Addiction, agency, and the limits of the 'Epistemic Prerequisites Model'

Zoey Lavallee

How does addiction alter agency in ways that undermine a person's ability to reliably abstain from drug use, even in the face of severe consequences? Arthur Krieger's (2024) article addresses this problem by developing an account of how behavioral control over drug use is eroded in addiction and what is required to regain it. He introduces the Epistemic Prerequisite Model (EPM), arguing that addiction undercuts a person's *first-order ability* (ability$_1$) to reliably abstain from drug use, but that (most) addicted people retain the *second-order ability* (ability$_2$) to reliably abstain. That is, they can still develop what Krieger calls *epistemic prerequisites* (EPs), which he argues are individually necessary and jointly sufficient for reliable abstinence. These EPs are (1) addiction-specific metacognitive skills – noticing and intervening on implicit, automatic thoughts that exaggerate the value of drug use, and which are prolific in addiction; and (2) a particular kind of self-knowledge – theoretical knowledge of one's cues for addictive behavior and how to avoid these. On this view, which builds on Nordenfelt's (2007) delineation between first and second order ability, ability$_2$ to reliably abstain can become ability$_1$, through cultivating these metacognitive skills and self-knowledge.

Krieger's approach has significant merits. He offers an insightful analysis of some of the automated patterns of thought and behavior that addiction can involve, and I particularly appreciate his attention to how one's social environment plays a role in enabling or undermining addiction recovery. By his account, the ability to acquire the metacognitive skills and knowledge needed for reliable abstinence is largely dependent upon the social and economic resources available to an individual. Like Krieger, I take it that a good philosophical account of how addiction shapes agency should recognize the heterogeneous and interacting factors that influence addictive behavior, including neurobiological, psychological, social, and economic factors. That said, in this commentary I want to suggest that the EPM may provide too narrow an account of the ways that addiction disrupts agency over time, and consequently, of the strategies through which control over drug use can be regained. I am of the view that, when an addicted person's lifeworld is conditioned by things like social and economic deprivation, or disproportionate precarity or distress, these factors undermine behavioral control not only by, as Krieger proposes, "profoundly impair[ing] cognitive and metacognitive functioning" needed to cultivate particular kinds of practical and theoretical knowledge, but through multiple pathways.

To begin, I summarize Krieger's explanation of how addiction undermines behavioral control. He first argues that the neuroadaptations caused by the effects of long-term drug use on the brain's reward-learning system can undermine behavioral control by rendering individuals hypersensitive to environmental cues that have become associated with their drug use (Grisel 2019; Berridge & Robinson 2011). As Krieger describes it, encountering these cues – whether the sight of a needle, the smell of alcohol, or the presence of certain friends – triggers automatic, implicit thought processes, primarily through "cue-induced" (Pickard 2024) cravings, which bias a person toward drug use. Drawing on Berridge and Robinson (2011), these cravings are not instrumental desires aimed at the pleasurable effects of drug use or at some other psychological goal, but *intrinsic* desires for the drug itself, operating outside the agent's reflective control. It is in this way that 'liking' and 'wanting' come apart in addictive behavior. Krieger further argues that addiction undermines $ability_1$ via a second pathway: distorted low-salience evaluative thoughts. Following Sripada's (2022) "distortion model," he suggests that along with the effects of cue-induced cravings on thought and behavior, addiction overwhelms individuals with a high frequency of distorted thoughts that exaggerate the value of drug use, while downplaying the value of abstinence. Cognitive distortion, then, in addition to cue-reactivity, can systematically erode a person's $ability_1$ to reliably control drug use.

While I agree with Krieger that cue-induced cravings and distorted evaluative thinking can play significant roles in addiction and impede efforts at recovery, it is not clear that these mechanisms exhaust the routes through which behavioral control is undermined. Much drug use in addiction is goal-directed, meaning people often use drugs to serve specific personal purposes (e.g., Müller & Schumann, 2011; Lavallee & Osler 2024; Flanagan 2019). That is, drug use in addiction is not just a matter of being motivated by cue-induced cravings or distorted thoughts that exaggerate the value of drugs, rather it can also effectively serve identifiable functions; for example, regulating distressing emotions (e.g., using opioids to dull pain or alcohol to cope with social anxiety); escaping withdrawal symptoms (e.g., taking heroin or benzodiazepines to avoid painful withdrawal); facilitating social connection (e.g., stimulant use in party or work environments); or constructing identity (e.g., 'addict' as an identity or meaningful social role). Recognizing this aspect of addiction, Pickard (2024, 18) distinguishes between cue-induced intrinsic cravings, as described by Berridge and Robinson (2011) (see, also Robinson & Berridge 1993), and goal-focused cravings – ordinary but powerful desires that drug use continues to satisfy in addiction. Unlike intrinsic cravings, which operate via automatic reward-learning processes and environmental cues, goal-directed cravings are shaped by broader environmental and psychosocial factors (Lavallee 2020). For instance, a person who grows up experiencing chronic poverty, trauma, and social marginalization may come to rely on drug use as one of the only reliable and available strategies for coping with pervasive emotional distress. These instrumental desires, and the learned reliance on drug use to satisfy them, are shaped by environmental and psychosocial factors largely beyond individual control. In such cases, control over drug use is constrained not just by distorted cognition or cue-sensitivity, but by the individual's lack of viable alternative ways to meet fundamental psychological and social needs (Lavallee 2023; Lavallee & Osler 2024).

To be clear, Krieger does acknowledge that psychosocial and economic factors play a key part in addiction. However, he treats these as background conditions that influence the "automated patterns of thought and behavior" that he argues undermine reliable abstinence, rather than direct mechanisms of behavioral constraint. At one point, Krieger suggests that people in deprived environments may be more likely to turn to drugs to seek pleasure or pursue subjective well-being, and he notes that social and structural factors help to explain "what causes a particular individual to overuse to the point of addiction". This arguably gestures to how the kinds of environmentally situated, goal-focused cravings that I have in mind can undermine the ability to abstain from drug use, but he does not fully integrate the role of these instrumental desires into his framework. Ultimately Krieger argues that "addiction undermines ability$_1$ to reliably abstain by one of two routes": cue-exposure resulting in automatic behavior and low-salience cognitive distortions. If many addicted individuals' ability to reliably stop using drugs is impeded by entrenched reliance on drug use as an emotional and psychological scaffold, then I suggest we need an account of loss of ability$_1$ that extends beyond the two routes that Krieger identifies.

In my own view, the exercise of agency in addiction becomes systematically inclined toward drug-seeking and drug-taking behaviors through multiple pathways – including through the ways that addiction diachronically shapes an agent's values, instrumental desires, cares and concerns, and perceived affordances for alternative actions (Lavallee 2020, 2023; Lavallee & Osler 2024). Similarly, Henden (2023) argues that autonomy can be impaired in addiction not only insofar as addicted people are propelled by automated behavior or fail to exercise effortful control over cue-reactivity. By considering the emotion-related roles that drugs can play in addiction, Henden proposes that goal-directed drug use– for its emotional effects – can, over time, result in a reorientation of a person's values, making "the values associated with the emotion-related role drug use has in the addict's life, central to their conception of the good" (2023, 2). In such cases, "loss of autonomy may therefore consist, not so much in a lack of ability to do what they truly want to do, but more in a lack of ability to truly want *a different life*, one that does not involve the regular consumption of addictive drugs" (2023, 8). Krieger articulates important pieces of the puzzle of how behavioral control is undermined in addiction, but I would argue that a more pluralistic approach is needed to account for the full range of ways that one's ability to reliably sustain recovery from addiction is impeded, including by goal-directed cravings and their psychosocial and structural determinants.

This, in turn, has implications for how we think about addiction recovery. Because Krieger's analysis focuses primarily on addictive behaviors that are explained by low-salience distorted cognition and cue-sensitivity, his account of reliable abstinence consists of two corresponding epistemic strategies. Addiction-specific metacognitive skills are needed to intervene on cognitive distortions, and to address automatic behavior driven by cue-sensitivity, one needs knowledge of their own addiction cues and how to avoid them. If Krieger's account of loss of ability$_1$ presents a non-exhaustive list of agency-undermining mechanisms in addiction, then the epistemic prerequisites he outlines may not always be sufficient for recovery – though they may at least sometimes be. Scaffolding ability$_1$ will typically require addressing environmental and psychosocial determinants that inform goal-focused cravings and that limit or enable alternative possibilities for satisfying these desires, beyond drug use. Plausibly, a person could undergo addiction treatment programs, and cultivate the metacognitive skills and self-knowledge that Krieger proposes to address cue-reactivity and cognitive distortion in addiction, yet because of powerful goal-directed cravings informed by their environment, continue to engage in "the pattern of drug choices characteristic of addiction, however compromised or distorted [...] these may prove to be" (Pickard 2022, 998).

Ultimately, Krieger's epistemic model offers valuable insights into how addiction disrupts behavioral control. I hope that my points of commentary constructively contribute to his analysis, inviting further consideration of how alternative mechanisms – such as goal-directed cravings and diachronic shifts in values – fit within his framework, alongside the broader structural constraints they reflect. I take it that Krieger and I are in agreement about many things here, including that people with addictions, in order to recover, will most often need both new coping skills, and an environment that enables changed behavior by supplying better alternatives to ongoing drug use.

Bibliography

Berridge, K. C., & Robinson, T. E. (2011). Drug addiction as incentive sensitization. In Poland, J. & Graham, G. (eds.), *Addiction and responsibility*, 21-54. Cambridge: The MIT Press.

Flanagan, O. (2018). Identity and addiction. In *The Routledge handbook of philosophy and science of addiction* (pp. 77-89). Routledge.

Grisel, J. (2019). *Never enough: The neuroscience and experience of addiction.* New York: Doubleday.

Henden, E. (2023). Addiction and autonomy: Why emotional dysregulation in addiction impairs autonomy and why it matters. *Frontiers in Psychology*, 14, 1081810.

Krieger, A. (2024). The epistemic prerequisites of reliable abstinence in addiction. *Philosophy, Psychiatry, & Psychology.*

Lavallee, Z. (2020). Addictive craving: There's more to wanting more. *Philosophy, Psychiatry, & Psychology*, 27(3), 227-238.

Lavallee, Z. (2023). Affective scaffolding in addiction. *Inquiry*, 1-29.

Lavallee, Z., & Osler, L. (2024). Affordances and the shape of addiction. *Philosophy, Psychiatry, & Psychology*, 31(4), 379-395.

Müller, C., & Schumann, G. (2011). Drugs as instruments: A new framework for non-addictive psychoactive drug use. *Behavioral and Brain Sciences*, 34(6), 247–259.

Nordenfelt, L. (2007). Rationality and compulsion: Applying action theory to psychiatry. *Oxford University Press*, USA.

Pickard, H. (2024). Craving for drugs. *Mind & Language*, 39(5), 598-626.

Pickard, H. (2022). Is addiction a brain disease? A plea for agnosticism and heterogeneity. *Psychopharmacology*, 1-15.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, 18(3), 247–291.

Sripada, C. (2022). Impaired control in addiction involves cognitive distortions and unreliable self-control, not compulsive desires and overwhelmed self-control. *Behavioural Brain Research*, 418, 113639.