A generalizable computational mechanism underlying the interaction between momentary craving and decision-making

Kaustubh R. Kulkarni[1, 2], Laura A. Berner[1, 2], Daniela Schiller[1, 2, 3], Vincenzo G. Fiore[1, 2], Xiaosi Gu[1, 2, 3]

[1]*Center for Computational Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY*
[2]*Nash Family Dept of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY*
[3]*Dept of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY*

Abstract

All addictive disorders consist of two key, interacting features: persistent subjective craving and maladaptive choice behaviors. However, the computations underlying the bidirectional relationship between craving and decision-making remain unknown. Here, we tested a hypothesis that momentary craving and value-based decision-making dynamically influence each other across addictive domains. We sampled momentary craving as three groups of participants (alcohol drinkers, cannabis users, and binge eaters; total n=202) performed a probabilistic learning task in which they received either monetary or group-specific addictive cue rewards. Using computational modeling, we found that momentary craving dynamically biased towards faster learning during addictive but not monetary decision-making across all groups. Conversely, we also found that prediction errors modulated perceived craving. Computational parameters underlying these latent mechanisms were reliably predictive of clinical severity. These findings provide a computational mechanism for the interaction between craving and choice behavior that is generalizable across addictive domains.

Humans can become addicted to a variety of stimuli, ranging from foods to drugs. Two elements are essential across addictive processes: *craving*, the strong subjective desire for the stimulus; and *decision-making*, the objective and observable choices associated with the addiction. While it is universally acknowledged that craving and addictive decision-making are tightly intertwined, the computational mechanism underlying such interaction is not clear. Historically, craving is considered a passively conditioned response to drugs (Tiffany, 1995) that also involves neuroadaptive response resulting from a shift in homeostatic state following persistent substance use (Robinson & Berridge, 1993). As such, the cue-reactivity framework – the most influential characterization of craving – emphasizes the elicitation of craving in response to learned addictive cues (Carter & Tiffany, 1999; Drummond 2000; Ekhtiari et al., 2022; Garrison & Potenza, 2014; Chase et al., 2011). This framework has been empirically tested across a number of addictive disorders and cue modalities (Brand et al., 2019; Antons et al, 2020; Limbrick-Oldfield et al., 2017; Filbey et al., 2009), identifying a distributed network involving the striatum, midbrain, anterior insula, and cingulate cortex amongst others as important for encoding cue-induced craving. In brief, cue-reactivity remains an important framework for explaining and understanding many empirical findings about craving.

The limitations of a purely cue-reactivity view of addiction, however, have become increasingly evident in recent years. Empirically, there are various findings related to craving that cannot be explained by cue-reactivity alone. For example, forced abstinence in addicted rodents (Grimm et al., 2001; Li et al., 2015) and humans (Parvaz et al., 2016; Bedi et al., 2011) paradoxically leads to increased craving and drug-seeking behaviors, in the absence of drug-associated cues, opposite to what a cue-reactivity model would predict. Other recent work also suggests that prior expectations about rewards to be received play an important role in evoking craving (Gu et al, 2016), an observation not well-explained by the cue-reactivity model. Collectively, these previous work suggests that the relationship between craving and choice behaviors – especially through a reinforcement learning lens – might not be as clear as previously thought.

A second major question is the generalizability of mechanisms related to craving and addictive behaviors. Although typically studied in the context of substance use disorders (SUDs), both craving and addictive choices have been increasingly recognized and examined in other disorders with addictive features, such as binge eating and behavioral addictions (e.g., gambling, social media use). These conditions and SUDs share features like persistent craving, unsuccessful attempts to quit, withdrawal effects, and heightened impulsivity, and may be explained by the same mechanistic framework (Kulkarni et al., under review). Specifically related to craving, recent work also demonstrates that acute social isolation can induce craving and related neural changes in the midbrain in a way that is similar to food craving (Tomova et al, 2020). Even within drug craving, very few studies have directly compared craving mechanisms across substances. As such, a domain-general computational account for craving remains unknown, hindering our understanding of the interaction between subjective states and explicit choice behaviors at a mechanistic level across addictive domains.

Recently, the nascent field of computational psychiatry has provided compelling theoretical frameworks for modeling both choice behaviors and subjective states, including craving (Kao et al, 2023; Redish, 2004; Sprevak & Columbo, 2019; Gu et al., 2019). Indeed, with the high degree of complexity, heterogeneity, and multidimensionality of symptomatology in psychiatric disorders, a deeper understanding of these latent mental states may be key to breakthroughs in targets for clinical interventions. Theoretical and empirical models abound for algorithmic explanations of how subjective states interact with valuation (Eldar et al., 2016; Zorowitz et al., 2020; Emanuel & Eldar, 2023; Vinckier et al, 2018; Rutledge et al., 2014; Sharot et al., 2009). In particular, expected value (EV) and prediction error (PE) have been frequently identified as important drivers of changes in subjective states. For example, Rutledge et al., 2014 demonstrated that EVs and PEs predicted mood more accurately than received rewards alone, suggesting that oscillation of internal states are dependent not on outcomes alone,

but also on expectations (and violations of expectations) about these outcomes. Conversely, subjective states may also modulate valuation, e.g., Eldar et al., 2016 outlined a theoretical model in which mood may increase 'momentum' of value updating, providing a plausible mechanism for how mood drives dynamic changes in learning. Momentary craving is likely entangled with addictive decision-making in similar ways, yet only a handful of empirical studies have examined this important relationship (Gu et al., 2015; Gu et al., 2016), and a computational mechanism linking these two constructs remains noticeably missing. Additionally, no studies have empirically examined the interaction between craving and learning processes specifically, despite well-established theoretical accounts of addiction that connect the two (Redish 2004; Zhang et al., 2009; Kalhan et al., 2023), instead focusing on trial-to-trial individuated decision-making (Rutledge et al., 2014), which likely involves different cognitive demands. To our knowledge, this is the first empirical study directly examining the interaction between craving and learning processes in addictive disorders.

Here, we hypothesized that momentary craving and decision-making will shape each other in a bidirectional fashion. To test this hypothesis, we developed a task in which human participants made choices to obtain either monetary or addictive (i.e., alcohol, cannabis, food) outcomes and self-reported their craving during decision-making. We examined three groups of participants (total n=202; see **Supp. Table 1** for participant characteristics): heavy alcohol drinkers (n=68), cannabis users (n=65), binge eaters (n=69). The task consisted of a modified two-armed bandit (**Fig. 1**), where participants selected one of two machines (80% reward rate), and saw the outcome of either a coin (in the money condition) or their pre-selected addictive cue (in the addictive cue condition). Momentary craving and mood were both sampled during the task (33% and 20% of the trials, respectively), and a joint modeling approach was used to simultaneously fit both choice and craving data per session. We found that, across all groups, increased craving dynamically modulated learning rate to drive faster learning from addictive cues, but not money cues. Additionally, computationally defined learning signals (i.e. expected value and prediction error) contributed to flucturations in perceived craving above and beyond simple cue-elicited effects. Finally, computational parameters derived from these models provided greater power than model-agnostic metrics for predicting clinical severity scores. Together, these results validate a generalizable computational mechanism linking momentary craving with value processing and decision-making.

RESULTS

Participants successfully learned to make optimal decisions

In order to ensure that computational models could be successfully fit, it was first necessary to ensure that participants could effectively perform the task with a high degree of success. Choice optimality was defined as the percentage of choosing the correct option (i.e. choosing the machine with the higher reward rate of 80%). Overall, participant choices were highly similar to the true reversal structure of the task (**Supp. Fig. 1a**), and choice optimality was significantly higher than chance (50%) (**Supp. Fig. 1b**; all groups: $P<0.001$), confirming that participants successfully learned to exploit the machines for maximum reward. Furthermore, choice optimality was not significantly different across participant groups ($F=1.1$, $P=0.33$), across task conditions ($F=0.47$, $P=0.49$), or their interaction ($F=0.26$, $P=0.77$), ensuring that findings related to craving-choice relationship would not be attributable to distinct task performance per se.

Momentary craving dynamically changed during the task and did not correlate with mood

Next, we sought to verify that that participants indeed experienced dynamic changes in their craving during the task. To do so, we calculated the mean and variability in self-reported cravings during the task across groups and conditions. We found craving was greater in the addictive cue condition compared to the money condition (Repeated measures ANOVA; $F=26.8$, $P<0.001$); there was no interaction effect between group and condition in terms of mean craving ($F=0.17$, $P=0.84$). We also observed a large amount of variability in craving ratings within-subjects (**Supp. Fig. 2b**), validating engagement during

the task and dynamic changes in perceived craving in response to outcomes. Craving variances were greater than zero for all groups and all conditions (all groups and conditions; $t>7$, $P<0.001$). Variances did not differ by group ($F=1.58$, $P=0.20$), condition ($F<0.01$., $P=0.95$), or their interaction ($F=0.29$, $P=0.74$).

It is possible that participants' craving ratings might be inherantly correlated with mood, as negative affect has been found to be associated with drug and food craving (Shen et al., 2012; Christensen & Pettijohn, 2001). To assess this, participant mood ratings were sampled ("what is your mood right now?") for 20% of the trials. We found that, across groups, within-participant craving ratings were not significantly correlated (**Supp. Fig. 3**; mean correlations; alcohol: $r=-0.01$, $P=0.94$; cannabis:$r=-0.16$, $P=0.35$; binge:$r=0.06$, $P=0.69$; social media: $r=0.22$, $P=0.18$), though some outliers showed strong positive and negative associations. This indicated that in most group-condition pairings, craving ratings contained distinct information from mood ratings. Binge-eaters in the addictive cue conditions were a notable exception, displaying strong positive correlations across participants (mean correlation=0.337, $t=6.66$, $P<0.001$), a finding that merits further investigation. We also directly correlated craving, mood, and decision behaviors with clinical severity scores directly (**Supp. Fig. 4**), discovering only an predictable positive relationship between severity and overall craving ratings.

A joint-modeling approach to identify bidirectional links between craving and decision-making

Our next objective was to construct computational models that bidirectionally linked choice behavior and momentary craving, and to compare these models to gold-standard alternatives. We composed four candidate decision model classes ($tdrl$, $\alpha bias$, $\beta\ bias$, *heuristic;* see **Table 1a** for a list of models and explanations). The first three models were derived from classic temporal difference reinforcement learning (TDRL) models that assume agents learn from outcomes and track values and prediction errors and incorporated asymmetric learning from positive and negative prediction errors (Gershman et al., 2015). Two of these models ($\alpha\ bias, \beta\ bias$) included key computational mechanisms for incorporating craving into either value updating or decision policy. Additionally, we specified a null model representing decision-making based on a simple lose-switch rule (*heuristic*).

In parallel to these decision-making candidates, we constructed two orthogonal model classes to estimate momentary craving (**Table 1b**). We recently proposed a Bayesian framework defining craving as a posterior inference from a top-down prior expectation about physiological bodily states following an outcome, and a bottom-up signal representing the likelihood that desired target was received (Gu & Filbey, 2017). We tested this theoretical framework in the present study, albeit instantiated by a reinforcement learning system rather than a Bayesian system. The overarching theoretical claims tested here were largely the same; namely that craving arises as combination of bottom-up (cue-reactivity) as well as top-down (prior expectations) signals. In the cue-elicited craving (CEC) class, only outcomes elicited perceived craving, while in the expectation-influenced (EI) class, expected values and prediction errors also influenced perceived craving. Importantly, craving and decision models were jointly implemented (see **Methods** for details), allowing for interdependent modulation of craving and valuation. The six joint models used for estimation and model comparison are listed in **Table 1c**.

Simulated choices and cravings match true behavior with high fidelity

To assess the fidelity of fit by these computational models, we generated 2000 simulations of choice behavior and cravings per model class and calculated the degree of alignment between simulations and true behavior. Overall, all decision model classes generated simulated behavior matching true behavior significantly better than chance (**Supp. Fig. 5**; all models, all conditions: $t>10$, $P<0.001$). Heuristic models generally performed significant worse than all other model alternatives across participants (all models, all conditions: $F>6.5$, $P<0.001$) as expected. Each craving model class also generated simulated cravings that were significantly correlated with true cravings across all participants and conditions (**Supp. Fig. 6**) (*CEC* – mean (s.d.) $r=0.20$ (0.18), $t>100$, $P<0.001$; *EI* – mean (s.d.) $r=0.38$ (0.25), $t>11$, $P<0.001$).

EI correlations were significantly high than CEC (all groups – t>11, P<0.001), providing a first indication that expectations likely provide some explanatory power for perceived craving.

Momentary craving biases towards faster learning specifically in response to addictive outcomes
Of our candidate models linking craving and valuation, which best explains the choices made by participants? How do our models compare to established gold-standard alternatives? Following model comparison (**Fig. 3a**), we found that the best-performing model was contingent on the experimental condition. In the money condition, the $tdrl$ model, in which craving did not modulate learning, was the best performer. In the addictive cue condition, the $\alpha\ bias$ model, in which momentary craving modulated learning rate, performed best. This suggested that the computation that best explained decision-making was contingent on the "addictive" relevance of the decision framework and that cravings were more likely to directly modulate decisions that were directly related to the addictive target.

What can the computational parameters derived from our model tell us about the underlying decision-making processes during this task? Constraining our analysis to the best-performing $\alpha\ bias$ model in the addictive cue condition, we found moderate learning rates across groups (**Fig. 4a**; learning rate mean (s.d.); alcohol – 0.50 (0.14), cannabis – 0.48 (0.14), binge – 0.47 (0.13)), with no significant difference between learning from positive vs. negative prediction errors (alcohol – t=0.69, P=0.49; cannabis – t=-1.9, P=0.06; binge – t=-0.49, P=0.62). Interestingly, however, modulation factor ('mod') which controls the direction and magnitude of craving's influence on learning rate, was significantly positive in all groups (mean (s.d.), significance; alcohol – 1.06 (0.57), P<0.001; cannabis – 0.95 (0.29), P<0.001; binge – 0.91 (0.54), P<0.001), indicating that increased craving positively modulated learning rate, ostensibly leading to faster value updating during addictive decision-making. Together, these findings suggest an important role for craving in dynamic modulation of learning. First, cravings influence learning ($\alpha\ bias$), rather than influencing decision policy ($\beta\ bias$), or being uninvolved in decision-making entirely ($tdrl$). Secondly, as craving increases, it biases toward *faster learning* from addictive outcomes. These effects provide a mechanistic explanation for why drug-related decision-making goes awry in an accelerated fashion towards the alter stages of addiction.

Trial-wise prediction errors significantly modulate perceived craving
Though our results corroborate a directional effect of craving on learning, it remains unclear whether the reverse interaction is true, i.e., do prior expectations influence perceived craving? Systematic craving model comparison revealed that expectation-influenced craving models indeed outperformed cue-elicited craving models as hypothesized (**Fig. 3b**). Importantly, the dominance of value-influenced models generalizes across conditions and groups, suggesting that incorporation of prior expectations into craving computation employs a shared domain-general mechanism. Our findings suggest that prior expectations about anticipated rewards, derived from values tracked during decision-making, play an essential role in evocation of craving, above and beyond simple cue-induced effects. These results provide the first empirical evidence for our theoretical framework which proposes that craving arises as a posterior inference from prior expectations and received outcomes.

As with the decision-making analysis, we derived distributions of computational parameters from the best performing model to interpret the processes underlying elicitation of craving during the task. The weight corresponding to cue-reactivity is significantly positive for alcohol and cannabis users (**Fig. 4b**; mean (s.d), significance; alcohol – 0.11 (0.57), P=0.03; cannabis – 0.16 (0.59), P=0.003), recapitulating the well-established cue-induced craving effect in these groups. More importantly, the weight associated with *prediction error* is also significantly positive across all groups (**Fig. 4b**; mean (s.d); alcohol – 0.08 (0.36), P=0.014; cannabis – 0.11 (0.29), P<0.001); binge – 0.15 (0.29), P<0.001)). Notably, cue reactivity weight is not significantly positive in binge eaters, suggesting that prediction errors may play a larger role than cues in eliciting craving in this group.

In sum, the models constructed here provide a means for disentangling two important components of perceived craving: cue-induced effects and expectation-induced effects. Our findings establish that while cue-induced effects may account for a modest fraction of craving, trial-to-trial prediction errors are also significant drivers of perceived craving, with increased prediction errors inducing increased momentary craving.

Model-derived parameters accounted for addiction severity better than model-agnostic measures
Though our models successfully disambiguate mental computations underlying craving, are they useful in predicting clinical severity? We tested a hypothesis that parameters derived from our models were predictively valuable above and beyond simple model-agnostic task metrics. A group-specific severity measure (ASSIST for alcohol, CAST for cannabis, EDE-Q for binge) was selected as the target, and regression analysis was restricted to the (best-performing) $\alpha\ bias$ model in addictive cue condition. Three classes of regression analyses were constructed: (1) Computational model-derived regression (Comp-R), in which only computational parameters were used to predict severity, (2) Model-agnostic regression (Agnostic-R), in which only task performance summary metrics (mean and s.d. of craving and mood) were used to predict severity, and (3) Joint regression (Joint-R), in which both sets of predictors were included. All regression analyses also included demographic controls for age, sex, and education level.

For each group, Joint-R, Comp-R, and Agnostic-R were compared and ranked by AIC and ELPD-WAIC scores (**Fig. 5**; **Supp. Table 2**). Joint-R categorically performed best across all groups (AIC (ELPD-WAIC); alcohol – 160.4 (-81.4), cannabis – 150.7 (-77.7), binge – 184.1 (-93.1)), and significantly improved prediction of severity compared to Agnostic-R (**Fig. 5a**; Joint-R vs. Agnostic-R improvement; alcohol – t=2.8 P=0.006; cannabis – t=3.3, P=0.001; binge – t=3.6, P<0.001). Joint-R also consistently outperformed Comp-R alone (**Supp. Table 2**) indicating model-agnostic and computational-model-dependent predictors provide orthogonal evidence that is more informative than either alone.

We also visualized the significantly predictive variables from the best-performing Joint-R model, seeking to interpret their relationship with clinical severity (**Supp. Table 3**; **Supp. Fig. 7**). Most interestingly, the 'mod' parameter is significantly predictive of clinical severity in alcohol users and cannabis users (β coefficient (significance); alcohol – 0.34 (P=0.013), cannabis – 0.32 (P=0.023)), but not in binge-eaters (0.01 (P=0.97)). Learning rate is also negatively associated in the cannabis group and binge-eating group (cannabis – -0.27 (P=0.02), binge – -0.20 (P=0.042)), but not in the alcohol group (0.12 (P=0.47)). Finally, inverse temperature is negatively associated with severity in alcohol users and binge-eaters (alcohol – -0.35 (P=0.003), binge – -0.27 (P=0.028)), but not cannabis users (-0.02 (P=0.87)). These findings may reflect the inherent differences between clinical groups, but further work will likely be necessary to elucidate the rationale behind heterogeneity in associations between parameters and clinical severity. In sum, a comparative regression analysis revealed that computational parameters indeed provide predictive power for clinical severity above and beyond simple model-agnostic measures.

DISCUSSION

In the present study, we identified a highly generalizable computational mechanism subserving the dynamic relationship between momentary craving and decision-making in a cross-sectional sample of participants with different primary addictive targets. Our findings provide strong support for the notion that decision-making and craving are computationally intertwined processes across addictive domains. Additionally, we challenged the notion that craving arises as a purely cue-elicited phenomenon, building on, and empirically testing, recent computational theories that both prior beliefs and outcomes play a role in perceived craving. Substantiating our hypothesis, we found that momentary craving dynamically influenced learning in a sequential decision-making task, and that prediction error, in addition to the traditional cue-induced effects drove perceived craving. At a broad level, these findings support the

expectation-influenced craving framework as a more plausible alternative to cue-induced craving, and provide a generalizable algorithm linking craving with decision-making.

Specifically, we demonstrate that increased craving leads to *faster learning* from outcomes, meaning that addiction-associated outcomes that follow a state of high craving may be more strongly valued by the brain. As a result, future decisions may be biased towards these outcomes, potentially leading to a self-reinforcing cycle of addictive behavior. This explanation corroborates similar findings in the field that addiction-associated rewards are distorted in value (Konova et al., 2018; Biernacki et al., 2021; Smith et al., 2021), and that drug-seeking behaviors paradoxically continue to persist, at least in the goal-oriented early phase of addiction, even as outcomes are devalued (Everitt & Robbins, 2005; Luscher et al., 2020). Additionally, we demonstrate that while cue-induced effects do exist, increased *prediction errors* generated during the task also significantly drive stronger cravings, adding a novel dimension to the traditional cue-reactivity model of framework. This result conforms with other findings that prediction errors are likely an important component of perceived craving in addiction at behavioral (Paulus et al., 2009; Verdejo-Garcia et al., 2012), computational (Naqvi et al., 2014; Stewart 2019), and neural (Wang et al., 2019; Gu et al., 2015) levels.

In our predictive modeling approach, we found that computational parameters reflect latent information that is predictive of clinical severity above and beyond model-agnostic metrics or demographics. Interestingly, inclusion of both model-agnostic and computational parameters yielded the best predictive power in substance users specifically, suggesting these components explain unique variance in SUD. To our knowledge, this provides the first empirical evidence that computationally driven parameters that are core to addiction can provide better predictions of addiction severity compared to model-agnostic measures. These findings provide compelling evidence for the value of computational modeling in uncovering latent information that may be highly valuable translationally.

Our finding of an algorithmic link between craving and decision-making has direct clinical implications: specifically, by reducing or managing cravings, individuals may be able to break the vicious cycle of addiction-associated decision-making that leads to poor clinical outcomes. Though interventions regarding craving have been proposed, tested, and examined widely (Li et al., 2018; Hone-Blanchet et al., 2014; Anton, 1999), our models provide the first comprehensive mechanisms that may explain their success, Further exploration will be necessary to envision a new line of therapeutic interventions focusing on the interplay between craving and decision-making that target both of these factors simultaneously.

This research is limited in that we do not have direct evidence supporting the neurobiological substrates that facilitate this mental computations proposed in this model. Though numerous efforts have identified the effect of decision-making on subjective states, such as happiness (Rutledge et al., 2014) or craving (Wang et al., 2019; Gu et al., 2015), there are, as of yet, no neuroimaging studies investigating a bidirectional interaction between the two. A next logical step would be to collect neural data (e.g., fMRI) in conjunction with our experimental design, and utilize machine learning or statistical approaches to find the neural patterns encoding the computational mechanisms outlined here.

CONCLUSION

Overall, this study highlights a reciprocal, dynamic link between decision-making and craving in addictive disorders. Our results demonstrate that momentary craving leads to faster learning in a decision-making task, and support the idea of expectation-influenced craving as a more plausible alternative to cue-induced craving, with prediction error being a key driver of perceived craving.

METHODS

Task design - instructions

We utilized a modified two-armed bandit task that incorporated reversals to encourage continuous learning during the experiment. At the start of the task, participants were presented with the instructions about how to play the game. Participants were presented with two slot machines and were encouraged to do their best to maximize their rewards from the machines, with the incentive of being rewarded with a greater bonus payment at the end of the game based on their final score.

In one of the conditions, participants received a monetary reward, presented as an image of a coin. In the other condition, they received a reward corresponding to an addictive cue, selected based on the addictive group to which they belonged (e.g., in the alcohol-using group, participants were rewarded with an image of alcohol). Additionally, during the start of the experiment, participants were able to select one of three possible addictive cues that were most tempting to them. For example, participants in the alcohol-using group were able to select from either a beer, wine, or liquor reward.

Participants were then also informed that they would intermittently be asked to assess and report their craving and mood. They were given specific definitions for craving ("An intense, conscious desire or wanting for something") and mood ("A non-specific, persistent general feeling about your current mental state, distinct from emotions, which are shorter-lived and specific to a particular thing"), in order to provide them with a deeper understanding of the concepts being measured, and to provide a concrete basis upon which their self-reports could be compared. Baseline craving and mood ratings were then assessed prior to presentation of the slot machine task.

Finally, participants were informed that one of the slot machines would be more rewarding than the other at all times, but the more rewarding slot machine might change over the course of the experiment. The true experimental structure, including the probabilities of reward, the reversal timings, and number of trials, was not revealed to the participants beyond the information listed above.

Task design - experimental structure

There were 60 trials per condition, and 5 practice trials at the start of the experiment for a total of 125 trials. The 5 practice trials were discarded during the computational modeling. Presentation of conditions was randomized over the cohort, such that half of the participants received the money condition first and the other half received the addictive cue condition first. The more rewarding slot machine presented a reward 80% of the time, while the less rewarding machine presented a reward 20% of the time. The more rewarding machine switched four times over the course of the experiment ("reversals"). First selection of the best machine (either right or left), was pseudorandomized and reversal timings were also pseudorandomized as either 12-12-11 or 13-12-10. Craving ratings were assessed approximately every three trials for a total of 20 craving ratings per condition. Mood ratings were assessed approximately ever 5 trials for a total of 12 mood ratings. Ratings were collected on a linear scale from 0-50, where participants used the arrow keys to move a cursor from left to right to select their answers, where the left end of the bar was labeled as 'Low' and the right end was labeled as 'High'.

Data collection

We first performed a high-throughput screening for 1000 participants on Prolific, an online data collection platform for recruiting participants across the world. We restricted our recruitment to participants with the following characteristics: USA residents, fluent English speakers, high approval rating on Prolific, no previous completion of any of our group's experiments. Furthermore, we recruited a

balanced sample of male and female participants. Participants completed three screening surveys implemented on Redcap. The first was the World Health Organization's Alcohol, Smoking and Substance Involvement Screening Test (ASSIST), which was developed as a screening tool to help primary health professionals detect and manage substance use and related problems. ASSIST allows for screening of use of substances such as alcohol, cannabis, tobacco, stimulants, inhalants, and others. Second, we used the Eating Disorder Examination Questionnaire (EDE-Q), a well-validated instrument designed to capture self-reported eating disorder behaviors and attitudes. We primarily utilized this screening tool as a way to identify participants with binge episodes in the past month. Binge episodes were defined as meeting two criteria: (1) consuming an unusually large amount of food that most people would feel is larger than one meal, and (2) feeling a sense of having lost control while eating.

From the screening surveys, we recruited cohorts for alcohol use, cannabis use, and binge-eating. Potential participants were identified with the following criteria: (1) The alcohol use cohort reported moderate to heavy use of alcohol weekly as identified by ASSIST, with no comorbid usage of any other substances. (2) The cannabis use cohort reported moderate to heavy use of cannabis weekly, with low to no use of alcohol, and no other substance usage. (3) The binge-eating cohort reported at least one binge episode within the past 4 four weeks.

From the full eligible cohort for each group, we collected a total of 40 participants' data for the slot machine game. Participants with substance use were additionally asked to complete further group-specific questionnaires. Alcohol users completed an Alcohol Dependence Scale, and an Alcohol Use Disorder Identification Test. Cannabis users completed a Cannabis Severity of Dependence Scale, and a Cannabis Abuse Screening Test.

We followed the exact protocol listed above for a replication sample, starting with 1000 screened participants, filtering of participants into potential cohorts for data collection, collection of 40 samples from each group of interest, and collection of auxiliary clinical questionnaires for the substance use groups. The data from the two samples was combined into our final analysis sample reported in the manuscript.

Quality control

Before analysis, participant data from each group underwent a series of quality control checks. First, we included one attention check during the experiment, with instructions about halfway through the task to select the highest option if the participant was paying attention. Participants that did not pass this attention check were excluded. Second, since the optimal machines and reversal timings are known, it is possible to construct the vector of optimal choices during the experiment. A participant's raw choice behaviors were compared to the optimal choice vectors for each condition, and only participants with greater than 40% optimality were selected, ensuring that they learned to play the game well, and were not simply randomly responding in the task, with a relatively low bar of exclusion reducing the chance that randomness of reward structure did not unnecessarily exclude well-performing participants. Finally, we z-scored the reported craving ratings during the task and excluded the participant if the standard deviation of the ratings did not exceed 1, suggesting that there was very low variability in the rating scores. Since the computational modeling for the behavioral data simultaneously fits the choice behavior and predicts craving, we could not properly fit the data for participants that essentially reported no variability in their craving. Moreover, we found that most participants with extremely low craving variability seemed to exclusively report a craving of 25/50, which was the default value in the rating scale, suggesting that they were not properly responding to the prompts. Following quality control, the participant counts were: alcohol (n=68), cannabis (n=65), binge (n=69).

Model-agnostic analysis

1. **Calculation of mean/SD of craving and mood:** We calculated the means and variances of reported cravings for each individual across groups and conditions. These were used to assess whether simple summary statistics of overall trends of craving were associated with clinical measures. The same process was repeated for reported moods. Individual distributions of cravings and moods were aggregated and reported across groups and conditions.
2. **Redcap scoring:** We scored the Redcap surveys for each group according to a questionnaire-specific scoring. For the alcohol group, we utilized 3 surveys: ASSIST, AUDIT, and ADS. ASSIST scores ranged from 0-40 (<3 – low severity, 4-26 – medium severity, >26 high severity). AUDIT scores ranged from 0-34 (<7 – low severity, 8-14 – medium severity, >14 high severity). ADS scores ranged from 0-54 (0 – No risk, 1-13 – low risk, 14-21 – moderate risk, 22-30 – substantial risk, >30 – severe risk). For the cannabis group, utilized 3 surveys: ASSIST, CAST, and SDS. ASSIST-Cannabis scores ranged from 0-36. CAST scores ranged from -6-20 (<3 – low risk, 3-6 – moderate risk, >6 high risk). SDS scores ranged from -5-13 (<=3 – low risk, >4 – high risk). For the binge-eating group, we utilized a brief Loss of Control over Eating Scale (LOCES-B) and a modified EDE-Q. LOCES-B scores ranged from 7-34, and EDE-Q scores ranged from 2-131. Finally, social media use was surveyed using a custom survey, which assess the number of services used, frequency of daily check-ins, size of the social networks, amount of daily usage, and impact on daily life. Scores ranged from 0-16. The distributions of all these surveys were visualized across groups. Additionally, we calculated the difference between clinical severity scores before the task (during screening surveys) and after the task to assess the stability of these clinical measures.
3. **Choice optimality:** We qualitatively assessed the performance of participants to ensure that they were able to learn the structure of the task well. To do this, we compared the choice of machine for all participants in a group to the optimal choice that could be made at that time, where the optimal choice was defined as the machine with higher probability of reward. Participants with lower than 50% optimality were excluded from further analysis because they were unable to learn the task structure. Finally, we plotted the distributions of optimality of remaining participants to demonstrate that participant optimality was well above chance.
4. **Qualitative performance checks:** We averaged the choice of machine across participants by group, controlled for randomization of the order of the best machine presented first (either left or right) to qualitatively assess group-level tracking of reversals across the task. We then visualized the overall distributions of craving and mood during the task within individuals and across groups to ensure that there was a realistic distribution of cravings and moods reported during the task.
5. **Low-level sanity check correlations:** We correlated group-specific clinical scores (ASSIST for alcohol and cannabis, EDEQ for binge-eating, and usage survey for social media) with percent optimality, total score, and baseline and mean craving/mood ratings. We also performed a Spearman correlation between reported mood and craving ratings within participant to check the covariance of the two.

Modeling analysis

There were five classes of decision-making.

1. Standard temporal difference learning ($tdrl$) – Values were computed using a standard temporal difference learning rule, given below.

$$V_t = V_{t-1} + \alpha(r_t - V_{t-1})$$

2. Additive learning bias models ($\alpha\ bias$) – Identical to the specification above, but learning rate at every trial was dynamically modulated by an addictive bias dependent on momentary craving and a modulation factor ('mod').

$$V_t = V_{t-1} + \alpha_t(r_t - V_{t-1})$$
$$\alpha_t = \alpha_{static} + mod \times craving_t$$

3. Multiplicative learning bias models – Identical to the specification, but the bias was multiplicative rather than addictive.

$$V_t = V_{t-1} + \alpha_t(r_t - V_{t-1})$$
$$\alpha_t = \alpha_{static} \times mod * craving_t$$

4. Reward modulation models – Identical to the $\alpha\ bias$, but the additive bias term was applied to the reward magnitude rather than learning rate. This class of models in most similar to Biernacki et al., 2022.

$$V_t = V_{t-1} + \alpha_t(r_t - V_{t-1})$$
$$\alpha_t = \alpha_{static} \times mod * craving_t$$

5. Temperature bias ($\beta\ bias$) – craving additively biased the inverse temperature parameter. The value update equations were identical to the $tdrl$ model.

$$\beta_t = \beta_{static} + mod * craving_t$$
$$decision_t \sim policy(\beta_t, V_t)$$

We found that Models 2, 3, and 4 were highly similar in terms of mechanism and model performance, and thus simplified our results down to three candidates ($tdrl$, $\alpha\ bias$, $\beta\ bias$), which encompassed three distinct and plausible model alternatives.

Two learning rate flavors were implemented for each decision-making class: (1) a standard TDRL with single learning rate, and (2) dual learning rate based on valence of the prediction error (Niv et al., 2012; Gershman 2015). Generally, the dual learning rate was found to be more successful than the single learning rate model, and the final analysis used only the dual learning rate model.

There were two variants of craving models.

1. Cue-elicited craving (CEC) - In this class, task outcomes were the only predictors of perceived craving, with a cue-reactivity weight modulating the degree of cue-induced response.

$$craving_t = w_{cue} * outcome_t$$

2. Expectation-influenced craving (EI) – In this class, task outcomes, expected values (EV), and prediction errors (PE) all influenced momentary craving.

$$craving_t = w_{cue} * outcome_t + w_{EV} * EV_t + w_{PE} * PE_t$$

Finally, there were seven variants governing the number of expected values (EV) and prediction errors (PE) from previous timepoints that influenced craving computation.

1. 0-step – Only the current trial EV and RPE were used in the computation of craving.
2. 1-step mean – the means of the current and previous trial EV and RPE, respectively were used
3. 1-step decay – the weight of the influence of EV and RPE exponentially decayed for the previous trial

4. 1-step separate – separate weights were estimated for the current and previous trial's EV and RPE
5. 2-step mean – the mean of the current and the previous two trials was used
6. 2-step decay – the weight of the influence of EV and RPE exponentially decayed for the previous two trials
7. 2-step separate – separate weights were estimated for the current and previous two trials' EV and RPE

Based on model performance, theoretical considerations, and previous empirical findings, we decided to select (6) 2-step decay as a sole representative from these variants, as it had the best balance between degrees of freedom, minimization of overfitting, out-of-sample performance, and ecological plausibility. All results in the main figures are derived from this variant.

For final reported model comparison, we had five alternatives, generated as unique combinations of decision-making and craving models, since these models were fit to behavioral responses jointly.

1. $tdrl + CEC$ – Standard TDRL decision-making with cue-elicited craving only
2. $\alpha\ bias + EI$ – Additive bias on learning rate from craving with expectation-influenced craving
3. $\beta\ bias + EI$ – Additive bias on inverse temperature from craving with expectation-influenced craving
4. $tdrl + EI$ – Standard TDRL decision-making with expectation-influenced craving
5. $heuristic + CEC$ – a simple heuristic for decision-making with cue-elicited craving only

Model implementation

Models were implemented in pymc, a Python library for Bayesian parameter estimation. Reported craving ratings and task choice behavior was fit simultaneously. Each model is instantiated with reasonable priors for its parameters as specified in the table below.

Model	α	β	$craving_0$	w_{cue}	w_{EV}	w_{PE}	mod
$heuristic + CEC$	-	-	$N(0,1)$	-	-	-	-
$tdrl + CEC$	$N(0,1)$	$HalfN(10)$	$N(0,1)$	-	-	-	-
$tdrl + EI$	$N(0,1)$	$HalfN(10)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	-
$\alpha\ bias + EI$	$N(0,1)$	$HalfN(10)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$
$\beta\ bias + EI$	$N(0,1)$	$HalfN(10)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$

Learning rate (+/-), and mod parameters were normalized to the (0, 1) interval using the inverse logit transformation.

For each sampled parameter combination during the Bayesian parameter estimation, the following procedure was used.

1. Values for each machine were initialized at 0.5 and craving was initialized at 0.
2. The algorithm looped over trials, adding to a grand sum log likelihood based on the log probability of true actions and reported craving based on estimated values and predicted craving.
3. Per trial, values were computed using a model-specific valuation rule, and likelihood of actions were assessed using probabilities derived from a softmax function.
4. Per trial, craving was computed using a model-specific craving computation rule, and the resultant vector of predictions were normalized to probabilities using the inverse logit

transformation. Likelihoods for these vectors were computed as the log probability mass of a binomial distribution (N=50, p=trial-wise probability).

5. The grand total log likelihood of actions and craving was defined as the joint LL.
6. A modified Metropolis-Hastings sampling method was used to traverse the parameter space to generate the posterior.
7. The model obtained a total of 4000 samples from the posterior.

Model fit metrics

Posterior samples of parameter sets were used to simulate choices and cravings for each participant following parameter fitting. The model ran a simulation for each parameter set for a total of 4000 datasets of simulated choices and cravings. The following checks were performed on the simulated data.

1. Means of these simulations, both choices and cravings, were visualized against true actions and cravings to qualitatively assess congruence.
2. Each choice simulation was given an accuracy score representing the number of simulated actions that matched true actions, where 50% was chance accuracy.
3. Each craving simulation was correlated with the true craving, resulting in a correlation score as well as the corresponding significance score (p-value).

Model comparison

For each model, information criteria scores were calculated to compare model performance. For both choice behavior and craving, we compared model performance by each model's expected log pointwise predictive density (ELPD), which was estimated using the widely applicable information criteria (WAIC). WAIC is a widely used, and well-established method of estimating level of model fit, penalizing for overfitting and number of parameters (Watanabe, 2010). ELPD-WAIC is calculated with the following equations:

$$\widehat{\mathrm{elpd}}_{\mathrm{waic}} = \widehat{\mathrm{lpd}} - \hat{\mathrm{p}}_{\mathrm{waic}}$$

$$\widehat{\mathrm{lpd}} = \sum_{i=1}^{n} log\left(\frac{1}{S}\sum_{s=1}^{S} p(y_i|\theta^s)\right)$$

$$p_{waic} = \sum V_{s=1}^{S}\left(log * p(y_i|\theta)\right)$$

$\widehat{\mathrm{lpd}}$ represents the computed log pointwise predictive density, calculated across all posterior sample across all participants. This score is penalized by a term ($\hat{\mathrm{p}}_{\mathrm{waic}}$), which represents an effective number of parameters determined by the posterior variance of the log predictive density for each data point. These model comparison metrics were automatically calculated by the pymc library. We additionally calculated deviance information criteria (DIC) for craving models to validate the WAIC model comparison. DIC is calculated with the following equations:

$$DIC = -2 * logp(\, y \mid \theta \,) + 2 * p_{DIC}$$

$$p_{DIC} = 2 * \left(logp(\, y \mid \theta \,) - \mathbb{E}_{post}\, logp(\, y \mid \theta \,)\right)$$

WAIC and DIC are often used as an information criterion for Bayesian models because unlike other common IC metrics, such as Bayesian information criterion (BIC), it penalizes for *effective* number of parameters, rather than raw number of parameters. For models with a large number of parameters, like the

ones we are testing, these measures are better able to distinguish high-fidelity model fit. Additionally, there is evidence than despite being very similar, WAIC provides more stable estimates of model fit and generalizability (Vehtari et al., 2016), and it is also calculated automatically by the pymc library, so all model results shown are determined by WAIC.

Parameter estimate distributions

Across groups, and within each condition (money or addictive cue) and model type (decision or craving), we identified the best performing model by the WAIC score. For this model, we plotted the distributions of parameter estimates across participants for all parameters utilized in the model. Though all model parameters were estimated jointly, we visualized decision-making parameters and craving parameters in separate sub-figures for easier summarization and interpretation. For each datapoint, we plotted the mean of all parameter samples obtained from Bayesian parameter estimation, although results were confirmed using the raw samples as well. For each parameter, we tested the directionality of the estimated effect by calculating statistical significance from zero. Finally, we performed significance testing between alpha parameter estimates obtained for positive and negative prediction errors. All significance testing was performed with parametric t-test (either independent or relative, depending on suitability of the samples) and confirmed with non-parametric Mann-Whitney U-tests.

Clinical score prediction

We used a multiple linear regression model to test the hypothesis that joint parameter estimates from best performing models can successfully predict clinical severity scores. For clinical severity scores, we decided to use a group-specific Redcap survey that had high variability across participants. For alcohol users, we used ASSIST, for cannabis users, CAST, and for binge eaters, EDE-Q. The best performing model identified by the model comparison procedure, restricted to the more salient addictive cue condition, was tested in three classes of regression analysis. In computational regression (Comp-R), model-specific parameter estimates were extracted from each participant's Bayesian traces. In model-agnostic regression (Agnostic-R), simple means and variances of craving and mood ratings were used to predict clinical severity. Finally, in joint regression (Joint-R), these approaches were combined, using both computational parameter estimates and craving/mood scores together to predict clinical severity. Participant demographic variables including age, sex, and level of education were included in all regression model classes. Models were implemented using the ordinary least squares (OLS) linear model in the 'statsmodels' Python library, as well as in 'bambi', a Bayesian linear modeling Python library. Models were ranked in performance using adjusted R-squared scores, AIC, and ELPD-WAIC scores. The best performing models were additionally refined with a stepwise elimination procedure by removing extraneous variables with significance values lower than a predetermined threshold ($P<0.05$) that were not contributing to prediction, and AIC scores were recalculated and compared to the original models. The final reported models represent the best-performing (by AIC and ELPD-WAIC) stepwise-regressed models. Variables significantly ($P<0.05$) contributing to prediction in the final models were visualized and interpreted within and across groups.

FIGURE CAPTIONS

Fig. 1. Specification of experimental procedure. Participants played a modified two-armed bandit task. Each session started with instructions and a choice of three reward options, of which the participant selected the one most tempting to them. Following five practice trials, two blocks of sixty trials were presented, in which the participant selected one of two machines. In the money block, the reward presented was an image of a coin. In the addictive cue block, the reward presented was the option selected by the participant at the start. Craving and mood were intermittently assessed throughout the block (20 craving ratings, 12 mood ratings), scored from 1 to 50. The more rewarding machine in each trial rewarded with 80% probability while the less rewarding machine rewarded at 20% probability. The optimal machine switched a total of 4 times (4 reversals) over the course of each block. The order and timings of the reversals was pseudorandomized over the cohort.

Fig. 2. Overview of computational modeling approach. At any trial, the momentary learning rate is dynamically modulated by the perceived craving at that trial. This, in turn, modulates the trial's expected value update. Conversely, the trial-wise expected value and trial-wise prediction error contribute to an expectation-based effect, which combines with a cue-induced effect to determine momentary perceived craving. Our proposed model can be contrasted against the classic TDRL and cue-reactivity models, which specify the following, (1) perceived craving does not modulate learning rate, (2) there is no expectation-based effect from expected value and prediction error, and perceived craving is determined purely through a cue-induced effect.

Fig. 3. Model comparison for choice behavior and craving. Model classes were compared across group and condition. (a) In the money condition (top row), standard $tdrl$ models, without the influence of craving on value update, performed best across groups (ELPD-WAIC; alcohol – 1388; cannabis – 1592; binge – 1909). In the addictive cue condition, however, the $\alpha\ bias$ model, where craving affects learning rate, performed best (ELPD-WAIC; alcohol – 1565; cannabis – 1605; binge – 1947). The heuristic model performed the worst across conditions and groups (ELPD-WAIC (money, addictive cue); alcohol – (1943, 2068); cannabis – (2041, 2010); binge – (2263, 2332)). (b) For the craving models, note that there are only two classes, representing a craving function elicited by cues alone, or by a combination of expectation-induced and cue-induced effects. In all group-condition pairings, expectation-influenced models performed the best (ELPD-WAIC (money, addictive cue); alcohol – (4276, 3867); cannabis – (3673, 3802); binge – (4392, 4386))

Fig. 4. Parameter distributions for addictive cue condition. There is a large variation in parameter distributions across participants across groups for the best-performing computational models, suggesting that, though the overall computational mechanisms are shared, there is a significant amount of phenotypic heterogeneity in the way individuals utilize values and craving to make trial-to-trial decisions. (a) There is no significant difference between learning rates for positive and negative prediction error trials (t-test; all groups – $t<1.9$, $P>0.06$). 'Mod' parameter is significantly positive across all groups (t-test; alcohol – $t=16.3$, $P<0.001$; cannabis – $t=32.0$, $P<0.001$; binge – $t=25.1$, $P<0.001$). (b) Cue-induced effect is significantly positive in alcohol users and cannabis users (t-test; alcohol – $t=2.24$, $P=0.027$; cannabis – $t=3.01$, $P=0.003$), but not in binge eaters (binge – $t=1.74$, $P=0.084$). Prediction error induced effect is significantly positive across all groups (t-test; alcohol – $t=2.50$, $P=0.014$; cannabis – $t=4.17$, $P<0.001$; binge – $t=6.02$, $P<0.001$).

Fig. 5. Regression analyses for clinical severity scores. Regression analyses were conducted to determine the efficacy of computational parameters in predicting clinical severity above and beyond simple model-agnostic metrics. Computational regression (Comp-R; green) and joint regression (Joint-R; orange), were visually and quantitatively compared to model-agnostic regression (Agnostic-R). (a) In each group, Joint-R significantly was superior to Agnostic-R (alcohol – $t=2.8$ $P=0.006$; cannabis – $t=3.3$,

P=0.001; binge – *t*=3.6, *P*<0.001, and explained a meaningfully larger proportion of the variance (**Table 3-3**; R^2 (best-performing vs. alternative); alcohol (R^2: 0.34 vs. 0.13); cannabis (R^2: 0.43 vs. 0.17); binge (R^2: 0.24 vs. 0.03)). (b) Bayesian model comparison validated that Joint-R outperformed both Comp-R and Agnostic-R (ELPD-WAIC; alcohol – -81.2; cannabis – -77.7; binge – -93.1).

SUPPLEMENTARY FIGURE CAPTIONS

Supplementary Fig. 1. Choice optimality. Choice optimality was measured by calculating the proportion of participant choices that matched the optimal choice at each trial (i.e., the machine that rewarded 80% of the time at that trial). (a) Choice optimality was assessed visually by averaging participant choices across the experiment. Blue lines represent the left machine, orange lines represent the right machine, and vertical gray lines represent the timings of reversals in reward structure. In all groups, participants were able to successfully learn the reversal structure, reflected in the periodic switches of left/right machine choices over the experiment. (b) Choice optimality was calculated for each participant, and stratified by group and condition. In all groups, choice optimality was significantly higher than chance (50%) (alcohol – t=14.6, P<0.001; cannabis – t=14.0, P<0.001; binge – t=13.0, P<0.001), but choice optimality was not significantly different across groups per condition (money – F=0.89, P=0.45; addictive cue – F=0.28, P=0.84).

Supplementary Fig. 2. Craving rating means and variances across groups. (a) Average craving ratings were similar for alcohol and cannabis users (alcohol – 18.3; cannabis – 18.7), but significant higher in binge eaters (27.1; F=21.8; P<0.001). Average craving ratings were significantly higher in the addictive cue condition across groups (alcohol: t=3.14, P=0.002; cannabis: t=2.46, P=0.016; binge: t=3.68, P<0.001). (b) Variances in craving ratings were not significantly different across groups (F=1.58, P=0.20) or conditions (F<0.01, P=0.95).

Supplementary Fig. 3. Correlations between severity scores, task performance and craving/mood ratings. Craving ratings and mood ratings were correlated with participant for all shared rating trials (12 across the whole session). Only the binge-eating group in the addictive cue condition achieved significance (t=6.66, P<0.001), displaying a strong positive relationship between mood and craving. All other group-condition pairing displayed no consistent directional relationship between the two, although some participants were definitely positive and negative outliers.

Supplementary Fig. 4. Correlations between severity scores, task performance and craving/mood ratings. (a) Group-specific clinical scores (alcohol – ASSIST; cannabis – CAST; binge – EDEQ) were correlated with task performance metrics (score and choice optimality). ASSIST was negatively correlated with choice optimality in the money condition (r=-0.33, P=0.008), while CAST and EDEQ were negatively correlated with choice optimality in the addictive cue condition (CAST: r=-0.37, P=0.003; EDEQ: r=-0.27, P=0.023). (b) Group-specific clinical scores were correlated with mood and craving (baseline rating and mean rating). ASSIST and CAST were positively correlated with craving ratings (ASSIST-baseline: r=0.36, P=0.003; ASSIST-mean: r=0.28, P=0.024; CAST-baseline: r=0.51, P<0.001; CAST-mean: r=0.40, P=0.001). Binge-eating severity is not correlated with craving (EDEQ-baseline: r=0.14, P=0.24; EDEQ-mean: r=0.01, P=0.91). Clinical scores are not associated with mood ratings (all groups: r<0.13, P>0.27).

Supplementary Fig.5. Simulated choice accuracy and qualitative match. To validate the goodness-of-fit for decision-making model classes, I employed a simulation-based qualitative and quantitative analysis. After fitting each participant's choice behaviors to our generative decision-making models, I simulated 4000 datasets of choice behavior per participant. (a) These simulated choices were plotted against true behavior to visually assess the fidelity of the model is recapturing participant behavior. Red dots represent true choice behavior for a sample participant in the binge-eating group, and blue dots represent simulated behavior. There is a high degree of similarity between simulations and true behavior, suggesting models do well in explaining true behavior. (b) Simulated choice accuracy was defined as the proportion of simulated choices that matched true choices. Across groups and conditions, average simulated choice accuracy was significantly higher than chance (50%) (money condition; alcohol –

t=10.2, $P<0.001$; cannabis – t=12.8, $P<0.001$; binge – t=9.7, $P<0.001$; addictive cue condition; alcohol – t=10.6, $P<0.001$; cannabis – t=8.1, $P<0.001$; binge – t=8.7, $P<0.001$).

Supplementary Fig. 6. Simulated craving correlations and qualitative match. Similar to the decision-making validation, I employed a simulation-based qualitative and quantitative analysis to test the goodness-of-fit for craving model classes. After fitting each participant's craving with our generative craving models, I simulated 4000 datasets of cravings per participant. (a) These simulated craving were plotted against true ratings to visually assess the fidelity of the model is recapturing participant behavior. The black line represents the true craving ratings from one sample participant in the binge eating group, and colored lines (green, orange, and blue) represent simulated behavior from three sample simulations. There is a high degree of correspondence between simulations and true ratings. (b) Simulated craving correlation was defined as correlation between simulated cravings and true craving ratings. Across groups and conditions, average simulated craving correlation was significantly higher than chance (r=0) (all groups and conditions – t>11.0, $P<0.001$). Important, simulated craving correlations was also significantly higher in value-influenced models compared to cue-elicited models only (all groups and conditions – t>11.0, $P<0.001$), providing another indication that value-influenced models are superior to models reliant on cue-induced effects only.

REFERENCES

1. Anton, Raymond F. "What Is Craving?" *Alcohol Research & Health* 23, no. 3 (1999): 165–73.
2. Antons, Stephanie, Matthias Brand, and Marc N. Potenza. "Neurobiology of Cue-Reactivity, Craving, and Inhibitory Control in Non-Substance Addictive Behaviors." *Journal of the Neurological Sciences* 415 (August 15, 2020): 116952. https://doi.org/10.1016/j.jns.2020.116952.
3. Bedi, Gillinder, Kenzie L. Preston, David H. Epstein, Stephen J. Heishman, Gina F. Marrone, Yavin Shaham, and Harriet de Wit. "Incubation of Cue-Induced Cigarette Craving During Abstinence in Human Smokers." *Biological Psychiatry*, Reduced Behavioral Flexibility in Addiction, 69, no. 7 (April 1, 2011): 708–11. https://doi.org/10.1016/j.biopsych.2010.07.014.
4. Biernacki, Kathryn, Silvia Lopez-Guzman, John C. Messinger, Nidhi V. Banavar, John Rotrosen, Paul W. Glimcher, and Anna B. Konova. "A Neuroeconomic Signature of Opioid Craving: How Fluctuations in Craving Bias Drug-Related and Nondrug-Related Value." *Neuropsychopharmacology* 47, no. 8 (July 2022): 1440–48. https://doi.org/10.1038/s41386-021-01248-3.
5. Brand, Matthias, Hans-Jürgen Rumpf, Zsolt Demetrovics, Daniel L. King, Marc N. Potenza, and Elisa Wegmann. "Gaming Disorder Is a Disorder Due to Addictive Behaviors: Evidence from Behavioral and Neuroscientific Studies Addressing Cue Reactivity and Craving, Executive Functions, and Decision-Making." *Current Addiction Reports* 6, no. 3 (September 1, 2019): 296–302. https://doi.org/10.1007/s40429-019-00258-y.
6. Carter, Brian L., and Stephen T. Tiffany. "Meta-Analysis of Cue-Reactivity in Addiction Research." *Addiction* 94, no. 3 (1999): 327–40. https://doi.org/10.1046/j.1360-0443.1999.9433273.x.
7. Eldar, Eran, Robb B. Rutledge, Raymond J. Dolan, and Yael Niv. "Mood as Representation of Momentum." *Trends in Cognitive Sciences* 20, no. 1 (January 1, 2016): 15–24. https://doi.org/10.1016/j.tics.2015.07.010.
8. Emanuel, Aviv, and Eran Eldar. "Emotions as Computations." *Neuroscience & Biobehavioral Reviews* 144 (January 1, 2023): 104977. https://doi.org/10.1016/j.neubiorev.2022.104977.
9. Ersche, Karen D., Claire M. Gillan, P. Simon Jones, Guy B. Williams, Laetitia H. E. Ward, Maartje Luijten, Sanne de Wit, Barbara J. Sahakian, Edward T. Bullmore, and Trevor W. Robbins. "Carrots and Sticks Fail to Change Behavior in Cocaine Addiction." *Science (New York, N.Y.)* 352, no. 6292 (June 17, 2016): 1468–71. https://doi.org/10.1126/science.aaf3700.
10. Everitt, Barry J., and Trevor W. Robbins. "Neural Systems of Reinforcement for Drug Addiction: From Actions to Habits to Compulsion." *Nature Neuroscience* 8, no. 11 (November 2005): 1481–89. https://doi.org/10.1038/nn1579.
11. Filbey, Francesca M., Joseph P. Schacht, Ursula S. Myers, Robert S. Chavez, and Kent E. Hutchison. "Marijuana Craving in the Brain." *Proceedings of the National Academy of Sciences of the United States of America* 106, no. 31 (August 4, 2009): 13016–21. https://doi.org/10.1073/pnas.0903863106.
12. Gershman, Samuel J. "Empirical Priors for Reinforcement Learning Models." *Journal of Mathematical Psychology* 71 (April 1, 2016): 1–6. https://doi.org/10.1016/j.jmp.2016.01.006.
13. Goldstein, Rita Z., and Nora D. Volkow. "Dysfunction of the Prefrontal Cortex in Addiction: Neuroimaging Findings and Clinical Implications." *Nature Reviews Neuroscience* 12, no. 11 (November 2011): 652–69. https://doi.org/10.1038/nrn3119.
14. Grimm, Jeffrey W., Bruce T. Hope, Roy A. Wise, and Yavin Shaham. "Incubation of Cocaine Craving after Withdrawal." *Nature* 412, no. 6843 (July 2001): 141–42. https://doi.org/10.1038/35084134.

15. Groman, Stephanie M., Bart Massi, Samuel R. Mathias, Daeyeol Lee, and Jane R. Taylor. "Model-Free and Model-Based Influences in Addiction-Related Behaviors." *Biological Psychiatry* 85, no. 11 (June 1, 2019): 936–45. https://doi.org/10.1016/j.biopsych.2018.12.017.
16. Groman, Stephanie M., Summer L. Thompson, Daeyeol Lee, and Jane R. Taylor. "Reinforcement Learning Detuned in Addiction: Integrative and Translational Approaches." *Trends in Neurosciences* 45, no. 2 (February 1, 2022): 96–105. https://doi.org/10.1016/j.tins.2021.11.007.
17. Gu, Xiaosi. "Incubation of Craving: A Bayesian Account." *Neuropsychopharmacology* 43, no. 12 (November 2018): 2337–39. https://doi.org/10.1038/s41386-018-0108-7.
18. Gu, Xiaosi, and Francesca Filbey. "A Bayesian Observer Model of Drug Craving." *JAMA Psychiatry* 74, no. 4 (April 1, 2017): 419–20. https://doi.org/10.1001/jamapsychiatry.2016.3823.
19. Gu, Xiaosi, Terry Lohrenz, Ramiro Salas, Philip R. Baldwin, Alireza Soltani, Ulrich Kirk, Paul M. Cinciripini, and P. Read Montague. "Belief about Nicotine Modulates Subjective Craving and Insula Activity in Deprived Smokers." *Frontiers in Psychiatry* 7 (2016). https://www.frontiersin.org/articles/10.3389/fpsyt.2016.00126.
20. Hone-Blanchet, Antoine, Tobias Wensing, and Shirley Fecteau. "The Use of Virtual Reality in Craving Assessment and Cue-Exposure Therapy in Substance Use Disorders." *Frontiers in Human Neuroscience* 8 (2014). https://www.frontiersin.org/articles/10.3389/fnhum.2014.00844.
21. Jansen, Jochem M., Joost G. Daams, Maarten W. J. Koeter, Dick J. Veltman, Wim van den Brink, and Anna E. Goudriaan. "Effects of Non-Invasive Neurostimulation on Craving: A Meta-Analysis." *Neuroscience & Biobehavioral Reviews* 37, no. 10, Part 2 (December 1, 2013): 2472–80. https://doi.org/10.1016/j.neubiorev.2013.07.009.
22. Kao, Chang-Hao, Gloria W. Feng, Jihyun K. Hur, Huw Jarvis, and Robb B. Rutledge. "Computational Models of Subjective Feelings in Psychiatry." *Neuroscience & Biobehavioral Reviews* 145 (February 1, 2023): 105008. https://doi.org/10.1016/j.neubiorev.2022.105008.
23. Katahira, Kentaro. "The Statistical Structures of Reinforcement Learning with Asymmetric Value Updates." *Journal of Mathematical Psychology* 87 (December 1, 2018): 31–45. https://doi.org/10.1016/j.jmp.2018.09.002.
24. Kober, Hedy, Peter Mende-Siedlecki, Ethan F. Kross, Jochen Weber, Walter Mischel, Carl L. Hart, and Kevin N. Ochsner. "Prefrontal–Striatal Pathway Underlies Cognitive Regulation of Craving." *Proceedings of the National Academy of Sciences* 107, no. 33 (August 17, 2010): 14811–16. https://doi.org/10.1073/pnas.1007779107.
25. Konova, Anna B., Kenway Louie, and Paul W. Glimcher. "The Computational Form of Craving Is a Selective Multiplication of Economic Value." *Proceedings of the National Academy of Sciences* 115, no. 16 (April 17, 2018): 4122–27. https://doi.org/10.1073/pnas.1714443115.
26. Koob, George F., and Michel Le Moal. "Plasticity of Reward Neurocircuitry and the 'dark Side' of Drug Addiction." *Nature Neuroscience* 8, no. 11 (November 2005): 1442–44. https://doi.org/10.1038/nn1105-1442.
27. Koob, George F, and Nora D Volkow. "Neurobiology of Addiction: A Neurocircuitry Analysis." *The Lancet Psychiatry* 3, no. 8 (August 1, 2016): 760–73. https://doi.org/10.1016/S2215-0366(16)00104-8.
28. Koob, George F., and Nora D. Volkow. "Neurocircuitry of Addiction." *Neuropsychopharmacology* 35, no. 1 (January 2010): 217–38. https://doi.org/10.1038/npp.2009.110.
29. Li, Peng, Ping Wu, Xue Xin, Yun-Li Fan, Gui-Bin Wang, Fan Wang, Meng-Ying Ma, et al. "Incubation of Alcohol Craving during Abstinence in Patients with Alcohol Dependence." *Addiction Biology* 20, no. 3 (2015): 513–22. https://doi.org/10.1111/adb.12140.

30. Li, Wen, Eric L. Garland, and Matthew O. Howard. "Therapeutic Mechanisms of Mindfulness-Oriented Recovery Enhancement for Internet Gaming Disorder: Reducing Craving and Addictive Behavior by Targeting Cognitive Processes." *Journal of Addictive Diseases* 37, no. 1–2 (April 3, 2018): 5–13. https://doi.org/10.1080/10550887.2018.1442617.
31. Limbrick-Oldfield, E. H., I. Mick, R. E. Cocks, J. McGonigle, S. P. Sharman, A. P. Goldstone, P. R. A. Stokes, et al. "Neural Substrates of Cue Reactivity and Craving in Gambling Disorder." *Translational Psychiatry* 7, no. 1 (January 2017): e992–e992. https://doi.org/10.1038/tp.2016.256.
32. Luijten, Maartje, Arnt F. Schellekens, Simone Kühn, Marise W. J. Machielse, and Guillaume Sescousse. "Disruption of Reward Processing in Addiction : An Image-Based Meta-Analysis of Functional Magnetic Resonance Imaging Studies." *JAMA Psychiatry* 74, no. 4 (April 1, 2017): 387–98. https://doi.org/10.1001/jamapsychiatry.2016.3084.
33. Lüscher, Christian, Trevor W. Robbins, and Barry J. Everitt. "The Transition to Compulsion in Addiction." *Nature Reviews Neuroscience* 21, no. 5 (May 2020): 247–63. https://doi.org/10.1038/s41583-020-0289-z.
34. Myers, Catherine E., Jony Sheynin, Tarryn Balsdon, Andre Luzardo, Kevin D. Beck, Lee Hogarth, Paul Haber, and Ahmed A. Moustafa. "Probabilistic Reward- and Punishment-Based Learning in Opioid Addiction: Experimental and Computational Data." *Behavioural Brain Research* 296 (January 1, 2016): 240–48. https://doi.org/10.1016/j.bbr.2015.09.018.
35. Robinson, Terry E., and Kent C. Berridge. "The Neural Basis of Drug Craving: An Incentive-Sensitization Theory of Addiction." *Brain Research Reviews* 18, no. 3 (September 1, 1993): 247–91. https://doi.org/10.1016/0165-0173(93)90013-P.
36. Rutledge, Robb B., Nikolina Skandali, Peter Dayan, and Raymond J. Dolan. "A Computational and Neural Model of Momentary Subjective Well-Being." *Proceedings of the National Academy of Sciences* 111, no. 33 (August 19, 2014): 12252–57. https://doi.org/10.1073/pnas.1407535111.
37. Schulte, Erica M., Carlos M. Grilo, and Ashley N. Gearhardt. "Shared and Unique Mechanisms Underlying Binge Eating Disorder and Addictive Disorders." *Clinical Psychology Review* 44 (March 1, 2016): 125–39. https://doi.org/10.1016/j.cpr.2016.02.001.
38. Shalev, Uri, Marisela Morales, Bruce Hope, Jasmine Yap, and Yavin Shaham. "Time-Dependent Changes in Extinction Behavior and Stress-Induced Reinstatement of Drug Seeking Following Withdrawal from Heroin in Rats." *Psychopharmacology* 156, no. 1 (June 1, 2001): 98–107. https://doi.org/10.1007/s002130100748.
39. Sharot, Tali, Benedetto De Martino, and Raymond J. Dolan. "How Choice Reveals and Shapes Expected Hedonic Outcome." *Journal of Neuroscience* 29, no. 12 (March 25, 2009): 3760–65. https://doi.org/10.1523/JNEUROSCI.4972-08.2009.
40. Smith, Ryan, Samuel Taylor, and Edda Bilek. "Computational Mechanisms of Addiction: Recent Evidence and Its Relevance to Addiction Medicine." *Current Addiction Reports* 8, no. 4 (December 1, 2021): 509–19. https://doi.org/10.1007/s40429-021-00399-z.
41. Tiffany, Stephen T. "Potential Functions of Classical Conditioning in Drug Addiction." In *Addictive Behaviour: Cue Exposure Theory and Practice*, 47–71. The Wiley Series in Clinical Psychology. Oxford, England: John Wiley & Sons, 1995.
42. Tiffany, Stephen T., and Jennifer M. Wray. "The Clinical Significance of Drug Craving." *Annals of the New York Academy of Sciences* 1248 (February 2012): 1–17. https://doi.org/10.1111/j.1749-6632.2011.06298.x.
43. Vinckier, Fabien, Lionel Rigoux, Delphine Oudiette, and Mathias Pessiglione. "Neuro-Computational Account of How Mood Fluctuations Arise and Affect Decision Making." *Nature Communications* 9, no. 1 (April 26, 2018): 1708. https://doi.org/10.1038/s41467-018-03774-z.

44. Volkow, Nora D., Roy A. Wise, and Ruben Baler. “The Dopamine Motive System: Implications for Drug and Food Addiction.” *Nature Reviews Neuroscience* 18, no. 12 (December 2017): 741–52. https://doi.org/10.1038/nrn.2017.130.
45. Zorowitz, Samuel, Ida Momennejad, and Nathaniel D. Daw. “Anxiety, Avoidance, and Sequential Evaluation.” *Computational Psychiatry (Cambridge, Mass.)* 4 (2020): 10.1162/cpsy_a_00026. https://doi.org/10.1162/cpsy_a_00026.

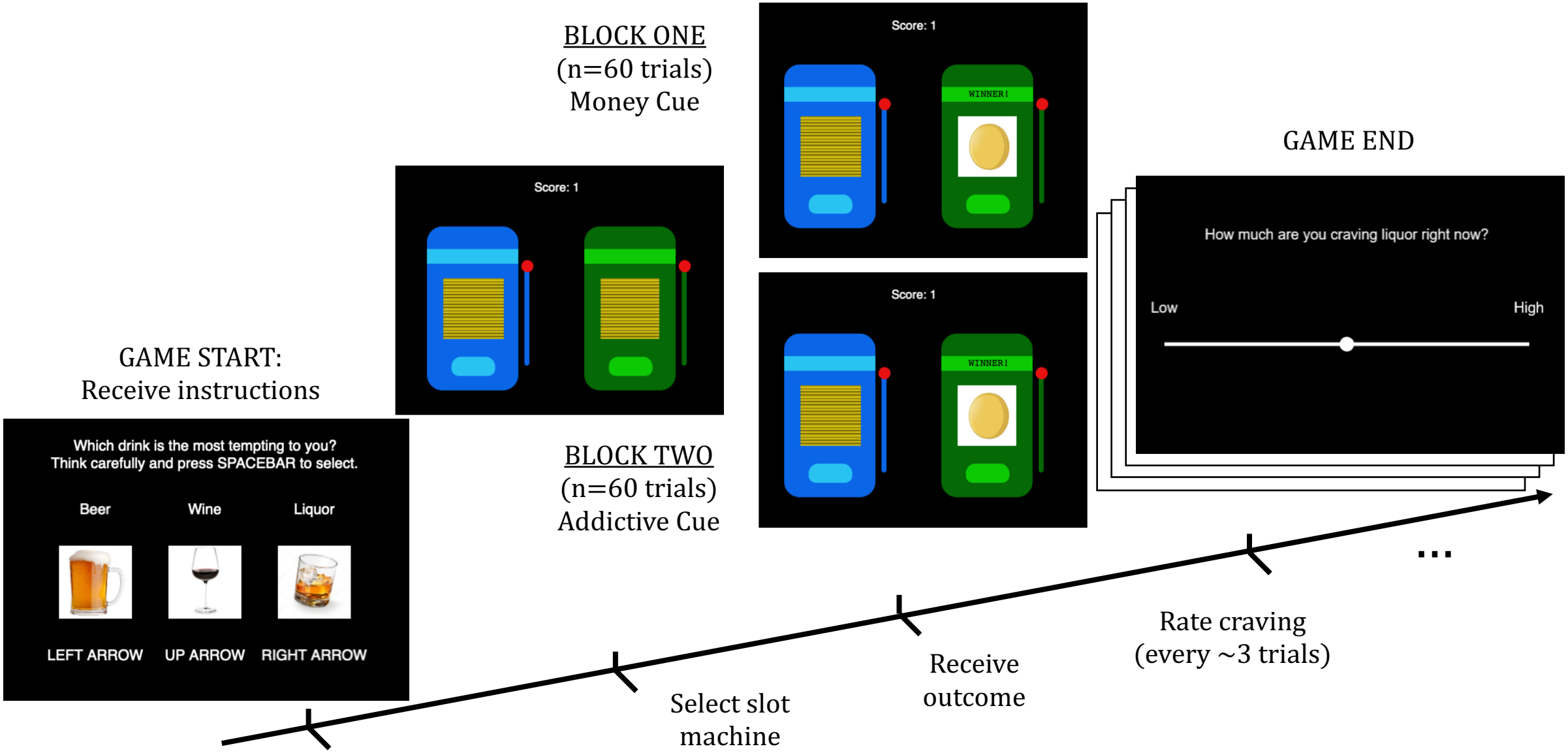


Fig. 1. Specification of experimental procedure.

Fig. 2. Overview of computational modeling approach.

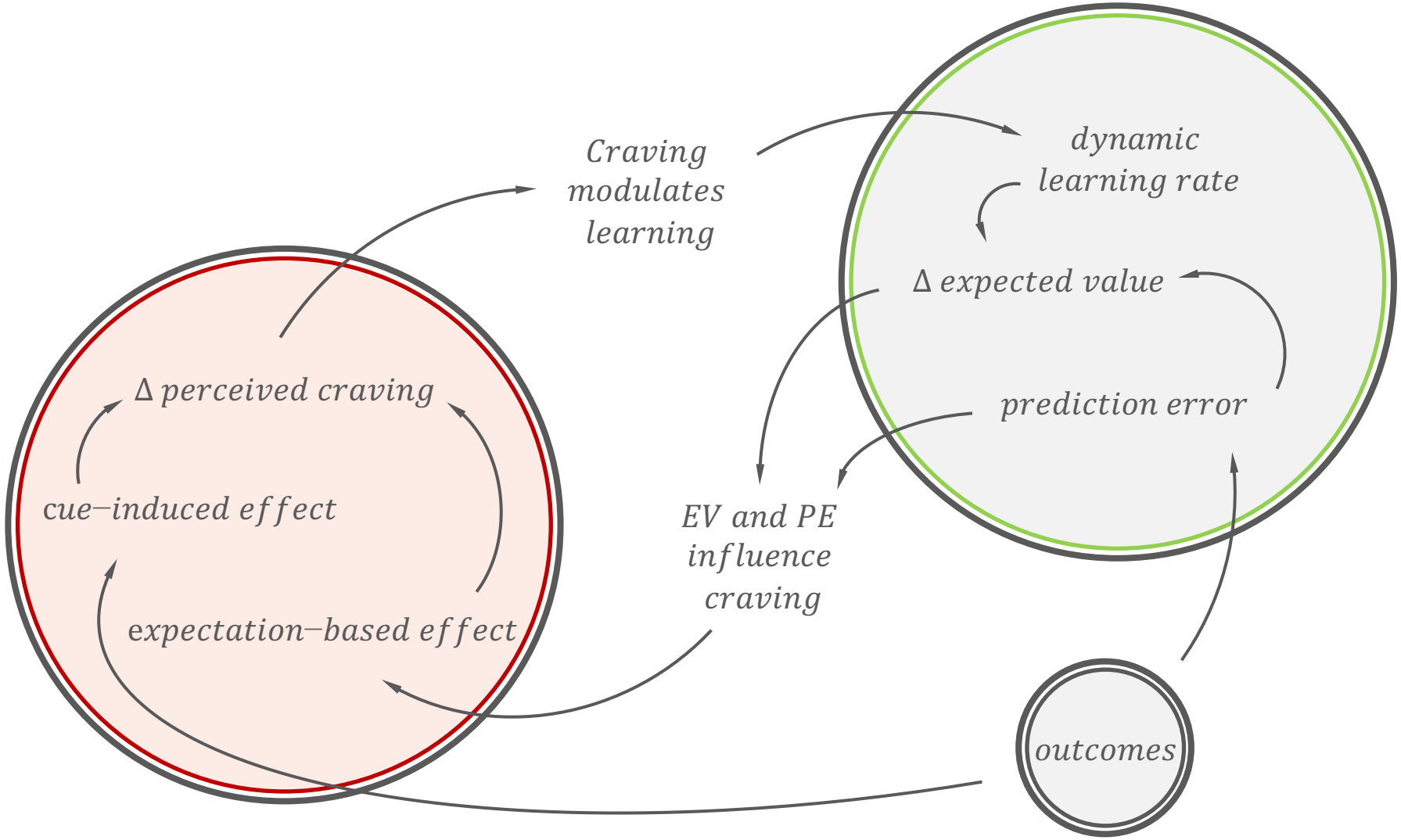


Table. 1a. Model specification for decision-making and craving computations.

Decision Model	Description	Value update scheme	Decision policy
Learning rate bias ($\alpha\ bias$)	Momentary craving modulates the learning rate at each trial.	TDRL with ***dynamic*** learning rates	Softmax with static temperature
Temperature bias ($\beta\ bias$)	Momentary craving modulates the inverse temperature	TDRL with static learning rates	Softmax with ***dynamic*** temperature
Standard TDRL	Craving does not modulate the value update or action policy	TDRL with static learning rates	Softmax with static temperature
Heuristic	Decisions made with simple switching rule.	–	Win-stay-lose-two-shift

Table. 1b. Model specification for decision-making and craving computations.

Craving Model	Description	Components of craving computation
Expectation-influenced (EI)	Prior expectations and cue-elicited effects both contribute to perceived craving	Expected value, prediction error, and cue-elicited effect
Cue-elicited craving (CEC)	Only cues contribute to perceived craving	Cue-elicited effect only

Joint models	Abbreviation	Decision Model	Craving Model
$TDRL + CEC$	*tdrl*	Standard TDRL	Cue-elicited craving only
$\alpha\ bias + EI$	$\alpha\ bias$	Learning rate bias	Expectation-influenced
$\beta\ bias + EI$	$\beta\ bias$	Temperature bias	Expectation-influenced
$RW + EI$	*passive*	Standard TDRL	Expectation-influenced
$Heu + CEC$	*heu*	Heuristic	Cue-elicited craving only

Fig. 3. Model comparison for choice behavior and craving.

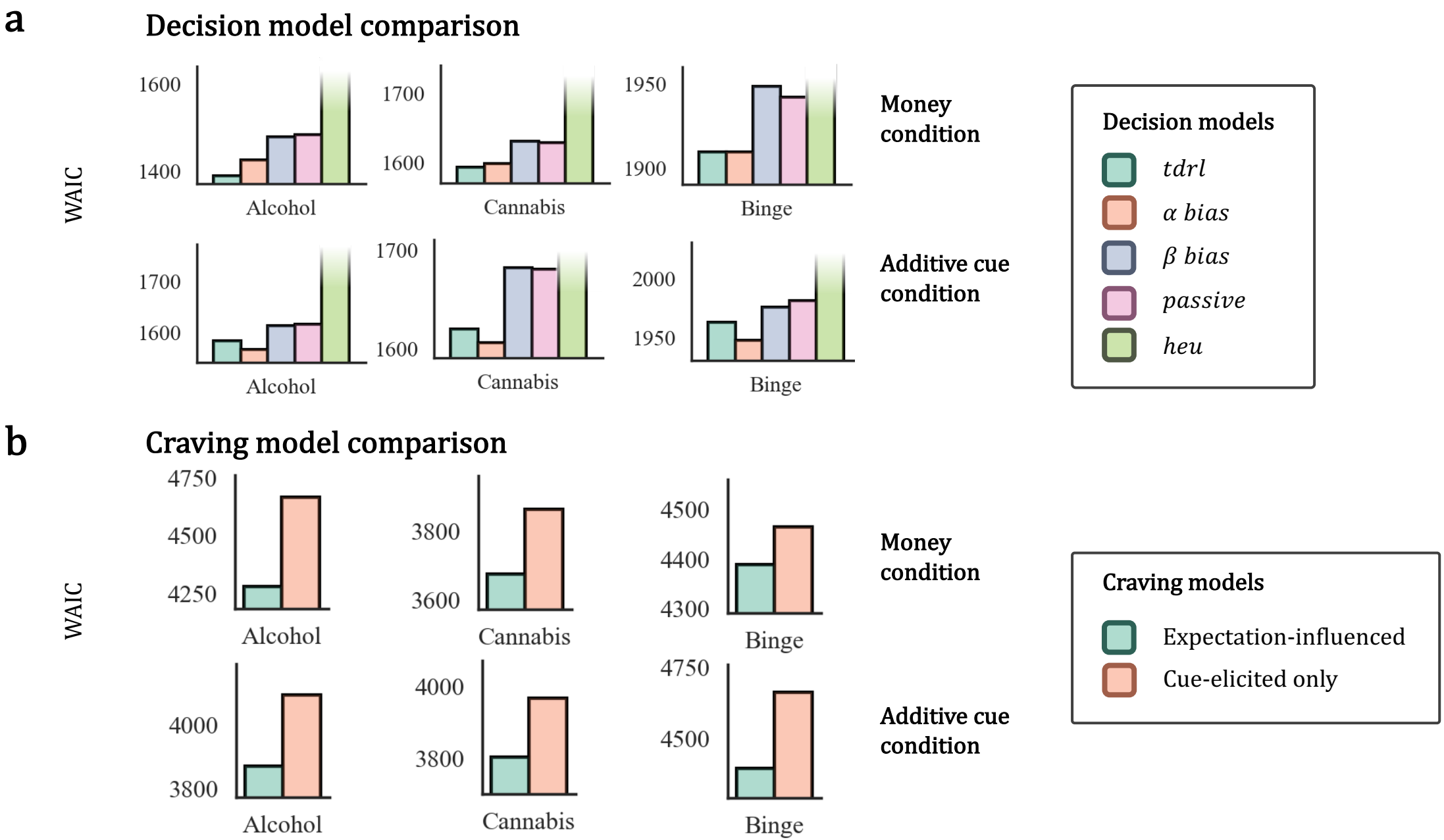


Fig. 4. Parameter distributions for addictive cue condition.

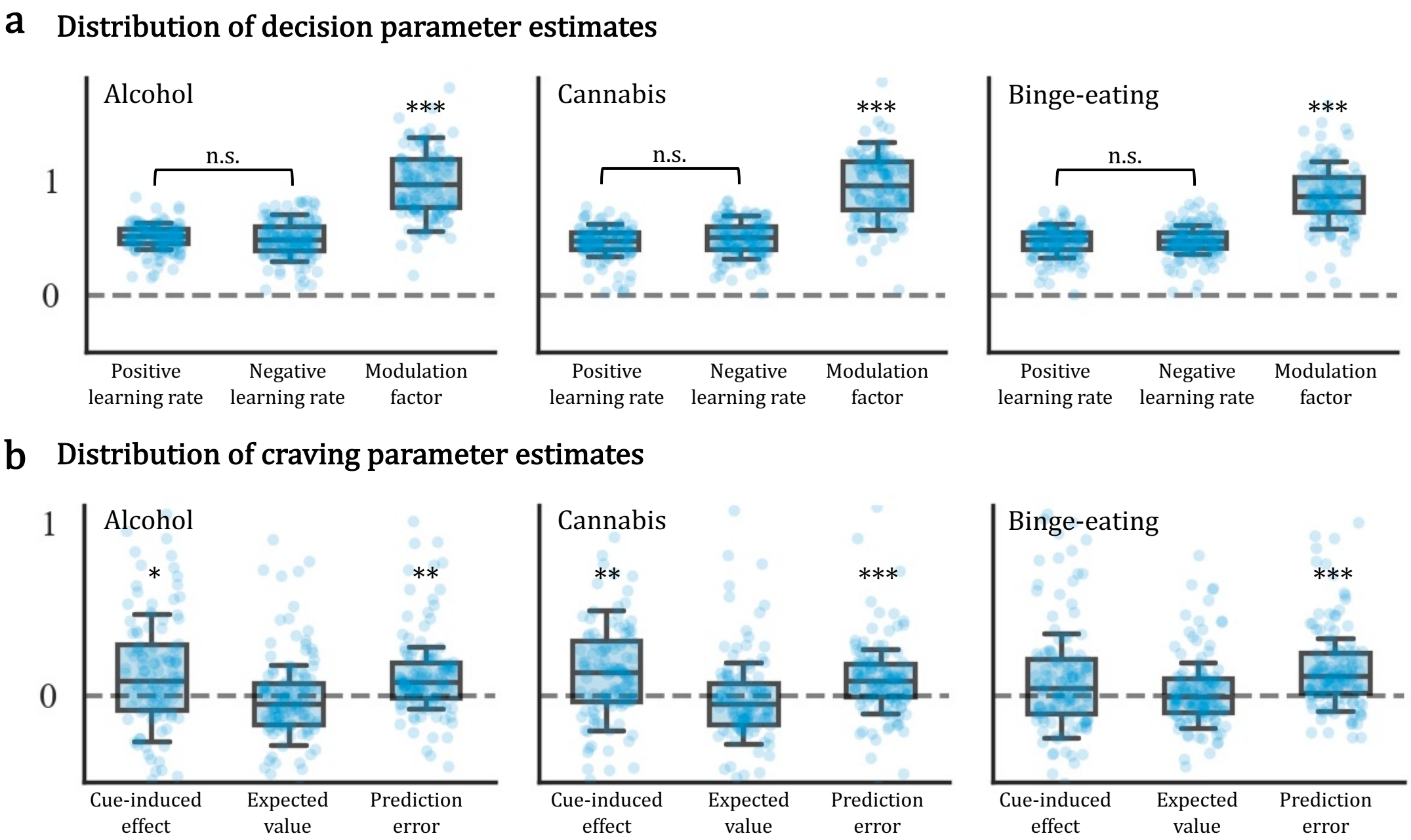
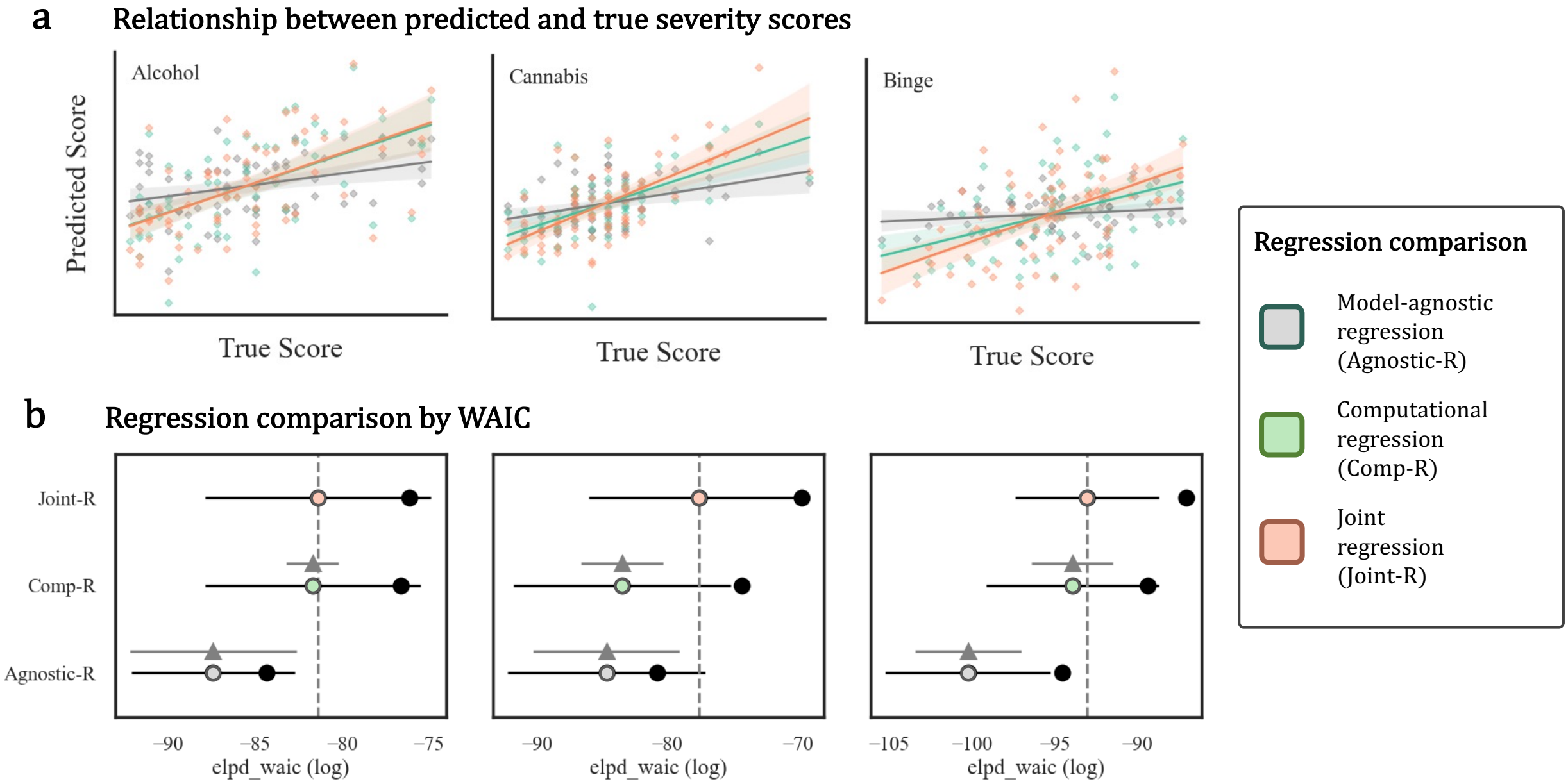


Fig. 5. Regression analyses for clinical severity scores.

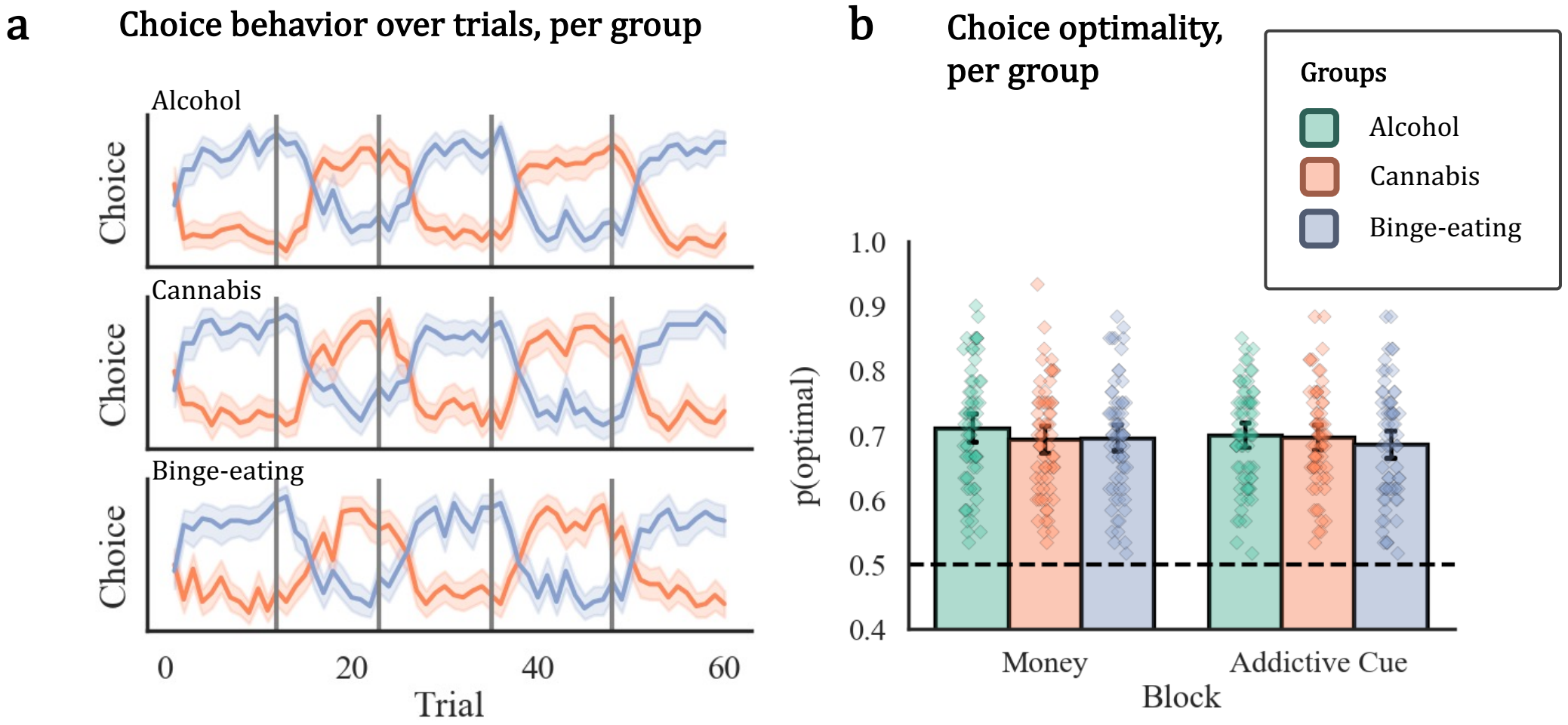


Supplementary Figures

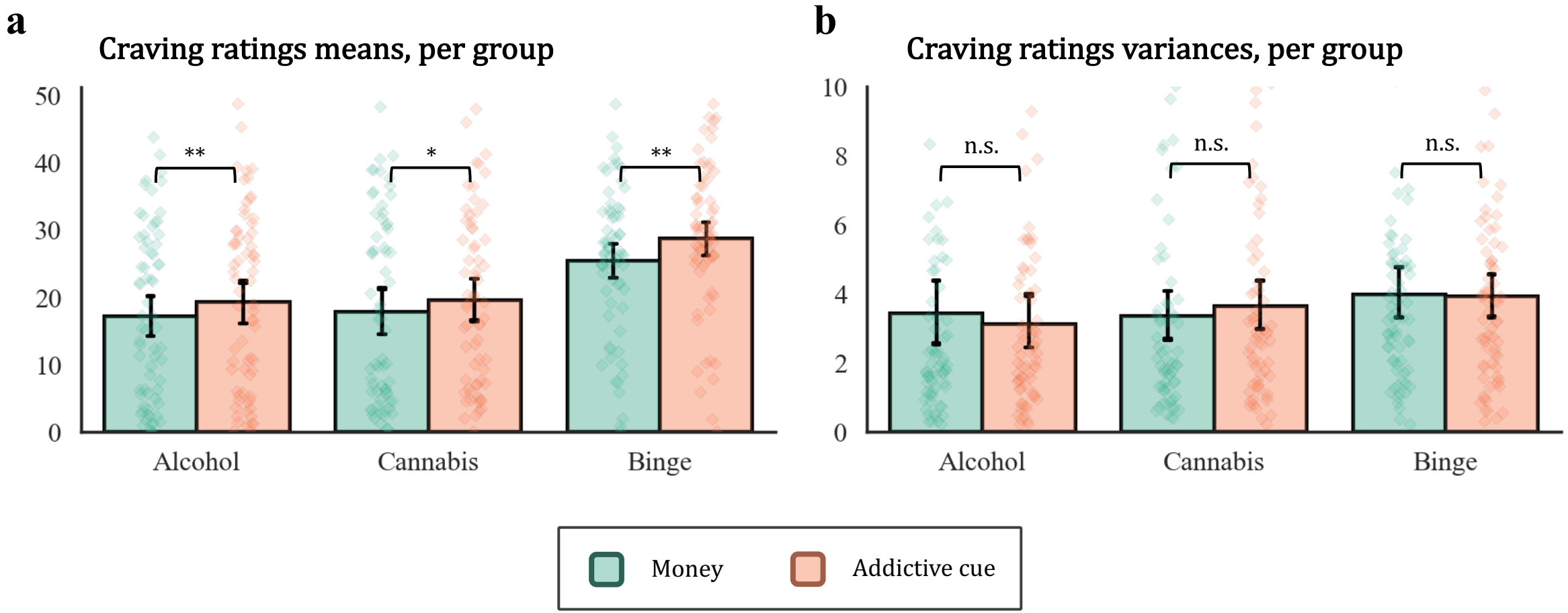
Supplementary Table 1. Original sample demographics and clinical measure scores.

Variable	Alcohol (n=68)	Cannabis (n=65)	Binge (n=69)
Age (s.d.)	41.0 (13.4)	38.7 (14.2)	38.1 (15.4)
Sex (%)			
Male	42 (61.8%)	34 (52.3%)	26 (37.7%)
Female	25 (36.7%)	31 (47.7%)	43 (62.3%)
Other	1 (1.5%)	0 (0%)	0 (0%)
Education level (%)			
≤ High school	8 (11.8%)	7 (10.7%)	16 (23.2%)
College	43 (63.2%)	43 (66.2%)	41 (59.4%)
≥ Graduate	17 (25.0%)	15 (23.1%)	12 (17.4%)
ASSIST (s.d.)	17.1 (8.4)	14.0 (7.9)	-
AUDIT (s.d.)	6.8 (6.6)	-	-
CAST (s.d.)	-	7.7 (3.6)	-
EDE-Q (s.d.)	-	-	72.6 (28.4)
LOCES (s.d.)	-	-	18.7 (6.1)

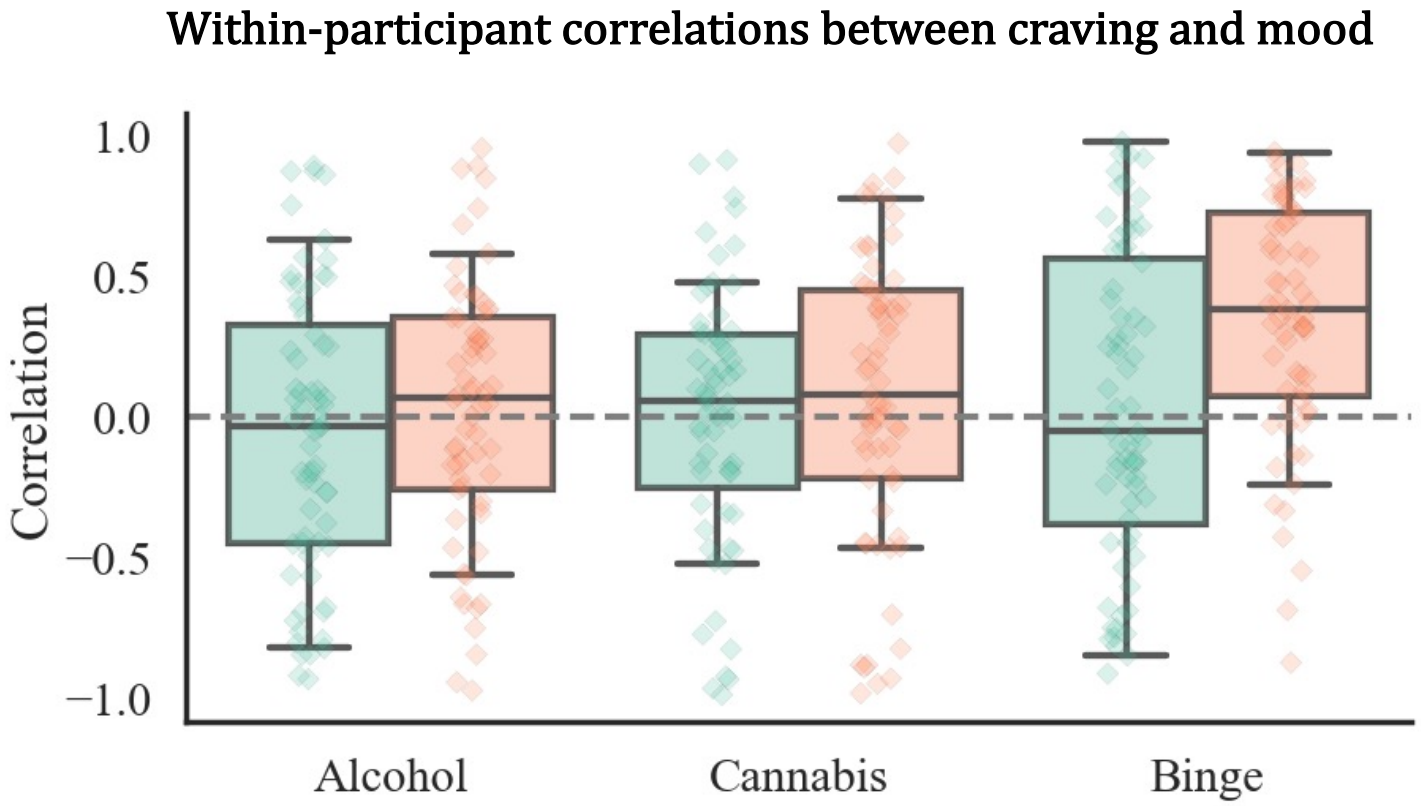
Supplementary Fig. 1. Choice optimality.



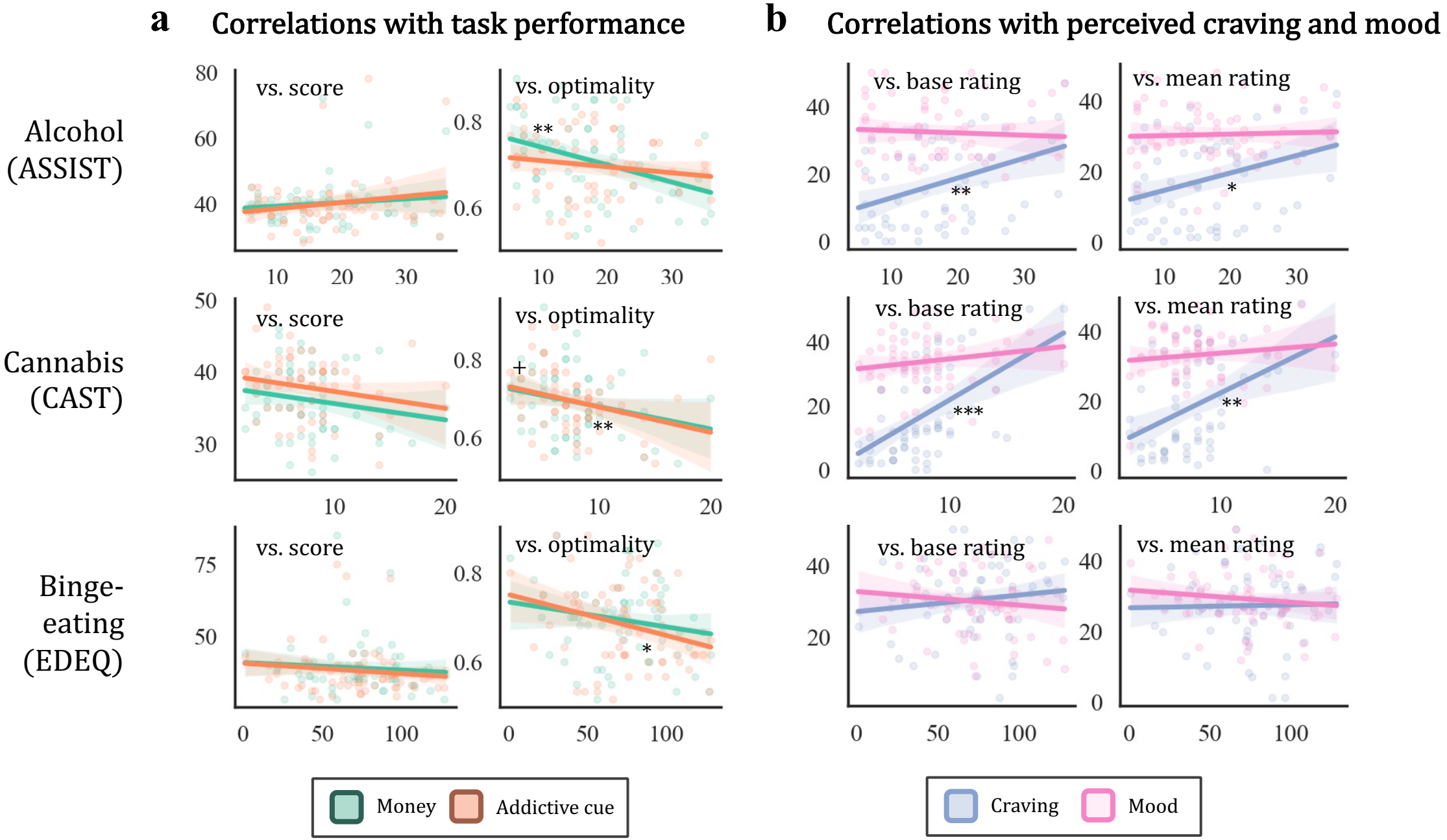
Supplementary Fig. 2. Craving rating means and variances across groups.



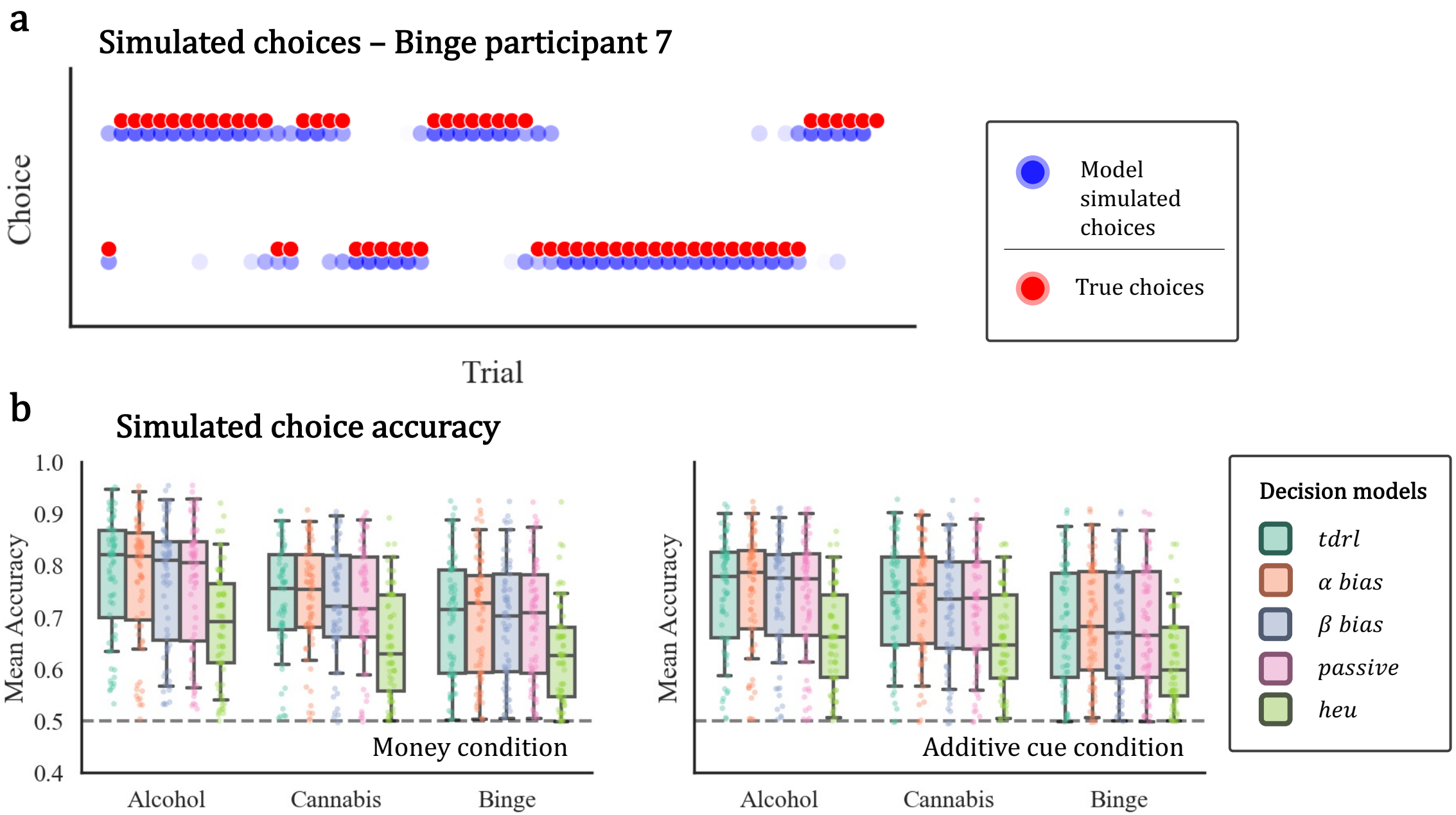
Supplementary Fig. 3. Correlations between craving and mood ratings.



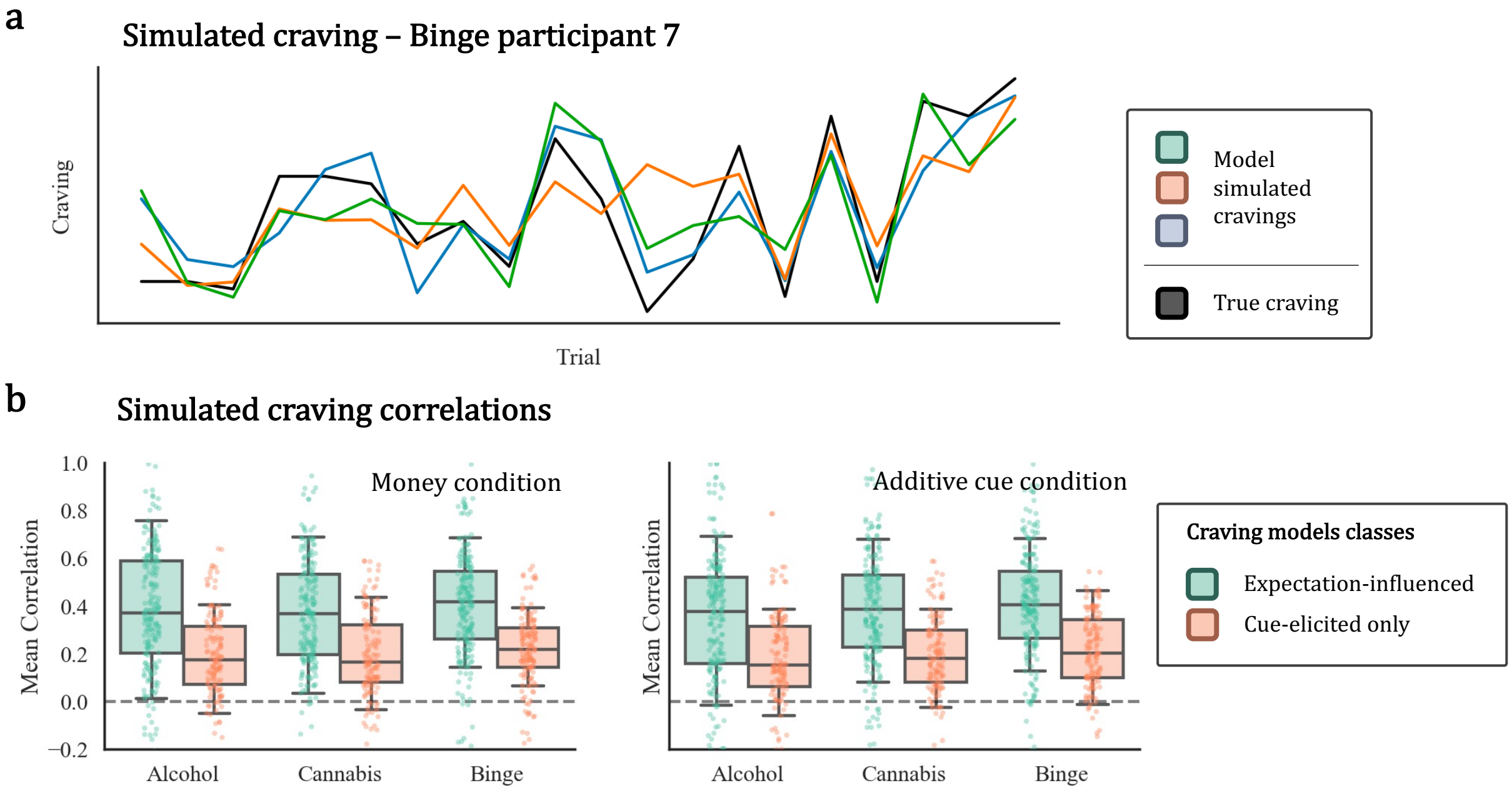
Supplementary Fig. 4. Correlations between severity scores, task performance and craving/mood ratings.



Supplementary Fig. 5. Simulated choice accuracy and qualitative match.



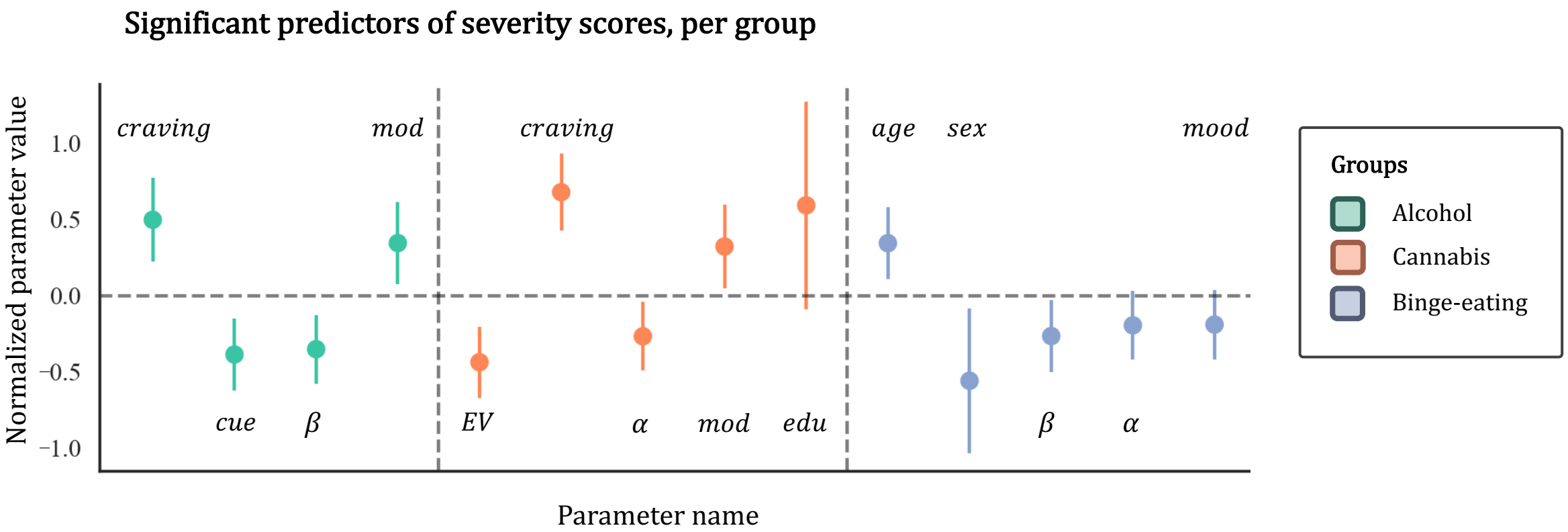
Supplementary Fig. 6. Simulated craving correlations and qualitative match.



Supplementary Table 2. Model comparison for clinical severity regressions.

Group	Regression Model	R2	R2_adj	AIC	ELPD-WAIC
Alcohol	Joint-R	0.336503	0.289942	160.514027	-81.400854
	Comp-R	0.326617	0.279362	161.431067	-81.698966
	Agnostic-R	0.127891	0.098328	173.464236	-87.434490
Cannabis	Joint-R	0.431466	0.379781	150.664138	-77.739369
	Comp-R	0.330638	0.269787	160.623288	-83.457107
	Agnostic-R	0.166317	0.152187	166.014470	-84.596829
Binge	Joint-R	0.235683	0.173034	184.129968	-93.074263
	Comp-R	0.167034	0.127369	185.892656	-93.954178
	Agnostic-R	0.029822	-0.032770	198.109256	-100.266604

Supplementary Table 2. Parameter weights for best performing regression per group.

Group	Best-performing regression	Parameter	Coef.	Std.Err.	t	P>|t|	[0.025	0.975]
Alcohol	Joint-R	mean_craving_ratings	0.4975	0.1364	3.6483	0.0006	0.2244	0.7705
		cec_weight	-0.3873	0.1185	-3.2679	0.0018	-0.6246	-0.1500
		beta	-0.3523	0.1127	-3.1259	0.0028	-0.5781	-0.1266
		mod	0.3448	0.1338	2.5772	0.0126	0.0769	0.6128
Cannabis	Joint-R	mean_craving_ratings	0.6790	0.1268	5.3533	0.0000	0.4248	0.9331
		weight_one	-0.4380	0.1162	-3.7685	0.0004	-0.6709	-0.2051
		alpha_pos	-0.2676	0.1129	-2.3710	0.0213	-0.4938	-0.0414
		mod	0.3222	0.1376	2.3411	0.0229	0.0464	0.5981
		DEMO_edu_High_School	0.5929	0.3406	1.7408	0.0873	-0.0896	1.2754
Binge	Joint-R	DEMO_age	0.3441	0.119	2.902	0.005	0.107	0.581
		DEMO_sex_Male	-0.5601	0.238	-2.358	0.022	-1.035	-0.085
		beta	-0.2661	0.118	-2.252	0.028	-0.502	-0.030
		alpha_pos	-0.1957	0.113	-1.737	0.087	-0.421	0.030
		mean_mood_ratings	-0.1906	0.114	-1.669	0.100	-0.419	0.038

Supplementary Fig. 7. Significant predictors from regression, per group.