

How Does Drug Use Shift the Balance Between Model-Based and Model-Free Control of Decision Making?

Kurt M. Fraser and Patricia H. Janak

What influences the decisions we make? Do we construct rich maps of possible actions and their outcomes to select the best future option, using so-called model-based strategies, or do we default to something more automatic that reflects what generally has been best based on past experience, processes that are referred to as model-free strategies? These questions have motivated experiments designed to tease apart differences in these two decision-making strategies. In humans, a major advance toward addressing this question came from the development of a two-stage decision-making task that allows simultaneous determination of the relative weighting of the model-based and model-free strategies in use (1). However, the absence of an easy-to-implement analogous rodent task that mirrors the features and outcomes of the human two-stage task was a major roadblock for translational mechanistic studies of decision making. This was a critical missing piece, because dysfunctions in model-based behaviors have been associated with psychiatric illnesses like addiction (2). As a result, answering the question of “the chicken or the egg” in the dysfunctions of model-based control in people with addiction has been unresolved.

To overcome this, Groman *et al.* (3,4) developed a novel two-stage decision-making task for rats that is rapidly acquired and that mimics the human task. In this task, rats press levers to gain the opportunity to nosepoke in ports for reward—earning a reward requires a sequence of two choices. The choice between two levers is the first stage, and the choice between two lighted nosepoke ports is the second stage. Each lever is associated with a high probability of access to one pair of nosepoke ports and a low probability of access to a different pair of nosepoke ports. Thus, on any given trial, upon pressing a lever, the rat experiences a relatively common or relatively rare transition from stage one to stage two depending on which pair of nosepokes lights up. Once a pair of nosepoke ports is available, the rat’s choice of which port to respond to is shaped by distinct probabilities of reward availability that shift over session blocks. These probabilities assigned to each nosepoke port are weighted so that the common transition after pressing one lever has a higher probability of reward than the common transition after pressing the other lever.

Earning a reward in this task is not simple—how do rats go about it? Animals exhibiting an exclusively model-free strategy should stick to their past rewarded choice on the current trial. In contrast, animals using a model-based strategy should alter their current choice based both on the previous trial’s outcome and on whether the first stage lever they chose led to the

common or rare set of nosepoke ports associated with that choice. Amazingly, Groman *et al.* (3,4) report that rats trained in this task exhibit contributions of both model-based and model-free strategies to their decision making, just as humans do on the equivalent task. This contrasts with other reports of a rodent two-step task in which rats use primarily model-based strategies after a long training period (5).

The use of the two-stage task developed by Groman *et al.* (3,4) allows for assessment of the degree to which both model-free and model-based processes contribute to decision making. This is a major advance over typical experimental approaches in animals such as reward devaluation and contingency degradation to test if behavior is sensitive to changes in reward value; if it is not sensitive, that behavior is taken as consistent with the use of model-free strategies (6). These approaches allow only for an either/or assessment rather than a relative weighting of each strategy in cases where both might be used. This is important in the context of addiction, where prominent hypotheses hold that behavior transitions from being model-based to model-free, contributing to compulsive drug intake (2,7). However, it has proven difficult to disentangle in human subjects whether observed behavior reflects strategy predispositions, whether drug use alters either model-based or model-free processes, or whether both occur.

To directly assess this, Groman *et al.* (4) use their two-stage task in rats to determine model-free and model-based contributions to behavior before and after a stint of drug taking, and report these findings in the current issue of *Biological Psychiatry*. Rats were allowed to self-administer methamphetamine, a gold standard procedure to model volitional drug use in animals. To quantify individual variability in drug self-administration, Groman *et al.* (4) obtained estimates of two parameters: one relating to the initial reinforcing properties of intravenous methamphetamine, and another relating to the escalation of intake across the 2 weeks of drug taking. This approach allowed a comparison of previous contributions of model-free and model-based strategies to decision making for a natural reward, sugar, with each of these metrics of drug taking. Interestingly, Groman *et al.* (4) found a negative relationship between the predrug degree of model-free strategy and the amount of methamphetamine administered at the beginning of self-administration training. Digging into the data, Groman *et al.* (4) were able to determine that subjects with low model-free behavioral control were actually less sensitive to receipt of food reward in the two-stage decision making task, yet ended up taking more methamphetamine. The transition

SEE CORRESPONDING ARTICLE ON PAGE 936

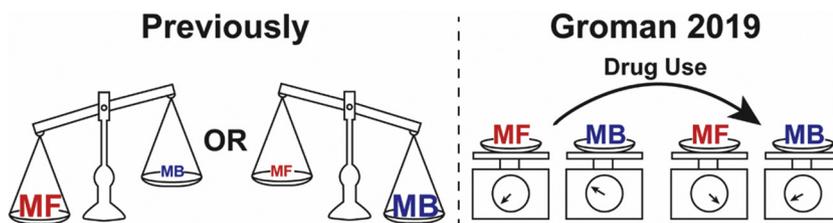


Figure 1. Balance between model-free (MF) and model-based (MB) control in decision making. Previous studies were primarily able to assess the contributions of either MB or MF processes to behavior. In contrast, Groman *et al.* (4) show a shared contribution of both MF and MB strategies before drug use, show that the MF system is related to initial drug reinforcement, and show that these processes are independently shifted after drug use.

from recreational drug use to addiction is proposed to involve changes in the balance of model-free and model-based decision making, so one may have expected that more model-free behavioral control and/or less model-based behavioral control would have predicted escalating drug intake, acting as predispositions or risk factors for addiction vulnerability (7). However, there was no relationship between either model-free or model-based strategies and escalation of methamphetamine intake over 2 weeks of drug taking, which is surprising given the evidence from previous studies using either/or evaluations of model-free and model-based strategies.

Next, the rats were retested in the two-stage task to relate drug taking and the influence of drug exposure on the weighting of model-based and model-free processes. Previous research suggested that drug use in animals promotes a shift to model-free influenced behavior [reviewed in (6,8,9)]. However, Groman *et al.* (4) found that this was not the case and that contributions of both model-free and model-based strategies to decision making were decreased after drug use. This pattern could arise from a deficit in a common arbitrating system that would predict an equivalent shift in both processes or instead could result from shifts in two independent systems determining individually the weighting of model-based and model-free influences. Surprisingly, Groman *et al.* (4) find evidence for the latter, because the degree of change for model-free and for model-based contributions to decision making after drug use were unrelated at the level of the individual subject. This pattern of independent change in these two systems suggests the need for a reevaluation of the idea that behavioral control is a winner-take-all competition between these two (Figure 1). This finding contrasts with findings from previous human imaging studies suggesting that there may exist an arbitrator that controls the preference for a model-based or model-free strategy (10).

Groman *et al.* (4) highlight the utility of well-designed preclinical behavioral procedures in advancing our understanding of complex cognitive processes and their dysregulation. The use of this procedure has provided tantalizing evidence that individual variation in model-free tendencies may impact initial drug taking, and that chronic drug taking may impair the appropriate use of both model-free and model-based decision-making processes. Finally, because the use of each strategy was independently impaired by methamphetamine, Groman *et al.* (4) have provided evidence that these strategies are not always in a trade-off or winner-take-all situation. In the future, it will be of interest to measure behavior in the rodent two-stage task after extensive experience with drug taking beyond the relatively short 2-week period used here, to determine if the described changes are

stable or if further alterations occur. In addition, comparisons with other drugs of abuse will be needed to test for generality of these findings.

The rodent two-stage task will also prove a boon to understanding how different neural circuits contribute to model-based and model-free decisions. In previous work, Groman *et al.* (3) found that dopamine tone in the orbitofrontal cortex and nucleus accumbens was related to the degree of model-based influence on behavior in the two-stage task. The encoding of actions and outcomes is disrupted in both of these regions after drug use (8), so it will be an exciting future direction to understand the contributions of these structures to ongoing behavior in the two-stage task. Understanding how dopamine in these regions, as well as in the more commonly studied dorsal striatal territories, contributes to these diverse strategies will also be an important area for further study, because dysregulated dopamine signaling is a core feature of addiction. The approach of Groman *et al.* (4) provides a framework for investigating dysregulated decision making in a variety of disorders and for advancing findings along the pipeline from research to treatment for those with psychiatric illness.

Acknowledgments and Disclosures

This work was supported by National Institute on Drug Abuse Grant Nos. F31 DA046136 (to KMF) and R01 DA035943 (to PHJ) and National Institute on Alcohol Abuse and Alcoholism Grant Nos. R01 AA026306 and R01 AA027213 (to PHJ).

The authors report no biomedical financial interests or potential conflicts of interest.

Article Information

From the Department of Psychological and Brain Sciences (KMF, PHJ), Krieger School of Arts and Sciences, and The Solomon H. Snyder Department of Neuroscience (PHJ), Johns Hopkins School of Medicine, Baltimore, Maryland.

Address correspondence to Patricia H. Janak, Ph.D., Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218; E-mail: patricia.janak@jhu.edu; or Kurt M. Fraser, MA, Department of Psychological and Brain Sciences, Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218; E-mail: kfraser3@jhu.edu.

Received Apr 15, 2019; accepted Apr 15, 2019.

References

- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215.
- Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, *et al.* (2015): Disorders of compulsivity: A common bias towards learning habits. *Mol Psychiatry* 20:345–352.

3. Groman SM, Massi B, Mathias SR, Curry DW, Lee D, Taylor JR (2019): Neurochemical and behavioral dissections of decision-making in a rodent multistage task. *J Neurosci* 39:295–306.
4. Groman SM, Massi B, Mathias SR, Lee D, Taylor JR (2019): Model-free and model-based influences in addiction-related behaviors. *Biol Psychiatry* 85:936–945.
5. Miller KJ, Botvinick MM, Brody CD (2017): Dorsal hippocampus contributes to model-based planning. *Nat Neurosci* 20:1269–1276.
6. Vandaeele Y, Janak PH (2018): Defining the place of habit in substance use disorders. *Prog Neuropsychopharmacol Biol Psychiatry* 87(Pt A):22–32.
7. Everitt BJ, Robbins TW (2016): Drug addiction: Updating actions to habits to compulsions ten years on. *Annu Rev Psychol* 67:23–50.
8. Stalnaker TA, Takahashi Y, Roesch MR, Schoenbaum G (2009): Neural substrates of cognitive inflexibility after chronic cocaine exposure. *Neuropharmacology* 56(Suppl 1):63–72.
9. Barker JM, Taylor JR (2014): Habitual alcohol seeking: Modeling the transition from casual drinking to addiction. *Neurosci Biobehav Rev* 47:281–294.
10. Lee SW, Shimojo S, O’Doherty JP (2014): Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81:687–699.