter hyperintensities in subjects with cocaine and opiate dependence and healthy comparison subjects. *Psychiatry Res. Neuroimaging.* 2004; 131:135–145.
- Madden GJ, Petry NM, Badger GJ, Bickel WK. Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: drug and monetary rewards. *Exp. Clin. Psychopharmacol.* 1997; 5:256–262. DOI: 10.1037/1064-1297.5.3.256 [PubMed: 9260073]
- Mangieri RA, Cofresí RU, Gonzales RA. Ethanol seeking by long evans rats is not always a goal-directed behavior. *PLoS One.* 2012; 7:1–13. DOI: 10.1371/journal.pone.0042886
- Mangieri RaCofresí RU, Gonzales Ra. Ethanol exposure interacts with training conditions to influence behavioral adaptation to a negative instrumental contingency. *Front. Behav. Neurosci.* 2014; 8:220.doi: 10.3389/fnbeh.2014.00220 [PubMed: 24987342]
- McGuier NS, Padula AE, Lopez MF, Woodward JJ, Mulholland PJ. Withdrawal from chronic intermittent alcohol exposure increases dendritic spine density in the lateral orbitofrontal cortex of mice. *Alcohol.* 2015; 49:21–27. DOI: 10.1016/j.alcohol.2014.07.017 [PubMed: 25468278]
- McKim TH, Bauer DJ, Boettiger CA. Addiction History Associates with the Propensity to Form Habits. *J. Cogn. Neurosci.* 2016; 28:1024–38. DOI: 10.1162/jocn_a_00953 [PubMed: 26967944]
- Miles FJ, Everitt BJ, Dickinson A. Oral cocaine seeking by rats: action or habit? *Behav. Neurosci.* 2003; 117:927–38. DOI: 10.1037/0735-7044.117.5.927 [PubMed: 14570543]

- Monterosso JR, Aron AR, Cordova X, Xu J, London ED. Deficits in response inhibition associated with chronic methamphetamine abuse. *Drug Alcohol Depend.* 2005; 79:273–277. DOI: 10.1016/j.drugalcdep.2005.02.002 [PubMed: 15967595]
- Morein-Zamir S, Robbins TW. Fronto-striatal circuits in response-inhibition: Relevance to addiction. *Brain Res.* 2015; 1628:117–129. [PubMed: 25218611]
- Morein-Zamir S, Simon Jones P, Bullmore ET, Robbins TW, Ersche KD. Prefrontal hypoactivity associated with impaired inhibition in stimulant-dependent individuals but evidence for hyperactivation in their unaffected siblings. *Neuropsychopharmacology.* 2013; 38:1945–1953. DOI: 10.1038/npp.2013.90 [PubMed: 23609131]
- Morris LS, Dowell NG, Cercignani M, Harrison NA, Voon V. Binge drinking differentially affects cortical and subcortical microstructure. *Addict. Biol.* 2017; doi: 10.1111/adb.12493
- Nelson A, Killcross S. Amphetamine Exposure Enhances Habit Formation. *J. Neurosci.* 2006; 26:3805–3812. DOI: 10.1523/JNEUROSCI.4305-05.2006 [PubMed: 16597734]
- Nimitvilai S, Lopez MF, Mulholland PJ, Woodward JJ. Chronic Intermittent Ethanol Exposure Enhances the Excitability and Synaptic Plasticity of Lateral Orbitofrontal Cortex Neurons and Induces a Tolerance to the Acute Inhibitory Actions of Ethanol. *Neuropsychopharmacology.* 2016; 41:1112–27. DOI: 10.1038/npp.2015.250 [PubMed: 26286839]
- Nimitvilai S, Uys JD, Woodward JJ, Randall PK, Ball LE, Williams RW, Jones BC, Lu L, Grant KA, Mulholland PJ. Orbitofrontal neuroadaptations and cross-species synaptic biomarkers in heavy drinking macaques. *J. Neurosci.* 2017; 37:133–17. DOI: 10.1523/JNEUROSCI.0133-17.2017
- Noël X, Van der Linden M, d'Acremont M, Bechara A, Dan B, Hanak C, Verbanck P. Alcohol cues increase cognitive impulsivity in individuals with alcoholism. *Psychopharmacology (Berl).* 2007; 192:291–298. DOI: 10.1007/s00213-006-0695-6 [PubMed: 17279375]
- Nordquist RE, Voon P, de Mooij-van Malsen JG, Joosten RNJMA, Pennartz CMA, Vanderschuren LJMJ. Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *Eur. Neuropsychopharmacol.* 2007; 17:532–540. DOI: 10.1016/j.euroneuro.2006.12.005 [PubMed: 17275266]
- Ostlund SB, Balleine BW. On habits and addiction: An associative analysis of compulsive drug seeking. *Drug Discov Today Dis Model.* 2008; 5:235–245.
- Ostlund SB, Maidment NT, Balleine BW. Alcohol-Paired Contextual Cues Produce an Immediate and Selective Loss of Goal-directed Action in Rats. *Front. Integr. Neurosci.* 2010; 4
- Otto AR, Gershman SJ, Markman AB, Daw ND. The Curse of Planning. *Psychol. Sci.* 2013; 24:751–761. DOI: 10.1177/0956797612463080 [PubMed: 23558545]
- Palmatier MI, Evans-Martin FF, Hoffman A, Caggiula AR, Chaudhri N, Donny EC, Liu X, Booth S, Gharib M, Craven L, Sved AF. Dissociating the primary reinforcing and reinforcement-enhancing effects of nicotine using a rat self-administration paradigm with concurrently available drug and environmental reinforcers. *Psychopharmacol.* 2006; 184:391–400.
- Pickard H, Ahmed SH. How do you know you have a drug problem? The role of knowledge of negative consequences in explaining drug choice in humans and rats. In: Heather N, Segal G, editors *Addiction and Choice*. Oxford University press; 2015.
- Pierce RC, Vanderschuren LJMJ. Kicking the habit: The neural basis of ingrained behaviors in cocaine addiction. *Neurosci. Biobehav. Rev.* 2010; 35:212–219. DOI: 10.1016/j.neubiorev.2010.01.007 [PubMed: 20097224]
- Rescorla RA. A Pavlovian analysis of goal-directed behavior. *Am. Psychol.* 1987; 42:119–129. DOI: 10.1037/0003-066X.42.2.119
- Robbins TW, Everitt BJ. Drug addiction: bad habits add up. *Nature.* 1999; 398:567–570. [PubMed: 10217139]
- Robinson TE, Berridge KC. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* 1993; 18:247–91. [PubMed: 8401595]
- Ron D, Barak S. Molecular mechanisms underlying alcohol-drinking behaviours. *Nat. Publ. Gr.* 2016; 17:576–591. DOI: 10.1038/nrn.2016.85
- Schmitzer-Torbert N, Apostolidis S, Amoa R, O'Rear C, Kaster M, Stowers J, Ritz R. Post-training cocaine administration facilitates habit learning and requires the infralimbic cortex and

- dorsolateral striatum. *Neurobiol. Learn. Mem.* 2015; 118:105–112. DOI: 10.1016/j.nlm.2014.11.007 [PubMed: 25460040]
- Schoenbaum G, Chang CY, Lucantonio F, Takahashi YK. Thinking Outside the Box: Orbitofrontal Cortex, Imagination, and How We Can Treat Addiction. *Neuropsychopharmacology.* 2016; 41:2966–2976. [PubMed: 27510424]
- Schoenbaum G, Takahashi Y, Liu TL, Mcdannald MA. Does the orbitofrontal cortex signal value? *Ann. N. Y. Acad. Sci.* 2011; 1239:87–99. DOI: 10.1111/j.1749-6632.2011.06210.x [PubMed: 22145878]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science (80-.)*. 1997; 275:1593–1599. DOI: 10.1126/science.275.5306.1593
- Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hägele C, Keller J, Jünger E, Kathmann N, Smolka M, Rapp MA, Schlagenhauf F, Heinz A, Huys QJM. Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology.* 2014; 70:122–131. DOI: 10.1159/000362840 [PubMed: 25359492]
- Sharpe MJ, Chang CY, Liu MA, Batchelor HM, Mueller LE, Jones JL, Niv Y, Schoenbaum G. Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nat. Neurosci.* 2017; :1–10. DOI: 10.1038/nn.4538
- Sjoerds Z, de Wit S, van den Brink W, Robbins TW, Beekman AT, Penninx BW, Veltman DJ. Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Transl Psychiatry.* 2013; 3:e337.doi: 10.1038/tp.2013.107 [PubMed: 24346135]
- Stalnaker TA, Takahashi Y, Roesch MR, Schoenbaum G. Neural substrates of cognitive inflexibility after chronic cocaine exposure. *Neuropharmacology.* 2009; 56:63–72. DOI: 10.1016/j.neuropharm.2008.07.019 [PubMed: 18692512]
- Sutton RS. Learning to predict by the methods of temporal differences. *Mach. Learn.* 1988; 3:9–44.
- Takahashi YK, Roesch MR, Stalnaker TA, Haney RZ, Calu DJ, Taylor AR, Burke KA, Schoenbaum G. The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron.* 2009; 62:269–280. DOI: 10.1016/j.neuron.2009.03.005 [PubMed: 19409271]
- Thraill EA, Bouton ME. Contextual control of instrumental actions and habits. *J. Exp. Psychol. Anim. Learn. Cogn.* 2015; 41:69–80. DOI: 10.1037/xan0000045 [PubMed: 25706547]
- Tiffany ST. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. *Psychol. Rev.* 1990; 97:147–168. DOI: 10.1037/0033-295X.97.2.147 [PubMed: 2186423]
- Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 2009; 29:2225–2232. DOI: 10.1111/j.1460-9568.2009.06796.x [PubMed: 19490086]
- Valentin VV, Dickinson A, O'Doherty JP. Determining the Neural Substrates of Goal-Directed Learning in the Human Brain. *J. Neurosci.* 2007; 27
- Volkow ND, Fowler JS, Wang G-J, Hitzemann R, Logan J, Schlyer DJ, Dewey SL, Wolf AP. Decreased dopamine D2 receptor availability is associated with reduced frontal metabolism in cocaine abusers. *Synapse.* 1993; 14:169–177. DOI: 10.1002/syn.890140210 [PubMed: 8101394]
- Volkow ND, Fowler JS, Wolf AP, Gillespi H. Metabolic studies of drugs of abuse. *NIDA Res. Monogr.* 1990; 105:47–53. [PubMed: 1876084]
- Volkow ND, Fowler JS, Wolf AP, Hitzemann R, Dewey S, Bendriem B, Alpert R, Hoff A. Changes in brain glucose metabolism in cocaine dependence and withdrawal. *Am. J. Psychiatry.* 1991; 148:621–626. DOI: 10.1176/ajp.148.5.621 [PubMed: 2018164]
- Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, Schreiber LRN, Gillan CM, Fineberg NA, Sahakian BJ, Robbins TW, Harrison NA, Wood J, Daw ND, Dayan P, Grant JE, Bullmore ET. Disorders of compulsivity: a common bias towards learning habits. *Mol. Psychiatry.* 2015; 20:345–352. DOI: 10.1038/mp.2014.44 [PubMed: 24840709]
- Walker SE, Peña-Oliver Y, Stephens DN. Learning not to be impulsive: Disruption by experience of alcohol withdrawal. *Psychopharmacology (Berl).* 2011; 217:433–442. DOI: 10.1007/s00213-011-2298-0 [PubMed: 21509502]

- Wang J, Ben Hamida S, Darcq E, Zhu W, Gibb SL, Lanfranco MF, Carnicella S, Ron D. Ethanol-mediated facilitation of AMPA receptor function in the dorsomedial striatum: implications for alcohol drinking behavior. *J. Neurosci.* 2012; 32:15124–32. DOI: 10.1523/JNEUROSCI.2783-12.2012 [PubMed: 23100433]
- Wassum KM, Izquierdo A. The basolateral amygdala in reward learning and addiction. *Neurosci. Biobehav. Rev.* 2015; 57:271–283. DOI: 10.1016/j.neubiorev.2015.08.017 [PubMed: 26341938]
- Zapata A, Minney VL, Shippenberg TS. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *J. Neurosci.* 2010; 30:15457–63. DOI: 10.1523/JNEUROSCI.4072-10.2010 [PubMed: 21084602]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Habits are generally defined as a stimulus-response association
- Habits are operationalized as an absence of goal-directed behavior
- Addictions result from an imbalance between goal-directed and habitual control
- Addictions are associated with deficits in goal -directed control
- Whether abnormal habits contribute to addictive behavior remains an open question

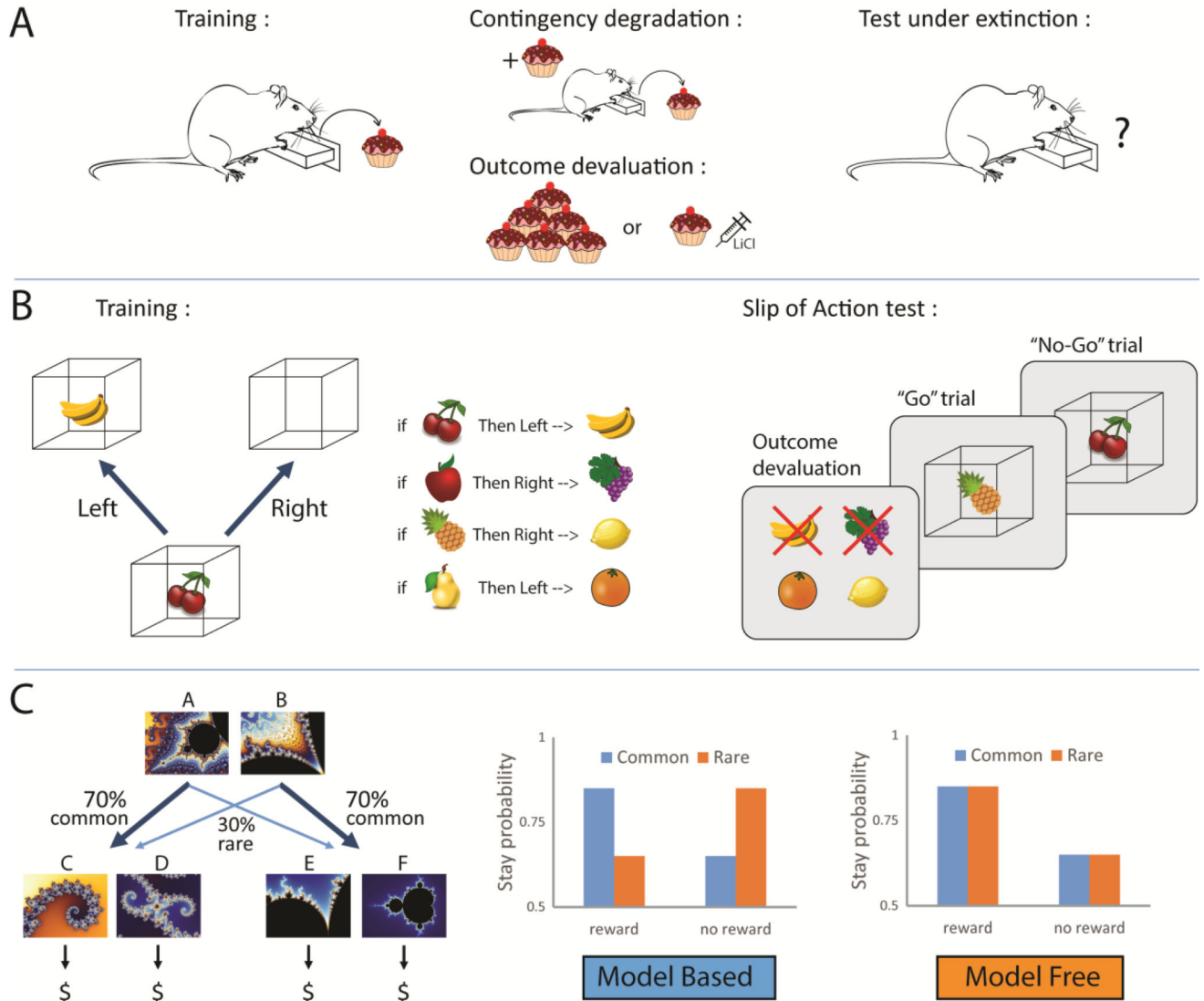
**Figure 1.**

Illustration of behavioral paradigms used to assess whether control of performance is goal-directed or habitual. A. In instrumental tasks, rodents are trained to make an action (i.e. lever press) to obtain a reward. During the test, the instrumental contingency between the response and the outcome can be degraded with non-contingent reward delivery (upper panel). It is also possible to devalue the outcome by giving ad libitum access to the reward to generate sensory-specific satiety, or by pairing reward consumption with lithium chloride injection to generate conditioned taste aversion (lower panel). After contingency degradation or outcome devaluation, performance is assessed under extinction. A reduction in responding demonstrates that performance is goal-directed while persistent responding indicates habitual behavior. B. In some human instrumental tasks, fruit pictures (the stimuli) are presented and signal which associated response (a left or right key press) earns points signaled by a subsequent picture of fruit inside a box (the outcome). If the wrong response is emitted, the box is empty. Some outcomes are then devalued (indicated by a red cross on the

fruit picture), and will now lead to a subtraction of points. During the test (right panel), subjects are presented with a rapid succession of stimuli, and are instructed to make the correct response for stimuli signaling still valuable outcomes (“go” trial), while they should refrain from responding for stimuli associated with devalued outcomes (“no go” trial). Adapted from (de Wit et al., 2012). C. In the sequential 2-stage decision task, one choice between 2 options at the first step leads commonly (70%) to a second pair of options but results occasionally (30%) in another set of options. At this second step, the selection of an option is rewarded or not, according to variable and unpredictable probabilities (left panel). Pure model-based agents are more likely to repeat the 1st step choice (i.e. ‘stay’) following rewarded trials after common transitions, but will switch to the alternative 1st step option after rare transitions (middle panel). In contrast, pure model-free agents will tend to repeat the same 1st step choice after rewarded trials, irrespective of the transition that preceded reward (right panel)(Daw et al., 2011).

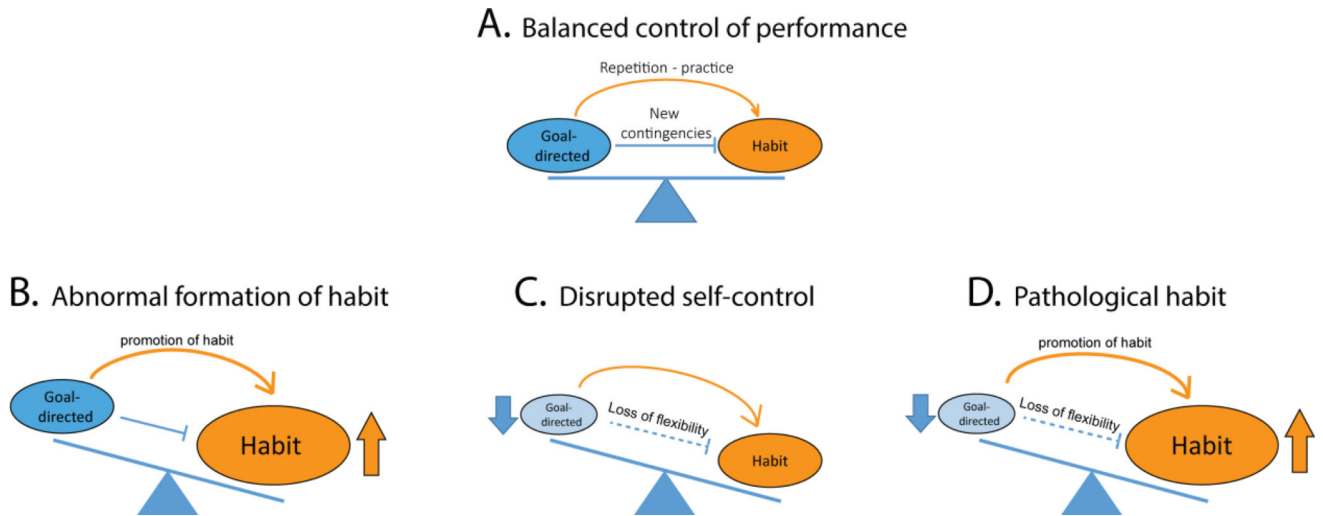


Figure 2.

Disentangling abnormal habit from deficits in goal-directed control in substance use disorders. A. Balanced control of performance in non-addicted individuals; Repetition and practice promote the formation of habit but in situations that warrant deliberation, such as changes in environmental contingency, individuals can re-exert goal-directed control over habit to flexibly adjust their responding. B. According to the first model, the unbalance between goal-directed and habitual control in SUD would result from accelerated formation of abnormal and strong drug habits without deficit in goal-directed behavior C. According to the “disrupted self-control” model, habit would not be abnormally strong but would be expressed more frequently as a result of impaired goal-directed and inhibitory control. D. The “pathological habit” model posits that impairments in goal-directed control combine with abnormal habit formation to promote compulsive drug use in SUD.