Goal-Directed and Habitual Control in Human Substance Use: State of the Art and Future Directions

Nuria Doñamayor\textsuperscript{a} Claudia Ebrahimi\textsuperscript{a} Viktoria A. Arndt\textsuperscript{b} Franziska Weiss\textsuperscript{b,c} Florian Schlagenhauf\textsuperscript{a} Tanja Endrass\textsuperscript{b}

\textsuperscript{a}Department of Psychiatry and Neurosciences CCM, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; \textsuperscript{b}Faculty of Psychology, Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany; \textsuperscript{c}Department of Clinical Psychology, Central Institute of Mental Health, University of Heidelberg/Medical Faculty Mannheim, Mannheim, Germany

Keywords
Addiction · Substance use disorder · Habits · Goal-directed behavior

Abstract
Theories of addiction posit a deficit in goal-directed behavior and an increased propensity toward habitual actions in individuals with substance use disorders. Control over drug intake is assumed to shift from goal-directed to automatic or habitual motivation as the disorder progresses. Several diagnostic criteria reflect the inability to pursue goals regarding reducing or controlling drug use and performing social or occupational functions. The current review gives an overview of the mechanisms underlying the goal-directed and habitual systems in humans, and the existing paradigms that aim to evaluate them. We further summarize the current state of research on habitual and goal-directed functioning in individuals with substance use disorders. Current evidence of alterations in addiction and substance use are mixed and need further investigation. Increased habitual responding has been observed in more severely affected groups with contingency degradation and some outcome devaluation tasks. Reduced model-based behavior has been mainly observed in alcohol use disorder and related to treatment outcomes. Motor sequence learning tasks might provide a promising new approach to examine the development of habitual behavior. In the final part of the review, we discuss possible implications and further developments regarding the influence of contextual factors, such as state and trait variations, and recent advances in task design.

Introduction
A loss of control over drug intake is a central characteristic and diagnostic criterion of addiction: individuals report being unable to stop drug intake despite foregone positive outcomes and devastating (future) negative consequences. While the clinical phenomenon of addiction is well described and a large body of neurobiological findings has been established, the exact mechanisms of the development and maintenance of addiction are still insufficiently understood. Drug intake has been conceptualized as a learned behavior, where drugs act as instrumental reinforcers \cite{1-4} and addiction as a disorder of aberrant choice behavior \cite{5-10}. Within this framework,
this loss of control over drug intake may be related to a shift away from goal-directed toward habitual behavioral control [11–16], i.e., drugs are initially consumed to achieve a certain hedonic goal or to avoid discomfort, but following repeated consumption, drug taking behavior becomes habitual, elicited by antecedent environmental stimuli and performed without considering the outcome. This gradual shift may contribute to automatic or habitual drug intake. While such an account has some face validity, the operationalization of the goal-directed and habitual systems and their contribution to specific (drug-related) choices in humans remains challenging.

Several paradigms have been used to disentangle goal-directed versus habitual control in humans. Outcome devaluation and contingency degradation tasks are well established in animal research and were subsequently transferred to human studies, whereas sequential decision-making tasks firmly based on computational modeling took the reverse path from human to animal research (for a review of the historical evolution of the constructs, see [13]). However, it needs to be taken into consideration that at least outcome devaluation and contingency degradation tasks, as well as the most commonly used computational parameter derived from sequential decision-making tasks, typically only assess the relative balance between the two systems (see [17–20]). Moreover, associative learning theories tend to define habitual control as the absence of goal-directed behavior [17, 18, 21, 22]. In this review, we will briefly introduce the different tasks used to assess those processes in humans and provide a systematic review of the available evidence regarding goal-directed versus habitual control in individuals who consume drugs of addiction (this review focuses exclusively on human research, for reviews of animal findings, see [15, 23]). Finally, we will synthesize the available evidence supporting the habit account of addiction and discuss open issues especially with respect to measuring habitual control in humans and the influence of contextual factors.

Current State of the Evidence for the Dual Systems Theory in Addiction

Dual systems accounts posit that action control is driven by the balance between two competing systems: a habitual and a goal-directed system [12, 22, 24]. The habitual system is driven by stimulus-response (S-R) associations, in which the outcome (O) merely reinforces said association, consistent with Thorndike’s law of effect [25]. Therefore, when behavior is driven by habits, the subject tends to repeat actions that had been reinforced during the learning phase and no longer considers current outcome characteristics. Habits are therefore highly automatic and efficient, which comes at the cost of behavioral flexibility [13, 26]. In contrast, goal-directed control involves knowledge about both the R-O relationship and current outcome value [13, 24]. When behavior is primarily controlled by the goal-directed system, subjects will adjust their actions to changes in both the motivational value of the outcome and the instrumental R-O contingency. Goal-directed behavior is thus actively deliberative and flexible, but requires more resources [13, 26]. Therefore, goal-directed and habitual behaviors differ in their sensitivity to changes in both the causal nature of the R-O relationship and the current value of the outcome [12, 27]. Classically, this has been assessed using two tasks: contingency degradation and outcome devaluation, respectively.

Outcome Devaluation Tasks

In outcome devaluation tasks, subjects first acquire an instrumental action, which is followed by the devaluation of one or more outcomes, while the remaining outcomes preserve the same value as in the instrumental learning phase. Outcome devaluation can be achieved through several types of manipulations. Translated from animal research [22, 28, 29], some paradigms have employed outcome-specific satiation, e.g., by selectively feeding participants one of the outcomes to satiety [30–32], or aversive conditioning, e.g., by showing one of the outcomes infested with vermin [33, 34] or pairing it with a bitter taste [34]. One of the most widely used outcome devaluation tasks, the so-called fabulous fruit game or slips-of-action task [35, 36] (Fig. 1a), makes use of instructed devaluation after an initial discrimination training, i.e., participants are simply instructed that certain outcomes are no longer valuable. Finally, outcome devaluation tasks end with a test phase in extinction, in which participants are presented with the stimuli associated to both still valuable and devalued outcomes, but no feedback is given following responses. During the test phase, if the participant’s behavior is driven by the habit system, they will respond equally to stimuli associated with both valuable and devalued outcomes, as this system is insensitive to changes in outcome value. If choices are guided by the value-sensitive goal-directed system, participants will withhold their actions to stimuli associated with devalued outcomes and only perform responses that lead to valuable outcomes. It has been extensively dem-
onstrated that both healthy animals and humans are sensitive to changes in outcome value, as reflected by decreased responding toward devalued outcomes [22, 29, 30, 35, 37].

We identified a total of ten outcome devaluation studies examining users of drugs of addiction: five examined nicotine, two alcohol, two cocaine, and one polysubstance use. The studies with cigarette smokers by Hogarth et al. [38–41] used instrumental learning paradigms, in which participants were prompted to press one of two buttons at will, each leading to a different outcome. The learning phase was followed by outcome devaluation and a subsequent test in extinction, revealing rather intact goal-directed control in smokers. In their initial study, Hogarth and Chase [38] reported two experiments in smokers. Participants were divided in two groups, for which either cigarette or chocolate outcomes were devalued by satiety (first experiment) or by issuing health warnings (second experiment), resulting in outcome-specific devaluation effects consistent with goal-directed behavior. In a between-subjects design, Hogarth [39] reported that devaluation through chocolate satiation reduced responses to obtain chocolate, whereas devaluation using a nicotine replacement therapy nasal spray attempting to induce nicotine satiation did not reduce choices for cigarettes. However, this devaluation insensitivity for cigarettes was only observed in individuals with higher smoking severity/desire, who might not have reached satiation by the offered dose of nicotine.

Based on findings showing that devaluation effects can be abolished by stress induction [42] and acute alcohol administration [43], Hogarth et al. [40] examined the effect of alcohol expectancy in smokers. After instrumental training for tobacco and chocolate outcomes, cigarettes were doubly devalued by health warnings and then smoking to satiety. Before proceeding to the extinction test, participants were presented with a glass of either water or an alcoholic beverage (beer or wine) to be consumed at the end of the experiment. The expectation of alcohol compared to water reduced devaluation sensitivity during the extinction test. The appraisal of the alternative reinforcer (alcohol) was suggested to limit the ability to retrieve drug (tobacco) value, which is required for goal-directed control, thus promoting a shift to habitual control. Likewise, Hogarth et al. [41] showed that sensitivity to devaluation by smoking satiety in daily smokers was reduced by the induction of negative mood [44] prior to the extinction test. Additional analyses showed that self-reported negative mood increased responding for cigarettes during the extinction test, whereas positive mood decreased it. This effect was taken as evidence that negative mood raised the expected value of drugs through incentive learning (since smoking is expected to alleviate negative mood), instead of being the result of automatic S-R associations (see [45]).

Recently, Hogarth et al. [46] reported two experiments with polysubstance users and unmatched healthy controls. In these experiments, devaluation of water by satiety and aversive devaluation of cola using health warnings, respectively, selectively reduced responding for either outcome in both groups similarly, suggesting intact goal-directed control in polysubstance users. A recent neuroimaging study compared abstinent individuals with alcohol use disorder (AUD) and healthy controls [34], utilizing outcome devaluation with both an aversive video and taste aversion of one of two possible snacks [33]. After devaluation, both groups showed comparably reduced choices for the devalued snack, suggesting intact goal-directed control in AUD [34]. In summary, these two studies by Hogarth et al. [46] and van Timmeren et al. [34] found no differences in goal-directed versus habitual control between patients with substance use disorder (SUD) and controls.

Four studies used instrumental S-R-O learning tasks with instructed devaluation and/or slips-of-action tests. Using the fabulous fruit game [47] in a neuroimaging design, Sjoerds et al. [48] examined individuals with AUD versus healthy controls. Accuracy in the outcome-devaluation test was significantly lower in AUD subjects, suggesting impaired response-outcome knowledge. Furthermore, the AUD group showed reduced activity in the ventromedial prefrontal cortex (PFC) which has been related to goal-directed learning [12, 47, 49], but increased activity within the posterior putamen, an area that has been associated with habitual behavior [12, 32]. Results did not differ between neutral (fruit) and alcohol-related images. Ersche et al. [14] examined individuals with cocaine use disorder (CUD) and healthy controls using appetitive [50, 51] and avoidance [52] instrumental learning tasks in combination with devaluation procedures. Participants with CUD performed worse than controls during both appetitive and avoidance instrumental learning phases, but devaluation differed between conditions: whereas both groups decreased avoidance responding for devalued outcomes, the CUD group exhibited comparable response rates for valued and devalued appetitive outcomes, suggesting a shift from goal-directed to habitual responding in CUD specific to appetitive but not aversive outcomes [14]. Subsequently, Lim et al. [53] reanalyzed the appetitive task using a hierarchical Bayesian learning
model. Higher values of the reinforcement sensitivity or inverse temperature parameter from the computational model, which indicates a greater influence of action values on choices, were related to increased habitual responding during the slips-of-action test. CUD patients further displayed reduced white matter integrity within the putamen-premotor cortex (habitual pathway, but the goal-directed pathway (anterior caudate to medial orbitofrontal cortex) did not differ between groups, although only healthy controls exhibited a positive association with learning rate. However, the authors considered that the overall findings did not provide strong support for deficits in either system and concluded that “drug addiction results in an imbalance between goal-directed and habitual control over behavior” [53]. Using variants of the task used by Ersche et al. [14], Luijten et al. [54] reported intact goal-directed control in cigarette smokers compared to controls, as reflected by decreased responding to devalued outcomes in both appetitive and avoidance tasks. However, dovetailing with the findings in CUD [14], participants with more severe nicotine dependence (assessed by the Fagerström Test for Nicotine Dependence [55]) committed more slips-of-action in the appetitive task, suggesting increased habitual control in appetitive contexts [54].

In summary, the studies reviewed here showed that sensitivity to devaluation appeared preserved in cigarette smokers [38, 41, 54], was modulated by factors such as the expectation of alcohol consumption [40] or negative mood [41], but correlated with smoking severity [39, 54]. Indeed, the use of discriminative stimuli in outcome devaluation tasks provides evidence of impaired goal-directed control in drug users compared to healthy controls [14, 48, 53]. However, increased habitual responding in SUD may be restricted to appetitive learning, in contrast to obsessive-compulsive disorder, in which this has been observed selectively for aversive learning [56].

Contingency Degradation Tasks

Contingency degradation tasks can be used to assess behavior when the causal relationship between an action and its outcome changes. In these tasks (Fig. 1b), an action is first paired with an outcome with a fixed probability, \( P(O|R) \), i.e., the probability that outcome \( O \) occurs given that action/response \( R \) is performed. Subsequently, the causal R-O relationship is degraded by increasing the probability that said outcome is presented in the absence of an action, \( P(O|\sim R) \). Therefore, when both probabilities are equal, the action has no effect on the likelihood of the outcome, so that the net R-O contingency, \( \Delta P = P(O|R) - P(O|\sim R) \), and the causal status of the action are zero [22, 57]. If behavior is driven by the goal-directed system, the participant’s response rate will correlate with \( \Delta P \), whereas habitual responding would lead to similar response rates irrespective of whether contingency is positive, negative, or fully degraded. Several studies have demonstrated that healthy participants are sensitive to contingency degradation, with higher response rates associated to positive \( \Delta P \)

---

Fig. 1. Schematic depiction of paradigms used in the study of goal-directed versus habitual behavior. a Slips-of-action task. In the instrumental learning phase, participants learn multiple (generally, 4–6) S-R-O associations. In the instructed outcome devaluation phase, all outcomes are shown simultaneously and two of them are devalued (crossed out), while the others remain valuable. In the slips-of-action test, the stimuli are sequentially presented in extinction: if behavior is controlled by the goal-directed system, the participant will only respond to stimuli leading to valuable outcomes and withhold their response to stimuli leading to devalued outcomes; whereas if behavior is controlled by the habitual system, the participant will respond to all stimuli irrespective of their current value. b Contingency degradation task. In positive contingency conditions, responses lead to an outcome with a fixed probability, \( P(O|R) \), while withheld responses are not rewarded. In degraded contingency conditions, the probability that the outcome is presented in the absence of an action, \( P(O|\sim R) \), is increased up to the point in which both probabilities are equal and thus the net R-O contingency, \( \Delta P = P(O|R) - P(O|\sim R) \), and the causal status of the response are zero. If behavior is controlled by the goal-directed system, response rates will strongly correlate with \( \Delta P \), whereas if behavior is controlled by the habitual system, response rates will be similar irrespective of the causal status of the action. c Two-step task. In an initial stage (grey), participants are asked to choose one of two stimuli, each of which leads to one of two second-stage states (green and yellow) with a certain probability (generally, \( p \leq 0.7 \) and to the other state with inverse probability (\( p \geq 0.3 \)). Each of the four second-stage stimuli is rewarded with a certain probability (generally, bound between 0.25 and 0.75) that changes over time according to a Gaussian random walk. If behavior is controlled by the model-based system, participants will take both rewards and task structure into account, repeating first-stage actions that led to a reward only after a common state transition, but also those that resulted in no reward following a rare transition; whereas if behavior is controlled by the model-free system, participants will simply repeat first-stage actions that led to a reward. d Sequence learning task. In the initial mapping stage, associations between visual stimuli and corresponding button-presses are overtrained. Later, certain S-R associations are switched. If behavior is controlled by the goal-directed system, participants will correctly perform the remapped button-presses; whereas if behavior is controlled by the habitual system, participants will commit errors by performing the button presses learned in the initial associations instead of the remapped responses.

(For figure see next page.)
values and lower to negative $\Delta P$ values [57–60]. Moreover, healthy subjects’ explicit causality ratings closely match their behavior [57–61], i.e., not only do participants rate their actions as more causal with increasing $\Delta P$ values, but response rates are also higher under conditions in which actions are perceived as more causal.

To date, only one study has used a contingency degradation paradigm in a sample of patients with CUD [62]. Ersche et al. [62] observed that both healthy controls and individuals with CUD decreased responding with decreasing $\Delta P$ values, though this decrease was less steep in CUD, indicating stronger habitual tendencies. Moreover, in contrast to nonaddicted controls, the CUD group failed to adjust responding between partially ($P[O|R] > P[O|\sim R]$) and fully ($P[O|R] = P[O|\sim R]$) degraded conditions. This was mirrored by causality ratings, with CUD patients rating their actions as similarly causal in both partially and fully degraded conditions. The duration of cocaine use correlated with increased habitual responding as well as with higher self-reported automaticity, with the latter also associated with lower magnetic resonance spectroscopy glutamate measures in individuals with CUD.

**Sequential Decision-Making Tasks**

More recently, computational accounts based on reinforcement learning algorithms [63] have characterized goal-directed and habitual behavior in terms of model-
based and model-free control, respectively [13]. Within this framework, one of the most widely used tasks is the so-called two-step task (Fig. 1c), originally developed by Daw et al. [20] and with recent reformulations (e.g., [64]). In the first stage, participants are asked to choose one of two stimuli, each of which leads to one of two second-stage states with a certain probability (generally, \( p \geq 0.7 \)) and to the other state with inverse probability (i.e., \( p \leq 0.3 \)). Each of the four stimuli in the two-second-stage states (two in each state) leads to a reward with a certain probability that varies according to a Gaussian random walk over the course of the task. Newer reformulations of the task include deterministic transitions [64] and/or varying reward magnitudes rather than probabilities [64, 65]. In the two-step task, model-free agents will select actions only based on previous reinforcement, thus repeating the first-stage action that led to a reward in the previous trial and switching if said action was not rewarded; whereas model-based agents will incorporate information about the environmental structure and its interaction with the outcomes resulting from the agent’s actions, thereby repeating first-stage actions that led to a reward only after a common state transition, but also those that resulted in no reward following a rare transition. The computational model most commonly applied to analyze this task uses an \( \omega \) parameter that weights the relative amount of model-based and model-free control employed, with \( \omega = 0 \) indicating a fully model-free strategy and \( \omega = 1 \) a fully model-based strategy [20]. More recent reformulations of the model forgo \( \omega \) and characterize model-free and model-based control in terms of the inverse temperature parameters \( \beta_{\text{MF}} \) and \( \beta_{\text{MB}} \), respectively (though note that these are algebraically equivalent to the original model in that \( \beta_{\text{MF}} = (1 - \omega) \beta \) and \( \beta_{\text{MB}} = \omega \beta \) [66]). Performance in two-step tasks is consistent with a mixture of model-free and model-based control [20, 64, 67–69].

We found nine studies using the two-step task, eight of which alcohol use. In a seminal study, Sebold et al. [70] demonstrated that recently detoxified (2–39 days) individuals with AUD, while not differing from healthy controls in their use of model-free strategies, displayed decreased model-based behavior. This decrease was specifically driven by an increased tendency to switch first-stage choices after unrewarded actions following rare transitions. However, patients also showed decreased cognitive speed (assessed by the Digit Symbol Substitution Test [71]) and when accounting for this difference, the model-based scores of AUD patients no longer significantly differed from healthy participants. In a subsequent prospective study, AUD patients were divided into those who had relapsed and those who had remained abstinent at 1-year follow-up [72]. Whereas, in this study, model-based control did not differ between patients and controls, individuals who relapsed showed an aberrant relationship between model-based control and alcohol expectancies (assessed by the Alcohol Expectancy Questionnaire [73]) compared with the other two groups: higher alcohol expectancies (i.e., the expectation of more positive effects from drinking alcohol) were associated with lower model-based control in AUD patients who relapsed at follow-up. Sebold et al. [72] suggested that the positive association between alcohol expectancies and model-based control in healthy and abstinent AUD participants might help these individuals achieve a goal-directed use of alcohol; but among those who relapsed, the authors posited that participants with high model-based control might underestimate the effects of alcohol, whereas low model-based control might facilitate alcohol consumption in those with high alcohol expectancies [72]. Moreover, the patient group that relapsed displayed decreased model-based prediction error signals within the medial PFC [72]. Similarly, Voon et al. [11] did not find differences in model-based control, as reflected by the computational parameter \( \omega \), between long-term abstinent (2 weeks to 1 year) AUD patients and healthy controls, although they did observe decreased second-stage learning rates (\( \alpha_2 \) parameter). Using a retrospective approach, however, these authors reported a positive association between model-based control and abstinence duration [11]. Together, the studies by Sebold et al. [72] and Voon et al. [11] suggest that model-based control could be related to treatment outcome, as it predicted relapse in combination with high alcohol expectancies, and increased values were observed with longer abstinence duration.

A number of studies have also used two-step tasks in subjects without a diagnosis of AUD but different levels of potentially problematic drinking behaviors. Model-based control, expressed by the \( \omega \) parameter, was found to be decreased in a sample of severe binge drinkers [74], as per criteria of the National Institute on Alcoholism and Alcohol Abuse [75], compared to healthy volunteers. Moreover, binge drinkers displayed decreased learning from first-stage choices (\( \alpha_1 \) parameter) and increased perseverative tendencies. In these subjects, model-based control was higher (both choice behavior and \( \omega \) values) and model-free choice behavior was lower as time passed since the last reported binge episode, pointing to a potential state effect of alcohol on model-based versus model-free behavior, as well as a possible amelioration of behav-
ioral control with alcohol use cessation [74]. Two recent studies [69, 76] explored whether the proposed imbalance between model-based and model-free control might be a predisposing factor for risky alcohol consumption. In a large sample of 18-year-old male social drinkers, Nebe et al. [69] found no association of alcohol consumption with model-based/model-free choice behavior, nor with model-based/model-free reward prediction errors within ventromedial PFC or ventral striatum. However, in a follow-up study in the same cohort, Chen et al. [76] reported a negative association between model-based behavior at baseline and binge-drinking trajectories over a 3-year period. Moreover, model-free reward prediction errors within ventromedial PFC and ventral striatum were associated with increased alcohol consumption over time. In line with the findings of Sebold et al. [72], alcohol expectancies mediated the relationship between model-based control and binge-drinking trajectory, such that only participants with low model-based control and expectations of mainly positive reinforcing effects from alcohol displayed increased binge-drinking behavior at follow-up [76]. Taken together, these findings suggest that a shift away from model-based toward model-free control could lead to more risky drinking behavior in young men, although the degree of model-based versus model-free control does not predict actual alcohol consumption in this population.

Further insights in the context of alcohol use have also been provided by two large online studies [66, 77]. Using the original two-step task together with several symptom questionnaires and factor analysis, Gillan et al. [66] reported decreased model-based choice behavior and \( \beta_{MB} \) values with higher scores in the Alcohol Use Disorder Identification Test (AUDIT [78, 79]). They also found a negative relationship between model-based control and compulsive behavior and intrusive thought, which was even more pronounced in what the authors defined as “putative” patients, i.e., participants scoring in the top 25% on the AUDIT. Consistent with studies in AUD patients [70, 72], Gillan et al. [66] observed no association between model-free behavior and alcohol use severity. In the second online study, Patzelt et al. [77] used a modified two-step task with deterministic transitions and high-/low-incentive trials [80]. They observed no association between model-based control and AUDIT scores, but relationships with compulsive behavior and intrusive thought. Moreover, the previously reported sensitivity of model-based control to incentive manipulations [80] was preserved across levels of alcohol use severity.

To our knowledge, to date, only the aforementioned study by Voon et al. [11] has used the two-step task in addictions other than AUD. This study showed that abstinent methamphetamine-dependent participants displayed both lower model-based control, as reflected by the \( \omega \) parameter, and increased stochasticity of their second-stage choices (\( \beta_2 \) parameter) compared to healthy controls. Beyond this, a recent study used the two-step task with two nonclinical student samples who used a range of substances, including alcohol, marihuana, and hallucinogens, but the authors did not observe clear associations with model-based choice behavior [81].

In summary, two-step studies with AUD patients have consistently shown no differences from healthy participants in model-free control [70, 72], which is considered the computational pendant of habitual behavior [13]. However, results pertaining model-based control, the computational formalization of goal-directed behavior [13], are mixed [11, 70, 72], suggesting that additional factors might influence changes observed in this component. Specifically, abstinence [11, 72], alcohol expectancies [72] and cognitive functioning [70] have been related with model-based control. A relationship between model-based control and time elapsed since the last drinking episode has also been reported in severe binge drinkers [74], a group particularly at risk for the development of AUD [82]. Yet evidence pointing to an association between model-based/model-free control and alcohol use in nonclinical samples is quite heterogenous [66, 69, 76, 77], suggesting a weak association that might require larger samples [66] or sophisticated longitudinal within-subject approaches [76] to be detected. Since alcohol consumption induces partially reversible changes in prefrontal-striatal regions [83, 84], which are also known to be essential for model-based control [20, 85, 86] as well as for more general cognitive functioning [87, 88], the question remains whether model-based deficits in AUD and their relationship to other factors are causal or result from alcohol addiction.

Sequence Learning Tasks

Based on the definition of habits as overlearned S-R associations, some researchers have used sequence learning or serial reaction time tasks, generally used to research motor skill learning, to investigate the development of habitual behavior. Typically, in these tasks (Fig. 1d), simple finger-tapping sequences and/or visuomotor associations are overtrained until they reach a certain level of efficiency but also habitual rigidity, hindering flexible, goal-directed responding. Task designs include finger-
tapping sequences, in which participants learn to press buttons in a specific order (indicated by visual stimuli or explicit instructions), or visuomotor associations, in which participants learn associations between specific stimuli (e.g., abstract letters or pictorials) and corresponding button-press responses [89–92]. While the link between these tasks and free-operant procedures (e.g., [38–41]) might not be evident, slips-of-action [35, 36] are thought to be likely related to motor behavior (for a review, see [93]). Given that the loss of flexible control over behavior is a crucial feature of addiction, investigating (motor) skills, especially in regards to the notion of automaticity [93], might be of particular relevance. Automatic motor skills, as overly fixed behavioral routines, could be used to map dissociable changes in neural activity during habit formation (e.g., striatal regions and striato-cortico-cerebral circuitry [94–97]) and, thus, offer valuable clues about the fundamental understanding of addiction.

We identified two studies that used serial reaction time tasks to examine habitual responding in addiction. McKim et al. [98] used a visuomotor S-R sequence learning and relearning task to examine habitual propensity in polysubstance users versus healthy controls. In a two-day procedure, participants first learned S-R associations for two sets of abstract visual stimuli, adding two novel sets on the second day. After training with all four sets on the second day, the S-R associations for one of the initial, well-trained sets and one of the novel sets were reversed. To quantify the degree of habitual behavior, the authors then compared the proportion of perseveration errors (i.e., responding in accordance with the original S-R rule) in well-learned S-R sets versus novel S-R sets following this instructed response devaluation. Polysubstance users learned new S-R associations equally well as healthy controls, indicating no impairment in goal-directed action selection. However, they committed a greater proportion of perseverative errors after the change of the S-R rule than healthy controls, indicating an impairment in overcoming well-learned S-R associations [98]. In a second study, McKim et al. [99] examined whether 10 Hz transcranial alternating current stimulation versus sham stimulation of the dorsolateral PFC would reduce perseverative responding in this task. Contrary to their prediction, perseveration errors in the well-learned S-R sets were unaffected by active stimulation in the SUD group, but increased in the control group. Moreover, the authors could not replicate the finding of increased perseveration errors as an indicator of habitual tendency in SUD, since these did not differ between groups under sham stimulation, which was potentially due to practice effects related to the repeated-measures design in the SUD group. Interestingly, however, 10 Hz active stimulation reduced perseveration errors in SUD individuals with a longer history of substance use, suggesting a reduction of habitual responding and thus an increase in flexible responding in these participants.

**Discussion**

Alterations of goal-directed versus habitual control play an important role in many models of addiction [1, 11] and animal research has shown powerful links between addiction and habitual propensity [16, 100]. The current review provides an overview of different tasks used to examine goal-directed versus habitual control in humans, and findings in participants with addiction and substance use. In recent years, classical tasks – outcome devaluation and contingency degradation – have been joined by paradigms originating from computational approaches, like sequential decision-making tasks, but also by simpler tasks trying to distill S-R-driven behavior, such as sequence learning tasks.

With the exception of two studies with polysubstance users [46] and individuals with AUD [34], studies using outcome devaluation tasks closely translated from animal research have been conducted in cigarette smokers [40, 41] and have generally shown that participants respond less for devalued outcomes, consistent with goal-directed behavior [40, 41, 45]. However, these studies also indicate goal-directed responding is influenced by specific factors, such as alcohol expectancy [40], negative mood [41] or smoking severity [39], which were shown to diminish devaluation sensitivity. In contrast, studies using slips-of-action tasks report decreased goal-directed control in substance users [14, 48, 53], although this appears restricted to appetitive learning and has not been observed for aversive learning [14]. This effect has been in part replicated in cigarette smokers, in whom smoking severity was anticorrelated with goal-directed behavior in appetitive learning [54]. The second of the classical paradigms, contingency degradation, has only been used in one study with SUD patients to date. This showed that individuals with CUD did not adapt their behavior to fully degraded reinforcement contingencies, indicating increased habitual versus goal-directed control [62]. Moreover, longer cocaine use duration was associated with increased habitual tendencies, suggesting that continued drug use may promote habit formation.
Although classical tasks provide some evidence for increased habitual versus goal-directed control, there are also some inconsistencies between studies. Simple, free-operant tasks translated from the animal literature, in which human participants are prompted to press either of the learned buttons at will after outcome devaluation (e.g., [34, 38, 39]), might be not sensitive enough to induce or reveal habitual responding in humans or to detect SUD-related changes in habitual versus goal-directed control. The interplay between the habitual and goal-directed systems has been demonstrated to be affected by the time available to compute the optimal response, and the structure of these simple, free-operant tasks might allow for more active deliberation (see also [15, 101]). Additionally, human subjects may quickly infer what is expected from them, e.g., they may reduce responding after outcome devaluation because they think that is what they are supposed to do. Another limitation of these tasks is that the stimulus that triggers the S-R association, which by definition drives the habitual system [12], may be difficult to identify. This has led some authors to even consider that the stimulus does not directly trigger the habitual response [101, 102]. In contrast, more complex tasks, in which participants are (over-) trained in multiple differential S-R-O associations [35, 36], seem to produce more reliable evidence for habitual responding [14, 48, 53, 101] and have revealed that participants with SUD are less able to withhold responses to stimuli associated with devalued outcomes than healthy controls [14, 48, 53], although the number of studies using these tasks is currently small. It has been suggested that these results could be explained by impaired contingency knowledge resulting from task disengagement [45], but experimental results might not support this [53]. Furthermore, across outcome devaluation tasks, cigarette smokers appear less likely to display impairments in habitual versus goal-directed control than individuals with other SUDs [38, 54]. The experimental samples of cigarette smokers ranged widely in smoking severity and, in contrast to other SUDs, participants were generally not seeking treatment and smoking is broadly accepted, which limits the negative social and occupational impact of the disorder. Indeed, only selective analyses of the most severe smokers have been able to demonstrate reduced goal-directed behavior [39, 54], which were otherwise diluted within the whole sample.

In contrast to outcome devaluation and contingency degradation tasks, derived from animal research, sequential decision-making tasks were originally devised for human studies and, in the context of addiction, the two-step task has so far been almost exclusively employed with individuals with AUD or subclinical samples with potentially hazardous drinking patterns. Overall, these studies only partially support the habit theory of addiction [4], as they consistently show no evidence of increased model-free control associated to AUD [70, 72], but some research points to deficits in model-based behavior ([70]; but see [11, 72]). Importantly, model-based control has been both prospectively and retrospectively associated with abstinence in AUD patients [11, 72], and negatively associated with problematic alcohol consumption in nonclinical samples ([66, 76]; but see [69, 77]). Decreased model-based prediction error signals within the medial PFC have also been reported in individuals with AUD that had relapsed [72]. However, whether an imbalance between model-based versus model-free control constitutes a vulnerability factor or develops over the course of AUD, or whether there even is an interaction between vulnerability and development, remains to be answered.

Although the equivalence of model-based and model-free control with goal-directed and habitual behaviors, respectively [64, 103], has been widely adopted, some researchers have raised a number of questions, specifically regarding the alignment of model-free and habitual behavior. Indeed, the notion of a model-free agent that adjusts its behavior based on previous reinforcement already contradicts the definition of habits as rigid S-R-driven behaviors [25, 104]. Furthermore, while animal research points to two distinct neural circuits underlying habit and goal-directed systems [12, 105], the brain signatures of model-free and model-based control largely overlap (e.g., [20, 72, 85]), mostly including regions of the goal-directed circuit [12]. It has thus been a long-standing matter of debate whether sequential decision-making tasks assess the same construct as classical paradigms [13, 106]. Studies examining this issue have only found correlations between devaluation sensitivity and model-based, but not model-free, measures [31, 107, 108], and some researchers have suggested that two-step tasks might not be suitable to study human habitual behavior at all [109, 110]. Altogether this might speak for both model-based and model-free behaviors being value-based and thus both aligning with goal-directed control [111, 112], with some authors proposing alternative taxonomies that might better align with the concepts of habitual and goal-directed behavior (for a review, see [111]), and in turn providing a possible explanation as to why studies in addiction have not been able to find changes in model-free control that would more strongly support the habit theory of addiction.
Notwithstanding, computational modeling approaches provide a powerful tool for mapping latent decision-making processes to overt behavior and neural function [113]. The generative reinforcement learning model described by Daw et al. [20] to analyze the two-step task and its reformulations [66] make use of the full trial-by-trial choice history, thereby capturing learning processes that occur over the course of multiple trials. In contrast, simpler nongenerative approaches only approximate this learning dynamic, e.g., by estimating stay probabilities based on information from the immediately preceding trial [20]. However, computational modeling requires methodological factors, such as model parameterization or parameter estimation, to be considered carefully in order to reliably measure these underlying mechanisms [114]. Further optimizations of the computational model used originally for the two-step task (see, e.g., [66, 115–117]) have led to the characterization of model-free and model-based control in terms of the inverse temperature parameters $\beta_{\text{MF}}$ and $\beta_{\text{MB}}$ instead of the weighting parameter $\omega$, and have yet to be used to analyze behavior from participants with SUD (with the exception of the dimensional study by Gillan et al. [66]).

Moreover, several researchers have noted how, under certain circumstances, a model-free agent can appear model-based [118–120] and vice versa [121]. Akam et al. [118] demonstrated that purely model-free strategies can result in the interaction between reward and state transition that resembles model-based control due to a correlation between action values at the start of the trial and subsequent events, and how the addition of further predictors to the stay probability analysis can be efficacious in dealing with such confounders [118, 120]. Furthermore, Akam et al. [118] used simulations to expose how, by taking advantage of dependencies in the task structure, model-free agents could develop “extended-state” strategies that appear similar to model-based behavior. Although the authors acknowledge that such strategies are unlikely to explain the model-based behavior reported in human literature, they suggested that limiting the number of trials, using neural data to differentiate strategies or introducing reversals in the transition matrix could further minimize the influence of “extended-state” strategies. However, the work by Akam et al. [118] raises the question of whether model-based control might have been overestimated in previous studies [11, 66, 69, 70, 72, 77, 81], which might have influenced potential group differences. On the other hand, Feher da Silva and Hare [121] demonstrated that purely model-based agents that used incorrect task models, such as erroneous assumptions about first-stage stimuli or different learning rates for common/rare transitions, appeared to display a mixture of model-based and model-free control. Furthermore, these authors showed experimentally that slight changes to the task instructions, namely framing the task in the context of a story and providing explanations for all task events (transitions, stimulus location, etc.), almost abolished model-free behavior in favor of primarily model-based choices. These studies [118–121] imply a need to reassess both task implementation and analysis tools when using sequential decision-making tasks. Of the SUD studies reviewed in this work, only the study by Patzelt et al. [77] applied some of the modifications described by Feher da Silva and Hare [121], reporting no association between model-based control and alcohol use severity. Future studies should implement the suggested adjustments to task instructions [121] and analyses [118, 121] to confirm the extent to which previous results in SUD hold.

More recent studies have started using motor sequence learning tasks, since when breaking down the definition of habit as a basic S-R association, these may be operationalized as visuomotor learning and relearning tasks. In this context, overtraining is used to attempt to induce habits in humans, mirroring the shift from goal-directed to habitual control demonstrated in animal studies ([12, 122, 123]; although see, e.g., [124–126]), and perseverative errors can be used to quantify habitual behavior. This approach appears simple, but the findings from both studies to date on individuals with SUD are not consistent. Whereas findings from the first study suggested that SUD patients committed more perseveration errors, indicating a difficulty to overcome well-learned S-R associations and a propensity toward habitual behavior [98], the authors did not replicate this finding in their second study [99]. Indeed, experimental habit induction through overtraining has proven elusive in humans, since behavior tends to remain goal-directed (see, e.g., [127]). A recent study by Hardwick et al. [101], however, successfully demonstrated experimentally induced S-R habits in healthy participants. Given that movement preparation precedes initiation [128] and that habitual responses are prepared rapidly but not necessarily initiated immediately, their expression could be overridden by the goal-directed system if given enough time [101]. Using a forced choice-paradigm, Hardwick et al. [101] observed that overtrained participants committed more perseverative errors than those without prior training when given short response preparation times, whereas this difference disappeared when they were given longer preparation times.
Such time restrictions might at least partially account for the fact that reduced goal-directed responding in SUD has been reported in the slips-of-action task (e.g., [14]), but not in free-operant tasks (e.g., [34, 38–41]). In contrast to the latter, slips-of-action tasks (e.g., [50, 51]) involve selecting among several potential actions to imperative stimuli under narrow time constraints. It is precisely these characteristics, for which slips-of-action have been deemed closely related to motor skills [93], that might allow to better capture the rapidly occurring habitual action selection. Moreover, the results of Hardwick et al. [101] indicate that habitual responses can in fact be experimentally induced by overtraining. While this might not solve the issue that most of the paradigms currently used, both classical and computational, can only evaluate the relative balance between goal-directed/model-based and habitual/model-free control [19, 127], it does offer promising prospects for future habit research, suggesting simple task adjustments, in the form of increased time pressure, might be able to uncover habitual responding before the dominant goal-directed system takes over.

As outlined above, some of the studies reviewed point to moderating factors that influence the balance between habitual and goal-directed control, such as negative mood [41], the prospect of drinking alcohol [40], working memory capacity [70] or the expectancy of positive effects from drinking alcohol [72]. While yet to be investigated in participants with SUD, another factor that has been shown to impair both sensitivity to devaluation [42, 129] and model-based control [115, 130] is stress, both induced during the experimental session and lifetime, chronic stress. Animal research has repeatedly shown that stress promotes compulsive drug seeking [for a review, see [131]] and, in humans, both acute and chronic stress are thought to be fundamental components in the development and maintenance of addiction, as well as in relapse [132, 133]. The level of impairment in goal-directed behavior also appears to be moderated by substance use severity, as reported in several studies [39, 54, 66]. It needs to be noted, however, that most human studies make use of stimuli and outcomes that are unrelated to drug use (e.g., abstract images, food, money, points). Thus, it possible that the above-described paradigms, as currently implemented, might not be able to demonstrate changes in goal-directed versus habitual behavior in participants with lower levels of drug use, as these may be limited to drug-related behaviors. Therefore, a potential fruitful approach might be to develop experimental designs using drug-related outcomes. Moreover, the propensity toward habitual behavior may also be moderated by the type of substance that is abused, as these also have distinct acute pharmacological effects on brain networks [134]. In conclusion, studies considering moderating contextual factors, such as those already identified as well as acute and chronic stress, in conjunction with inter-individual trait variability in symptom severity, type of substance used and temporal trajectories are necessary to fully understand the extent of the changes in habitual versus goal-directed behavior in human drug addiction [135].

The habit theory of addiction has been criticized by recent reviews (e.g., [15, 45]). Hogarth [45] proposed that drug-seeking behavior in addiction is best explained by excessive goal-directed drug choice instead of habits or compulsions. Based on his finding of reduced devaluation sensitivity with negative mood, he argues that increased responding toward devalued smoking cues is based on the learned expectation that smoking alleviates negative mood. In fact, Hogarth [45] argues that the behavior of participants with SUD remained goal-directed in most devaluation studies, although impairments have been observed in very complex tasks [14, 48, 53]. Similarly, Vandaele and Ahmed [15] conclude that human addiction cannot be explained by the habit account alone. They also criticize the current definition of habits and their experimental (and translational) implementation. They contend that most current experimental habit tasks test whether a potential habitual behavior is executed or not, whereas most everyday behavior, including drug-seeking and -taking, is very complex and consistent with a mixture of habitual/automatic and goal-directed choices. Vandaele and Ahmed [15] importantly suggest that there is a continuous arbitration between the habit and goal-directed systems and that they may be connected in a hierarchical manner during decision-making.

Conclusion

The current findings on habitual versus goal-directed control in addiction and substance use are mixed, and the relevance of habitual behavior for addiction has been questioned by recent reviews [15, 45]. Evidence of impairments in sensitivity to outcome devaluation and contingency degradation have been mainly observed in more severely affected individuals, with discrepancies among SUDs, whereas reduced model-based behavior has been primarily related with alcohol use and treatment outcomes in AUD. Interestingly, two large online studies with individuals with known SUD diagnosis have confirmed strong relationships between reduced model-
based responding and compulsive behavior. Indeed, it has been suggested that the development of compulsions in addiction may require a shift from goal-directed to habitual systems and reduced executive control over mal-adaptive behavior [16].

The reviewed studies evidence that individuals with SUD are capable of behaving in a goal-directed manner in most situations and future studies need to consider more complex interactions between habitual and goal-directed control. As suggested by Hogarth [45], certain behaviors in addiction may well be goal-directed and current evidence does not support that either system is unable to function. Studies integrating recent task developments and considering moderating contextual factors are still required to further investigate the changes in habitual versus goal-directed behavior associated with human drug use [135]. Likewise, it will be necessary to develop tasks that examine the interplay between both systems, and that will allow to determine when, why, and how they are involved in developing and maintaining addictive behaviors.

Conflict of Interest Statement
The authors have no conflicts of interest to declare.

Funding Sources
This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) with grants DO1915/3-1 to ND, SCHL1969/5-1 to FS, CRC940 (project C6) and EN906/8-1 to TE, and CRC/TRR265 (Project-ID 402170461 [115], projects B1 to CE, VAA, FS and TE, B3 to CE, and C4 to FW).

Author Contributions
Nuria Doñamayor: conceptualization, literature search and extraction of relevant literature, preparation and creation of figure, writing of original draft, and writing and editing of final version.
Claudia Ebrahimi: conceptualization, extraction of relevant literature, and writing of original draft.
Viktoria A. Arndt and Franziska Weiss: writing of original draft.
Florian Schlagenhauf: conceptualization and writing of original draft.
Tanja Endrass: conceptualization, writing of original draft, and writing and editing of final version.

References
1 Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci. 2005; 8(11):1481–9.
2 Ostlund SB, Balleine BW. On habits and addiction: an associative analysis of compulsive drug seeking. Drug Discov Today Dis Models. 2008;5(4):235–45.
3 Belin D, Belin-Rauscent A, Murray JE, Everitt BJ. Addiction: failure of control over mal-adaptive incentive habits. Curr Opin Neurobiol. 2013;23(4):564–72.
4 Everitt BJ, Robbins TW. Drug addiction: updating actions to habits to compulsions ten years on. Annu Rev Psychol. 2016;67(1):23–50.
5 Spanagel R. Aberrant choice behavior in alcoholism. Science. 2018;360(6395):1298–9.
6 Redish AD, Jensen S, Johnson A. A unified framework for addiction: vulnerabilities in the decision process. Behav Brain Sci. 2008; 31(4):415–37.
7 Ahmed SH. In search of addiction in the brains of laboratory animals. In: Heather N, Field M, Moss A, Satel S, editors. on behalf of the Addiction Theory Network. Evaluating the brain disease model of addiction hal-03368519; 2021.
8 Heyman GM. Addiction and choice: theory and new data. Front Psychiatry. 2013 2013-May-06;4:31.
9 Heyman GM. How individuals make choices explains addiction’s distinctive, non-eliminable features. Behav Brain Res. 2021;397: 112899.
10 Field M, Heather N, Murphy JG, Stafford T, Tucker JA, Witkiewitz K. Recovery from addiction: behavioral economics and value-based decision making. Psychol Addict Behav. 2020;34(1):182–93.
11 Voon V, Derbyshire K, Rüch C, Irvine MA, Worbe Y, Enander J, et al. Disorders of compulsivity: a common bias towards learning habits. Mol Psychiatry. 2015;20(3):345–52.
12 Balleine BW, O’Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neupropsychopharmacology. 2010;35(1):48–69.
13 Dolan RJ, Dayan P. Goals and habits in the brain. Neuron. 2013;80(2):312–25.
14 Ersche KD, Gillan CM, Jones PS, Williams GB, Ward LHE, Luijten M, et al. Carrots and sticks fail to change behavior in cocaine addiction. Science. 2016;352(6292):1468–71.
15 Vandaele Y, Ahmed SH. Habit, choice, and addiction. Neupropsychopharmacology. 2021; 46(4):689–98.
16 Lüscher C, Robbins TW, Everitt BJ. The transition to compulsion in addiction. Nat Rev Neurosci. 2020;21(5):247–63.
17 Schreiner DC, Renteria R, Gremel CM. Fractionating the all-or-nothing definition of goal-directed and habitual decision-making. J Neurosci Res. 2020;98(6):998–1006.
18 Vandaele Y, Janak PH. Defining the place of habit in substance use disorders. Prog Neuropsychopharmacol Biol Psychiatry. 2018;87:22–32.
19 Watson P, de Wit S. Current limits of experimental research into habits and future directions. Curr Opin Behav Sci. 2018;20:33–9.
20 Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans’ choices and striatal prediction errors. Neuron. 2011;69(6):1204–15.
21 Dickinson A, Balleine BW. Motivational control of goal-directed action. Anim Learn Behav. 1994;22(1):1–18.
22 Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neupropharmacology. 1998;37(4–5):407–19.
23 Lerner TN. Interfacing behavioral and neural circuit models for habit formation. J Neurosci Res. 2020;98(6):1031–45.
24 Dickinson A, Balleine BW. The role of learning in the operation of motivational systems. In: Gallistel R, editor. Stevens’ Handbook of Experimental Psychology. New York, USA: John Wiley & Sons, Inc.; 2002. p. 497–534.
25 Thorndike EL. Animal intelligence: Experimental studies. New York, USA: Macmillan; 1911.
Goal-Directed and Habitual Control in Substance Use

26 Dayan P. Goal-directed control and its antipodes. *Neural Netw.* 2009;22(3):213–9.

27 Dickinson A, Balleine BW. Causal cognition and goal-directed action. In: Heyes C, Huber L, editors. *Vienna series in theoretical biology The evolution of cognition*. Cambridge, USA: The MIT Press; 2000. p. 185–204.

28 Dickinson A, Nicholas DJ, Adams CD. The effect of the instrumental training contingency on susceptibility to reinforcing devaluation. *Q J Exp Psychol.* 1983 Feb;35(1b):35–51.

29 Adams CD, Dickinson A. Instrumental responding following reinforcing devaluation. *Q J Exp Psychol.* 1981 May;33(2b):109–21.

30 Valentin VV, Dickinson A, O’Doherty JP. Determining the neural substrates of goal-directed learning in the human brain. *J Neurosci.* 2007 Apr;27(15):4019–26.

31 Friedel E, Koch SP, Wendt J, Heinz A, Derserno L, Schlangenhaufe F. Devaluation and sequential decisions: linking goal-directed and model-based behavior. *Front Hum Neurosci.* 2014;8:587.

32 Tricomi E, Balleine BW, O’Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci.* 2009; 29(11):2225–32.

33 Morris RW, Quail S, Griffiths KR, Green MJ, Balleine BW. Corticostriatal control of goal-directed action is impaired in schizophrenia. *Biol Psychiatry.* 2015;77(2):187–95.

34 van Timmeren T, Quail SL, Balleine BW, Geurts DEM, Goudriaan AE, van Holst RJ. Intact corticostriatal control of goal-directed action in alcohol use disorder: a Pavlovian-to-instrumental transfer and outcome-devaluation study. *Sci Rep.* 2020;10(1):4949.

35 de Wit S, Niry D, Wariyar R, Aitken MRF, 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 Sep; 29(36):11330–8.

36 de Wit S, Corlett PR, Aitken MRF, 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 Sep; 29(36):11330–8.

37 Corbit LH, Niv H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsolateral striatum. *Biol Psychiatry.* 2012 91(1):385–98.

38 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(3):261–76.

39 Hogarth L, Goal-directed and transfer cue elicited drug-seeking are dissociated by pharmacotherapy: evidence for independent additive controllers. *J Exp Psychol Anim Behav Process.* 2012 Jul;38(3):266–78.

40 Hogarth L, Field M, Rose AK. Phasic transition from goal-directed to habitual control over drug-seeking produced by conflicting reinforcer expectancy. *Addict Biol.* 2013;18(1):88–97.

41 Hogarth L, He Z, Chase HW, Wills AJ, Troisi J, Leventhal AM, et al. Negative mood reverses devaluation of goal-directed drug-seeking favouring an incentive learning account of drug dependence. *Psychopharmacology.* 2015;232(17):3235–47.

42 Schwabe L, Wolf OT. Stress prompts habit behavior in humans. *J Neurosci.* 2009 Jun; 29(22):7191–8.

43 Hogarth L, Attwood AS, Bate HA, Munafò MR. Acute alcohol impairs human goal-directed action. *Biol Psychol.* 2012;90(2):154–60.

44 Veltin E. A laboratory task for induction of mood states. *Behav Res Ther.* 1968;6(4):473–82.

45 Hogarth L. Addiction is driven by excessive goal-directed drug choice under negative affect: translational critique of habit and compulsion theory. *Neuropsychopharmacology.* 2019;45(5):720–35.

46 Hogarth L, Lam-Cassettari C, Pacitti H, Currah T, Mahlberg J, Hartley L, et al. Impaired goal-directed treatment in treatment-seeking drug users indexed by outcome-devaluation and Pavlovian to instrumental transfer: critique of habit theory. *Eur J Neurosci.* 2019;50(3):2513–25.

47 de Wit S, Corlett PR, Aitken MRF, 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 Sep; 29(36):11330–8.

48 Sjoerds Z, de Wit S, van den Brink W, Robbins TW, Beekman ATF, Penninx BWJH, et al. Behavioral and neuroimaging evidence for over-reliance on habit learning in alcohol-dependent patients. *Transl Psychiatry.* 2013;3(12):E337.

49 Tanaka SC, Balleine BW, O’Doherty JP. Calculating consequences: brain systems that encode the causal effects of actions. *J Neurosci.* 2008 Jun;28(26):6750–5.

50 de Wit S, Standing HR, DeVito EE, Robinson OJ, Ridderinkhof KR, Robbins TW, et al. Reliance on habits at the expense of goal-directed control following dopamine precursor depletion. *Psychopharmacology.* 2012;219(2):621–31.

51 Worbe Y, Savulich G, de Wit S, Fernandez-Egea E, Robbins TW. Tryptophan depleting promotes habitual over goal-directed control of appetitive responding in humans. *Int J Neuropsychopharmacol.* 2015;18(10):pyv013.

52 Gillan CM, Morein-Zamir S, Urcelay GP, Sule A, Voon V, Apergis-Schoute AM, et al. Enhanced avoidance habits in obsessive-compulsive disorder. *Biol Psychiatry.* 2014;75(8):631–8.

53 Lim T-V, Cardinal RN, Savulich G, Jones PS, Moustafa AA, Robbins TW, et al. Impairments in reinforcement learning do not explain enhanced habit formation in cocaine use disorder. *Psychopharmacology.* 2019; 236(8):2359–71.

54 Luijten M, Gillan CM, de Wit S, Franken IHA, Robbins TW, Ersche KD. Goal-directed and habitual control in smokers. *Nicotine Tob Res.* 2020;22(2):188–95.

55 Heatherton TF, Kozlowski LT, Frecker RC, Fagerstrom K-O. The Fagerström test for nicotine dependence: a revision of the fagerstrom tolerance questionnaire. *Addiction.* 1991;86(9):1119–27.

56 Gillan CM, Apergis-Schoute AM, Morein-Zamir S, Urcelay GP, Sule A, Fineberg NA, et al. Functional neuroimaging of avoidance habits in obsessive-compulsive disorder. *Am J Psychiatry.* 2015;172(3):284–93.

57 Lijeholm M, Tricomi E, O’Doherty JP, Balleine BW. Neural correlates of instrumental contingency learning: differential effects of action–reward conjunction and disjunction. *J Neurosci.* 2011;31(7):2474–80.

58 Chatlish DL, Neunaber DJ, Wasserman EA. Response-outcome contingency: behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learn Motiv.* 1985;16(1):1–34.

59 Shanks DR, Dickinson A. Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Mem Cogn.* 1991;19(4):353–60.

60 Morris RW, Cyron C, Green MJ, Le Pelley ME, Balleine BW. Impairments in action–outcome learning in schizophrenia. *Transl Psychiatry.* 2018;8(1):54.

61 Vaghi MM, Cardinal RN, Apergis-Schoute AM, Fineberg NA, Sule A, Robbins TW. Action–outcome knowledge dissociates from behavior in obsessive-compulsive disorder following contingency degradation. *Biol Psychiatry Cogn Neuroimaging.* 2019;4(2):208–9.

62 Ersche KD, Lim TV, Murley AG, Rua C, Vaghi MM, White TL, et al. Reduced glutamate turnover in the putamen is linked with automatic habits in human cocaine addiction. *Biol Psychiatry.* 2021;89(10):970–9.

63 Sutton RS, Barto AG. *Reinforcement learning: an introduction.* Cambridge, USA: MIT Press; 1998.

64 Kool W, Cushman FA, Gershman SJ. When does model-based control pay off? *PLoS Comput Biol.* 2016;12(8):e1005090.

65 Berghäuser T, Bensmann W, Zink N, Endrass T, Bente C, Stock A-K. Alcohol hangover does not alter the application of model-based and model-free learning strategies. *J Clin Med.* 2020;9(5):1453.

66 Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife.* 2016;5:E11305.

67 Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND. Model-based choices involve prospective neural activity. *Nat Neurosci.* 2015;18(5):767–72.
76. Gläscher J, Daw ND, Dayan P, O'Doherty JP.
Model-free decisions in alcohol dependence. Biol Psychiatry. 2005; 85(5):498–503.

77. Patzelt EH, Janak PH. Habitual alcohol seeking: neural bases and possible relations to alcohol use disorders. Alcohol Clin Exp Res. 2016; 40(7):1380–9.

78. Wopshol MJ, Pfefferbaum A. Striatal and frontal model-based signatures are reduced in young adults: a 3-year prospective study. J Cogn Neurosci. 2012; 226(2):579–91.

79. Patzelt EH, Kool W, Millner AJ, Gershman SJ. Incentives boost model-based control across model-free decisions in alcohol dependence. J Stud Alcohol. 1987;48(5):483–91.

80. Crabbe JC, Harris RA, Koob GF. Preclinical studies of alcohol binge drinking. Ann N Y Acad Sci. 2011;1216(1):24–40.

81. Makris N, Oscar-Berman M, Jaffin SK, Hodge SM, Kennedy DN, Caviness VS, et al. Decreased volume of the brain reward system in alcoholism. Biol Psychiatry. 2008;64(3):192–202.

82. Sullivan EV, Deshmukh A, De Rosa E, Rosenbloom MJ, Pfefferbaum A. striatal and forebrain nuclei volumes: contribution to motor function and working memory deficits in alcoholism. Biol Psychiatry. 2005; 57(7):768–76.

83. Deserno L, Huys QJM, Boehme R, Buchert R, Heinze H-J, Grace AA, et al. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. Proc Natl Acad Sci U S A. 2015 Jan 20;112(5):1595–600.

84. Gläscher J, Daw ND, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 2010;66(4):585–95.

85. Davies SJ, Pandit SA, Feeney A, Stevenson BJ, Kerwin RW, Nutt DJ, et al. Is there cognitive impairment in clinically ‘healthy’ abstinent alcohol dependence? Alcohol Alcohol. 2005;40(6):498–503.

86. Moselly HF, Georgiou G, Kahn A. Frontal lobe changes in alcoholism: a review of the literature. Alcohol Alcohol. 2001;36(5):357–68.

87. Ariani G, Diedrichsen J. Sequence learning in humans. Curr Opin Behav Sci. 2018;20(1):110–6.

88. Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. Trends Cogn Sci. 2010; 14(5):208–15.

89. Du Y, Krakauer JW, Haith AM. The relationship between habits and motor skills in humans. Trends Cogn Sci. 2022;26(5):371–87.

90. Dayan P, Abbott LF. Theoretical Neuroscience: the cellular and computational basis of learning and behavior. The MIT Press; 2001.

91. Dayan P, Abbott LF. Theoretical Neuroscience: the cellular and computational basis of learning and behavior. The MIT Press; 2001.

92. Doyon J, Gabitov I, Vedant S, Lunghi O, Boulin A. Current issues related to motor sequence learning in humans. Curr Opin Behav Sci. 2018;20:89–97.

93. Magon S, Pfister A, Laura G, Lüthi M, Papadopoulo A, Kappos L, et al. Short timescale modulation of cortical and cerebellar activity in the early phase of motor sequence learning: an fMRI study. Brain Imaging Behav. 2020;14(6):2159–75.

94. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookenheimer SY. The neural correlates of motor skill automaticity. J Neurosci. 2005;25(22):5356–64.

95. McKim TH, Bauer DJ, Boettiger CA. Addiction history associates with the propensity to form habits. J Cogn Neurosci. 2016;28(7):1024–38.

96. McKim TH, Dove SJ, Robinson DL, Fröhlich F, Boettiger CA. Addiction history moderates the effect of prefrontal 10-Hz transcranial alternating current stimulation on habitual action selection. J Neurophysiol. 2021; 125(3):768–80.

97. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookenheimer SY. The neural correlates of motor skill automaticity. J Neurosci. 2005;25(22):5356–64.

98. McKim TH, Dove SJ, Robinson DL, Fröhlich F, Boettiger CA. Addiction history moderates the effect of prefrontal 10-Hz transcranial alternating current stimulation on habitual action selection. J Neurophysiol. 2021; 125(3):768–80.

99. Gardner B. A review and analysis of the use of “habit” in understanding, predicting and influencing health-related behaviour. Health Psychol Rev. 2015;9(3):277–95.

100. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci. 2005;8(12):1794–11.

101. Hardwick RM, Forrence AD, Krakauer JW, Haith AM. Time-dependent competition between goal-directed and habitual response preparation. Nat Hum Behav. 2019; 3(12):1252–62.

102. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookenheimer SY. The neural correlates of motor skill automaticity. J Neurosci. 2005;25(22):5356–64.

103. Hardwick RM, Forrence AD, Krakauer JW, Haith AM. Time-dependent competition between goal-directed and habitual response preparation. Nat Hum Behav. 2019; 3(12):1252–62.

104. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookenheimer SY. The neural correlates of motor skill automaticity. J Neurosci. 2005;25(22):5356–64.

105. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nat Rev Neurosci. 2006;7(6):464–76.

106. Doll BB, Simon DA, Daw ND. The ubiquity of model-based reinforcement learning. Curr Opin Neurobiol. 2012;22(6):1075–81.

107. Sjoerds Z, Dietrich A, Deserno L, de Wit S, Villringer A, Heinze H-J, et al. Slips of action and sequential decisions: a cross-validation study of tasks assessing habitual and goal-directed action control. Front Behav Neurosci. 2016 Dec 20;10:234.

108. Gillan CM, Otto AR, Phelps EA, Daw ND. Model-based learning protects against forming habits. Cogn Affect Behav Neurosci. 2015;15(3):323–36.

109. Pataki LK, Knowlton BJ. Subregional specificity in human striatal habit learning: a meta-analytic review of the fMRI literature. Curr Opin Behav Sci. 2018;20:75–82.

110. Ceceli AO, Tricomi E. Habits and goals: a motivational perspective on action control. Cogn Opin Behav Sci. 2018;20:110–6.
Goal-Directed and Habitual Control in Substance Use

111 Miller KJ, Ludvig EA, Pezzulo G, Shenhav A. Realigning models of habitual and goal-directed decision-making. In: Morris RW, Bornstein AM, Shenhav A, editors. Goal-directed decision making: computations and neural circuits. Academic Press; 2018. p. 407–28.

112 Miller KJ, Shenhav A, Ludvig EA. Habits without values. Psychol Rev. 2019;126(2): 292–311.

113 Stephan KE, Mathys C. Computational approaches to psychiatry. Curr Opin Neurobiol. 2014;25:85–92.

114 Brown VM, Chen J, Gillan CM, Price RB. Improving the reliability of computational analyses: model-based planning and its relationship with compulsivity. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020;5(6): 601–9.

115 Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Multiple memory systems as substrates for multiple decision systems. Neurobiol Learn Mem. 2015;117:4–13.

116 Doll BB, Shohamy D, Daw ND. Working-memory capacity protects model-based learning from stress. Proc Natl Acad Sci U S A. 2013;110(52):20941–6.

117 Doll BB, Bath KG, Daw ND, Frank MJ. Variability in dopamine genes dissociates model-based and model-free reinforcement learning. J Neurosci. 2016;36(10):52:1211–22.

118 Akam T, Costa R, Dayan P. Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. PLoS Comput Biol. 2015;11(12):E1004648.

119 Desfouli A, Balleine BW. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. PLoS Comput Biol. 2013;9(12):E1003364.

120 Smithenraar P, FitzGerald T, Romei V, Wright N, Dolan R. Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. Neuron. 2013;80(4):914–9.

121 Feher da Silva C, Hare TA. Humans primarily use model-based inference in the two-stage task. Nat Hum Behav. 2020;4(10): 1053–66.

122 Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA. Motivational control after extended instrumental training. Anim Learn Behav. 1995;23(2):197–206.

123 Adams CD. Variations in the sensitivity of instrumental responding to reinforcer devaluation. Q J Exp Psychol. 1982;34(2b):77–98.

124 Thrailkill EA, Trask S, Vidal P, Alcalá JA, Bouton ME. Stimulus control of actions and habits: a role for reinforcer predictability and attention in the development of habitual behavior. J Exp Psychol Anim Learn Cogn. 2018;44(4):370–84.

125 Garr E, Bushra B, Tu N, Delamater AR. Goal-directed control on interval schedules does not depend on the action–outcome correlation. J Exp Psychol Anim Learn Cogn. 2020;46(1):47–64.

126 Garr E, Padovan-Hernandez Y, Janak PH, Delamater AR. Maintained goal-directed control with overtraining on ratio schedules. Learn Mem. 2021 Dec;28(12):435–9.

127 de Wit S, Kindt M, Knot SL, Verhoeven AAC, Robbins TW, Gasull-Camos J, et al. Shifting the balance between goals and habits: five failures in experimental habit induction. J Exp Psychol Gen. 2018;147(7):1043–65.

128 Haith AM, Pakpour J, Krakauer JW. Independence of movement preparation and movement initiation. J Neurosci. 2016;36(10):3007–15.

129 Schwabe L, Höffken O, Tegenthoff M, Wolf OT. Preventing the stress-induced shift from goal-directed to habit action with a β-adrenergic antagonist. J Neurosci. 2011;31(47):17317–25.

130 Radenbach C, Reiter AMF, Engert V, Sjöerds Z, Villringer A, Heinez H-J, et al. The interaction of acute and chronic stress impairs model-based behavioral control. Psychoneuroendocrinology. 2015;53:268–80.

131 Bossert JM, Marchant NJ, Calu DJ, Shaham Y. The reinstatement model of drug relapse: recent neurobiological findings, emerging research topics, and translational research. Psychopharmacology. 2013;229(3):453–76.

132 Koob GF. A role for brain stress systems in addiction. Neuro. 2008;50(1):11–34.

133 Koob GF. Negative reinforcement in drug addiction: the darkness within. Curr Opin Neurobiol. 2013;23(4):559–63.

134 Lüscher C, Ungless MA. The mechanistic classification of addictive drugs. PLoS Med. 2006;3(11):E437.

135 Heinz A, Kiefer F, Smolka MN, Endrass T, Beste C, Beck A, et al. Addiction Research Consortium: losing and regaining control over drug intake (ReCoDe) – from trajectories to mechanisms and interventions. Addict Biol. 2020;25(2):E12866.