

Computational models of exploration and exploitation characterise onset and efficacy of treatment in methamphetamine use disorder

Alex H. Robinson  | Trevor T.-J. Chong  | Antonio Verdejo-Garcia 

Turner Institute for Brain and Mental Health,
Monash University, Melbourne

Correspondence

Antonio Verdejo-Garcia, Turner Institute for
Brain and Mental Health, Monash University,
18 Innovation Walk, Clayton, Melbourne,
Victoria 3800 Australia.
Email: antonio.verdejo@monash.edu

Funding information

Australian Medical Research Future Fund,
Grant/Award Number: MRF1141214;
Australian Research Council, Grant/Award
Numbers: DE 180100389, DP 180102383;
Australian Government Research Training
Program Scholarship

Abstract

People with Methamphetamine Use Disorder (PwMUD) spend substantial time and resources on substance use, which hinders their ability to explore alternate reinforcers. Gold-standard behavioural treatments attempt to remedy this by encouraging action towards non-drug reinforcers, but substance use often persists. We aimed to unravel the mechanistic drivers of this behaviour by applying a computational model of explore/exploit behaviour to decision-making data (Iowa Gambling Task) from 106 PwMUD and 48 controls. We then examined the longitudinal link between explore/exploit mechanisms and changes in methamphetamine use 6 weeks later. Exploitation parameters included reinforcement sensitivity and inverse decay (i.e., number of past outcomes used to guide choices). Exploration parameters included maximum directed exploration value (i.e., value of trying novel actions). The Timeline Follow Back measured changes in methamphetamine use. Compared to controls, PwMUD showed deficits in exploitative decision-making, characterised by reduced reinforcement sensitivity, $U = 3065$, $p = 0.009$, and less use of previous choice outcomes, $U = 3062$, $p = 0.010$. This was accompanied by a behavioural pattern of frequent shifting between choices, which appeared consistent with random exploration. Furthermore, PwMUD with greater reductions of methamphetamine use at 6 weeks had increased directed exploration ($\beta = 0.22$, $p = 0.045$); greater use of past choice outcomes ($\beta = -0.39$, $p = 0.002$) and greater choice consistency ($\beta = -0.39$, $p = 0.002$). Therefore, limited computational exploitation and increased behavioural exploration characterise PwMUD's presentation to treatment, while increased directed exploration, use of past choice outcomes and choice consistency predict greater reductions of methamphetamine use.

KEYWORDS

computational modelling, decision-making, explore/exploit, methamphetamine, predictive, substance use

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Addiction Biology* published by John Wiley & Sons Ltd on behalf of Society for the Study of Addiction.

1 | INTRODUCTION

People with Methamphetamine Use Disorder (PwMUD) spend substantial time and resources on their substance use, which hinders their ability to act towards alternative reinforcers.^{1,2} Clinically, we see this in the persistent preoccupation with methamphetamine use³; the decreased ability to complete non-drug-related goals⁴ and the impoverished health, social relationships and quality of life that accompany their presentation to treatment.⁵⁻⁷ Furthermore, despite treatment interventions attempting to shift people away from methamphetamine use and towards adaptive, non-drug-related activities (i.e., positive social groups, treatment attendance and personal goals),⁸ substance use often persists.^{9,10}

Computational models of the exploration/exploitation framework^{2,11} are a novel approach to investigate the mechanisms preventing behavioural change in PwMUD. These models assess whether an organism takes a new action to seek out greater rewards (exploration) or instead repeats familiar actions already associated with a known reward (exploitation).^{2,12} Exploration can be further separated into directed exploration (goal-oriented behaviours to gain information about an unfamiliar action) and random exploration (exploration without the conscious goal of gaining new information).^{2,13,14}

The prevailing view in the addiction literature is that persistent substance use is linked to a bias towards exploitative decision-making.^{2,15,16} However, evidence supporting this hypothesis is mixed and varies depending on the substance of concern and form of analysis (i.e., basic behavioural outcomes versus computational modelling).¹⁵⁻¹⁸ To overcome this challenge, more recent computational approaches now avoid the dichotomous, behavioural separation of explore/exploit biases, and instead provide explanations of how dysfunction may be driven in either process. For example, Smith et al.¹⁸ found that explore/exploit dysfunction was driven by hypo-sensitivity to loss, hypersensitivity to reward and increased random exploration amongst a mixed sample of Substance Use Disorders. In another example, Morris et al.¹⁵ found that people with Alcohol Use Disorder had reduced effectiveness in exploiting reward, driven by decreased exploration after reinforcement. As these two findings provide more nuanced explanations of explore/exploit decision-making (when compared to summary behavioural measures), computational modelling appears to be a promising measure to identify underlying cognitive processes relevant to PwMUD's treatment.¹⁹

The Iowa Gambling Task (IGT)²⁰ is one paradigm that appears well-suited to measure the mechanistic underpinnings of PwMUD's explore/exploit behaviour. This is because advantageous performance on the task requires an initial exploration of the available choices, followed by a subsequent exploitation of the most advantageous choices. Such a pattern is clinically relevant as it mimics the efforts of PwMUD to progressively explore and exploit non-drug reinforcers across their recovery. Furthermore, the IGT is a useful measure to identify decision-making differences between PwMUD and healthy controls and has been linked to clinical outcomes in PwMUD. For example, similar to other substance use disorders, early IGT studies were able to identify generalised and/or 'risky' decision-making in

PwMUD compared to controls.²¹⁻²⁴ Following these findings, the IGT has since identified relationships between generalised decision-making dysfunction and clinical outcomes in PwMUD, such as length of abstinence²⁵ and treatment response.^{26,27} However, while the IGT has been computationally modelled to identify specific decision-making mechanisms in other stimulant use disorders,^{28,29} such an approach has not yet been implemented in PwMUD, and the vast majority of the above findings are based on a single summary metric of performance (i.e., net score).

Recently, a computational model has been developed that can sensitively measure the underlying, computational parameters involved in explore/exploit decisions on the IGT. This model, the Value plus Sequential Exploration model (VSE),¹³ quantifies the strength of multiple, interlinked mechanisms involved in exploitation (i.e., sensitivity to reinforcement and number of previous choices used to guide current decision) and directed exploration (i.e., value associated to gaining new information and how frequently people reach this value). It then uses these individual parameters to compute the overall values of exploring and exploiting, which can then be used to predict participants' choice behaviour. The inclusion of these directed exploration parameters separates the VSE from previous models of the IGT²⁹⁻³² and allows a mechanistic understanding of PwMUD's behaviour on the task.

Our study had two aims. First, we sought to reveal the computational mechanisms driving maladaptive explore/exploit decisions amongst treatment-seeking PwMUD. This was achieved by applying the VSE model to behaviour on the IGT and then comparing the parameters characterising explore/exploit behaviour between PwMUD and drug-naïve controls. Second, we investigated whether the baseline explore/exploit parameters in PwMUD could predict changes in methamphetamine use six-weeks after treatment engagement—which is a vulnerable period for drug reduction amongst PwMUD.³³ Both these aims were exploratory and did not include any *a priori* hypotheses.

2 | MATERIALS AND METHODS

2.1 | Design

We applied the VSE model of the IGT to a large, previously collected, dataset comparing PwMUD and drug-naïve controls.³⁴ Our first aim of identifying group differences in explore/exploit computational parameters used a cross-sectional design. Our second aim of identifying whether baseline parameters of explore/exploit decisions predicted changes in methamphetamine use after 6 weeks used a prospective design.

2.2 | Participants

Data were available for 106 PwMUD and 48 control participants, of which 79 PwMUD (74.53%) returned for follow-up. PwMUD were

TABLE 1 Descriptive statistics of sociodemographic characteristics in PwMUD and controls

	PwMUD	Controls	Frequentist Mann. Whit.	Bayes factor Mann. Whit.
Sex (F/M)	27F/79M	12F/36M	$p = 0.95$	
Age	31.20 (7.25)	31.59 (8.67)	$p = 0.65$	$BF_{10} = 0.18$
Education (years)	12.93 (2.25)	13.26 (2.07)	$p = 0.32$	$BF_{10} = 0.29$
FSIQ	96.19 (10.96)	101.92 (11.36)	$p = 0.007$	$BF_{10} = 5.14$
Employed (any)	28 (26.92%)	26 (54.17%)	$p = 0.001$	
Depression (CES-D)	28.51 (12.30)	7.25 (5.93)	$p < 0.001$	$BF_{10} = 87\ 412.26$
Sociodem. Status	6.97 (2.41)	7.92 (1.60)	$p = 0.042$	$BF_{10} = 1.019$

Note: Values represent means with standard deviations in parentheses.
Abbreviation: CES-D, Centre for Epidemiologic Studies Depression Scale.

TABLE 2 PwMUDs baseline methamphetamine and other common substance use

	M	SD	Range
Severity of Dependence Meth	11.10	3.15	[1–15]
Severity of Dependence Cann.	2.63	4.24	[0–15]
Severity of Dependence Alch.	1.71	3.30	[0–14]
Methamphetamine			
Daily dose (grammes)	0.71	0.56	[0.03–4.00]
Frequency (days/month)	23.30	9.25	[4–31]
Duration of use (years)	6.95	4.90	[0.6–30]
Other substance use			
Cigarettes <i>p</i> month	407.13	257.50	[0–930]
Standard drinks <i>p</i> month	75.34	160.76	[0–775]
Cannabis <i>p</i> month (grammes)	24.89	42.21	[0–217]
	N	%	
Meth. route of Admin.			
Smoking	74	69.81%	
Injecting	25	23.59%	
Multiple	7	6.60%	

Note: Severity of Dependence Scale scores range between 0 and 15, with those >4 indicating problematic methamphetamine use,³⁷ ≥3 indicating likely Alcohol Use Disorder³⁸ and ≥4 indicating likely Cannabis Use Disorder.³⁹

recruited during the first 3 weeks of treatment from private and public treatment settings in Melbourne, Australia. These included inpatient rehabilitation and detoxification services, as well as outpatient counselling services. The inclusion criteria for PwMUD were (1) a primary diagnosis of methamphetamine dependence identified by the Structured Clinical Interview for the DSM-IV (SCID-IV)³⁵; (2) self-reported abstinence from methamphetamine for more than 48 h (to rule out acute drug effects) but less than 21 days (to prevent heterogeneity in recovery stages) and (3) no secondary substance dependence (excluding alcohol, cannabis and tobacco, measured by SCID-IV). Controls were required to have never used methamphetamine and to not meet criteria for any substance dependence. All participants were required to have never experienced (1) loss of

consciousness greater than 30 min (self-reported); (2) a diagnosis of bipolar, schizophrenia or other psychotic disorder (screened by the SCID-IV); (3) intellectual disability (screened by the Wechsler Abbreviated Scale of Intelligence II; WASI-II)³⁶ and (4) neurological disorders (self-reported). Table 1 presents each group's sociodemographic statistics. Table 2 presents substance use in PwMUD.

2.3 | Procedure

Participants completed the IGT as part of a larger cognitive assessment, from which results have been reported elsewhere.^{24,34,40,41} Testing sessions occurred at treatment centres or quiet, easily accessible spaces (e.g., community libraries). The follow-up assessments had a similar structure and nature to the baseline assessment but were briefer due to the omission of the clinical interview. Measures were administered by researchers with postgraduate training in clinical psychology and followed well-established guidelines to ensure reliability.⁴² Study procedures were approved by the Eastern Health Human Research Ethics Committee (E52/1213) and participants received \$40 AUD in grocery gift cards.

2.4 | Measures

2.4.1 | Sociodemographic measures

Participants self-reported their sex, age, years of education and employment status. The WASI-II estimated full-scale IQ. The Centre for Epidemiologic Studies Depression Scale (CES-D)⁴³ estimated depressive symptomology. Socioeconomic status was estimated by the Australian Bureau of Statistics' Socio-Economic Indexes for Areas.⁴⁴

2.4.2 | History of methamphetamine and secondary drug use

A modified version of the Interview for Research on Addictive Behaviours (IRAB)⁴⁵ estimated PwMUD's daily methamphetamine use

(grammes), duration of use (years), route of administration and secondary substance use. The Severity of Dependence Scale (SDS)⁴⁶ measured severity of dependence to alcohol, cannabis and methamphetamine.

2.4.3 | Iowa Gambling Task

The IGT is a computer-based measure that requires participants to make 100 selections from four decks of cards, with the goal of making as much game-based currency as possible. We used the ABCD version²⁰ in which decks C and D are 'good' (providing small wins but even smaller losses, leading to a net profit) and decks A and B are 'bad' (providing large wins but even larger losses, leading to a net loss). The likelihood of punishment also varies across decks. For example, decks A and C regularly enforce losses (50% likelihood), while decks B and D rarely enforce losses (10% likelihood).²⁸ Participants received standardised, verbal instructions outlining the task goal and its visual layout, with intentionally vague information on the reward structure ('some decks are better than others'). While the IGT has typically been measured using the net score (number of 'good' decks selected minus number of 'bad' decks; either measured once as a single outcome or measured in blocks of 20 trials across the task),⁴⁷ we analysed numerous additional measurements of choice behaviour and computational parameters using the VSE Toolbox.¹³

2.4.4 | Choice behaviour

Choice behaviour measures were used to provide an overall understanding of the overt behaviour of participants on the task. These would then be later used to provide a context to the VSE parameter findings. Beyond the traditional measure of net score, the VSE toolbox also provides behavioural measures of win/stay and lose/switch frequency, mutual information (higher values indicate that choices on subsequent trials are better predicted by the current choice), choice entropy (higher values indicate that participants evenly selected from all decks; lower values indicate that participants selected from only a few decks/one deck) and the frequency with which individuals chose three different decks over three consecutive trials ('sequential exploration 3') or four decks over four trials ('sequential exploration 4'). We note that the 'sequential exploration 3/4' outcomes have previously been referred to as 'directed exploration 3/4'.¹³ However, because high frequencies of switching decks may also represent random exploration (particularly if this is occurring in later trials with little improvement in performance), we believe this new label is more accurate. Finally, we also measured log median reaction time (after removing the first trial)⁴⁸ in order to measure task engagement. We used this approach as significantly faster responses may highlight a participant response strategy of 'rushing' at the cost of accuracy, while significantly longer responses may highlight 'mind-wandering' during the task.⁴⁹

2.4.5 | Exploration/exploitation modelling: Value plus sequential exploration model

The VSE model computes value weightings for both exploitation and directed exploration and estimates several parameters related to each process.

The Exploitation weight (*Exploit*) of a given deck, d , calculates the value of continuing to select that deck on a given trial, t . This value is influenced by an inverse decay parameter, Δ , which controls how many past outcomes are used to guide the decision ($\Delta \in [0, 1]$, with $\Delta = 0$ using only the outcome from the previous trial to update value and $\Delta = 1$ using the outcomes of all previous trials). The *Exploit* weight also incorporates a value function, v , which includes a parameter that represents the sensitivity of the individual to gains and losses on the current trial (reinforcement sensitivity, θ : $\theta \in [0, 1]$, higher values reflect greater sensitivity to both gains and losses):

$$Exploit^d(t+1) = Exploit^d(t) * \Delta + v(t), \quad (1)$$

$$v(t) = Gain(t)^\theta - Loss(t)^\theta. \quad (2)$$

A decay of previous exploit weightings is also fitted to unselected decks, using the same inverse decay parameter, Δ :

$$Exploit^d(t+1) = Exploit^d(t) * \Delta. \quad (3)$$

In contrast, the Exploration weighting (*Explore*) computes the value for participants to undertake directed exploration towards a specific deck on each trial.

Here, the *Explore* value of the selected deck is set to zero after selection (Equation 4), while unexplored decks are updated via a delta-rule to increase the drive towards exploration. This occurs via (1) a directed exploration bonus parameter (φ , unbounded), which sets the 'threshold' of how biased a person is towards directed exploration (i.e., positive values mean a person is more attracted to explore, while negative values indicate a person is more likely to continue selecting the current deck), and (2) a directed exploration learning rate (α : $\alpha \in [0, 1]$), which drives how quickly a person gets to the upper limit of exploration demand (Equation 5). In this manner, if two people had all other parameter values equal, the person with a higher exploration learning rate (α) would more frequently explore.

$$Explore^d(t+1) = 0, \quad (4)$$

$$Explore^d(t+1) = Explore^d(t) + \alpha * (\varphi - Explore^d(t)). \quad (5)$$

Finally, the overall probability of choosing a particular deck was determined by comparing the *Exploit* and *Explore* weights using a *softmax* function (Equation 6). The inverse temperature parameter, C (or consistency), reflects stochastic behaviour (with higher values representing more deterministic behaviour in line with the above computations).¹³

$$P(\text{Choice} = d) = \frac{e^{(\text{Explore}^d + \text{Exploit}^d) * C}}{\sum_{i=d}^4 e^{(\text{Explore}^i + \text{Exploit}^i) * C}}. \quad (6)$$

Using the parameters above, we can estimate (1) the relative strength of exploitative processes (i.e., inverse decay, Δ , and sensitivity to reinforcement, θ); (2) the relative strength of directed exploration processes (directed exploration bonus, φ , and directed exploration learning rate, α) and (3) the stochasticity of individuals' decisions (consistency, C). These five parameters were our primary outcomes of interest given they drive the explore/exploit behaviour described in the previous section.

The VSE model has been shown to have good to excellent parameter recovery. The creator of the model found a mean parameter recovery of $r = 0.81$, with all parameters falling within a recoverability range of $r = 0.67$ – 0.95 .¹³ A simulation based on our own sample size ($N = 149$) found a mean parameter recoverability of $\rho = 0.73$ (reinforcement sensitivity, θ , $\rho = 0.56$; inverse decay, Δ , $\rho = 0.83$; directed exploration bonus, φ , $\rho = 0.87$; directed exploration learning rate, α , $\rho = 0.58$; consistency, C , $\rho = 0.79$, $p < 0.001$ for all parameters). Table S1 reports the correlations matrix between the five parameters from our analysis. The small and non-significant correlations between exploitative-based and explorative-based parameters (i.e., reinforcement sensitivity, θ , and inverse decay, Δ , did not correlate with directed exploration bonus, φ , and/or directed exploration learning rate, α) highlight how exploitation and exploration are separable constructs in the model.

2.4.6 | Changes in methamphetamine use at six-weeks

The Timeline Follow Back interview (TLFB)⁵⁰ measured the number of days PwMUD used methamphetamine in the month before commencing treatment and in the month before the follow-up assessment. To aid accuracy, we provided several cues to aid recollection (i.e., including public holidays on the calendar; prompting participants to describe their patterns of use and identifying any deviations; and assuring confidentiality). We calculated reduction of methamphetamine use by subtracting follow-up TLFB scores from baseline TLFB scores.

2.5 | Statistical analyses

2.5.1 | Overview

When comparing groups, we applied Bayesian and frequentist independent t tests, Mann-Whitney U tests, and Chi-Square tests using JASP.⁵¹ For the traditional analysis of net score across 20-trial blocks, a 5 (number of blocks) \times 2 (groups) repeated ANOVA was used. A Bayes Factor (BF_{10}) of $<1/30$ indicated evidence in favour of the null, and >3 indicated evidence in favour of the alternative.⁵² Alpha was

set at $p = 0.05$. We implemented the VSE model using the VSE Toolbox¹³ in MATLAB 2017a.⁵³ Outliers were removed based on choice behaviour data and were defined as more than 3.29 standard deviations from the mean.⁵⁴

2.5.2 | Aim 1: Modelling and comparing explore/exploit mechanisms between groups

The VSE toolbox identified whether PwMUD and controls differed in their parameters of explore/exploit decisions. This toolbox uses a Variational Bayesian scheme, implemented via the VBA Toolbox.⁵⁵ To first see whether the VSE model was the best fitting model of participants' behaviour, we compared model fit between the VSE and other models of the IGT. These comparisons used the Akaike Information Criterion⁵⁶ and included (1) the VSE model; (2) a modified VSE model which includes separate parameters for sensitivity towards reward and losses; (3) the Expectancy Valence model²⁹; (4) the Prospect Valence Learning model³⁰; (5) the Value Plus Perseverance model³¹ and (6) the Outcome-Representation Learning model.³² For brief explanations of the other models, see the supporting information. We based each individuals' parameters on the best fitting model.

2.5.3 | Covarying for cognitive or demographic group differences

Because PwMUD can differ from controls in areas such as attention,^{24,57} working memory,⁴⁹ depression²⁴ and IQ,²⁴ we also investigated whether any of these four factors were explaining any of the computational, between-group findings. This is particularly important, as previous work with this sample of PwMUD found increased inattention and depression, in addition to decreased full-scale IQ, compared to controls.²⁴ Therefore, we obtained previously collected measures of inattention (omission errors on the Continuous Performance Test II; CPT-II Om Errors),⁵⁸ inattention/disinhibition (commission errors on the CPT-II; CPT-II Comm Errors)⁵⁸ and working memory (longest digit sequencing span from the Letter Number Sequencing task; LDSS)³⁶ used previously in this sample. We then ran correlations between the above five measures (intelligence, depression, inattention, inattention/disinhibition and working memory) and the VSE parameters of interest, controlling for multiple comparisons. If significant, they were included as covariates in an ANCOVA analysis. A brief description of the additional cognitive measures and their descriptive statistics are presented in the supporting information.

2.5.4 | Aim 2: Predicting reductions in methamphetamine use

We used multiple hierarchical regression to identify whether any VSE parameters could predict changes in PwMUD's methamphetamine use at 6-week follow-up. We chose the VSE parameters to predict

reductions in methamphetamine use, rather than summary behavioural statistics, as they provide a more mechanistic explanation of why methamphetamine use may change, over and above the raw behavioural outputs alone. Here, we used the stepAIC function⁵⁹ in R⁶⁰ to identify the best combination of parameters. We applied the stepwise regression both backwards (i.e., starting with a model including all parameters and then systematically removing parameters) and forwards (i.e., starting with the simplest model that includes only the intercept and then systematically adding parameters). The best fitting model was then recreated in JASP to determine Bayesian estimates.

3 | RESULTS

3.1 | Behavioural analyses

Five PwMUD represented extreme outliers in one or more behavioural domains (i.e., net score, win/stay, lose/switch, mutual information, choice entropy and sequential exploration over three/four trials) and were removed from analyses.

Visual inspection of Figure 1 shows control participants continuously improved their net score across the task. In contrast, PwMUD did not increase their score. Furthermore, controls showed a higher level of sequential exploration for three and four consecutive decks in the first 10 trials, before reducing their exploration (Figure 2). In contrast, PwMUD started with lower levels of exploration yet maintained these at a higher level throughout the task. Thus, controls appeared to experience an early period of exploration, followed by adaptive exploitation, while PwMUD appeared to struggle identifying which decks should have been exploited and instead appeared to engage in random exploration.

Traditional group comparisons yielded significant differences on net score, with controls performing better than PwMUD, $t(70.83) = 4.50$, $p < 0.001$, $BF_{10} = 2265.34$, as well as controls showing a greater improvement in their net score across the blocks of the task, $F(3.32, 500.74) = 6.07$, $p < 0.001$, $BF_{incl} = 100.95$. Using more detailed behavioural analysis, controls also showed greater mutual information, an indication towards autocorrelation between subsequent choices, $U = 3486$, $p < 0.001$, $BF_{10} = 89.22$. Conversely,

PwMUD showed greater choice entropy, indicating greater selection of different decks across the task, $U = 1488$, $p < 0.001$, $BF_{10} = 60.43$. Consistent with Figure 2, PwMUD exhibited greater sequential exploration across three consecutive trials, $U = 1729.5$, $p = 0.005$, $BF_{10} = 7.51$, though there was mixed evidence to whether this trend continued over four trials, $U = 1925$, $p = 0.043$, $BF_{10} = 1.75$. There were no significant group differences on win/stay ($U = 2726.5$, $p = 0.22$, $BF_{10} = 0.45$) or lose/switch behaviour ($U = 2054$, $p = 0.13$, $BF_{10} = 0.81$). Figure 3 visualises these results.

Log median reaction times did not differ between PwMUD ($M = 6.45$, $SD = 0.62$) and controls ($M = 6.38$, $SD = 0.45$; $U = 2369$, $p = 0.83$, $BF_{10} = 0.20$).

3.2 | Aim 1: Modelling and comparing explore/exploit mechanisms between groups

Compared to all other models, the VSE had the best fit in both groups (see Figures S1 and S2). From this, we found that PwMUD's behaviour on the IGT appeared primarily driven by deficits in exploitative processes.

First, we investigated whether any of the five covariates previously identified in Section 2.5.2 were particularly relevant to explaining between-group differences in the computational parameters of the VSE models (see Table S2). Here, inattention/disinhibition (CPT-II Comm Errors) correlated with directed exploration bonus (ρ ; $\rho = 0.30$, $p = 0.002$), and intelligence (WASI-II) correlated with inverse decay (Δ , $\rho = 0.26$, $p = 0.003$). Thus, both were included as covariates in their respective between-group analysis.

In comparing the computational parameters, PwMUD were first found to be less sensitive to reinforcement (both negative/positive) than controls, shown by a reduced reinforcement sensitivity (θ) parameter, $U = 3065$, $p = 0.009$, $BF_{10} = 3.22$. Second, there was a trend for PwMUD to use fewer previous choices to guide their current choice, as shown by a borderline difference in the inverse decay (Δ) parameter, $U = 3062$, $p = 0.010$, $BF_{10} = 2.48$. When including the covariate of intelligence ($F[1, 145] = 5.98$, $p = 0.02$, $BF_{incl} = 4.44$, $\eta^2 p = 0.04$), this between-group difference on inverse decay (Δ)

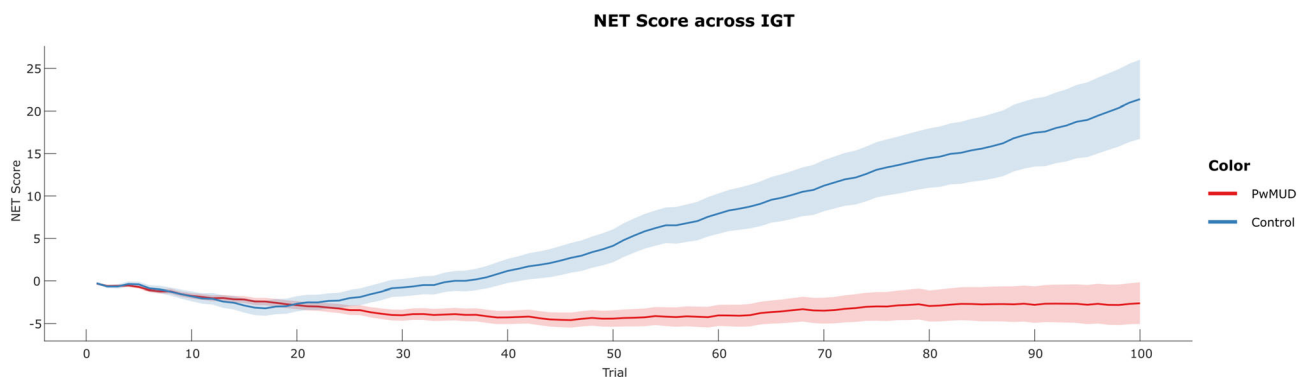
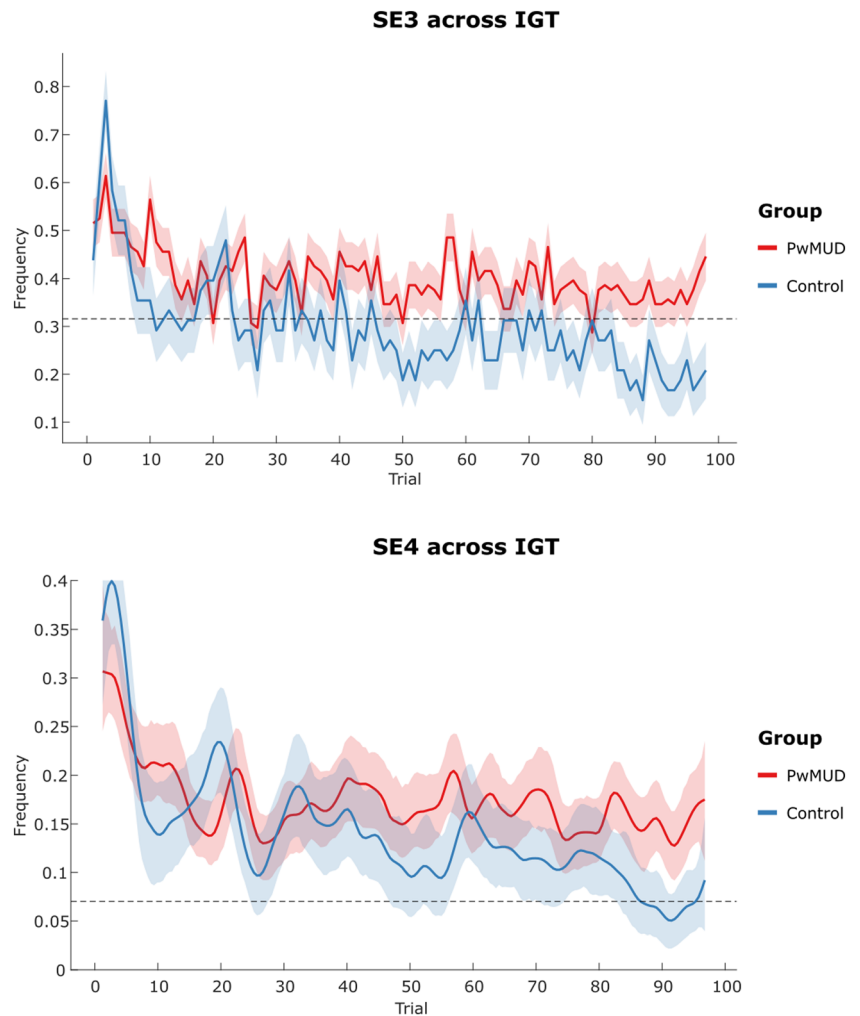


FIGURE 1 Cumulative net score of PwMUD (red) and controls (blue) across IGT. Note. Solid line represents mean; shaded area represents 95% confidence intervals

FIGURE 2 Comparison of behavioural indices of sequential exploration (SE3 and SE4) between PwMUD (red) and controls (blue). *Note.* SE3, sequential exploration 3. SE4, sequential exploration 4. These refer to how often participants chose three/four different decks across three/four subsequent trials (i.e., choosing deck A then C then B then D). Dotted line represents theoretical chance of each event (0.33 for SE3, and 0.09 for SE4). Solid line represents mean; shaded area represents 95% confidence intervals



remained similar ($F[1, 145] = 4.33$, $p = 0.04$, $BF_{incl} = 1.96$, $\eta^2 p = 0.03$). Third, PwMUD and controls did not significantly differ in the upper bound of their directed exploration value, shown by the directed exploration bonus parameter (φ ; $U = 2046$, $p = 0.13$, $BF_{10} = 0.39$). Moreover, when we included inattention as a covariate ($F[1, 146] = 10.47$, $p = 0.002$, $BF_{incl} = 31.52$, $\eta^2 p = 0.07$), the between-group differences on directed exploration bonus (φ) became null ($F[1, 146] = 0.27$, $p = 0.61$, $BF_{incl} = 0.22$, $\eta^2 p = 0.002$). Fourth, and in contrast to the previous result, PwMUD reached their upper bound of directed exploration slower than controls, shown by the directed exploration learning rate parameter (α), $U = 3094$, $p = 0.007$, $BF_{10} = 5.07$. Finally, choice stochasticity was similar between groups, shown by the consistency parameter (C), $U = 2525$, $p = 0.68$, $BF_{10} = 0.21$. Figure 4 visualises these group differences across the different parameters.

3.3 | Aim 2: Predicting reductions in methamphetamine use

The best fitting combination of VSE parameters (Table 3) significantly predicted changes in self-reported methamphetamine use between

baseline and 6-week follow-up, $F(3,69) = 5.93$, $p = 0.001$, adj. $R^2 = 0.17$. This model retained three parameters: inverse decay (Δ), directed exploration bonus (φ) and consistency (C), all of which significantly predicted changes in methamphetamine use at follow-up. Here, inverse decay (Δ) and consistency (C) negatively associated with reduced methamphetamine use (i.e., using fewer previous choices to guide decisions and having a greater choice stochasticity were associated with smaller reductions in methamphetamine use). Conversely, directed exploration bonus (φ) positively associated with reduced methamphetamine use (i.e., a greater tendency towards directed exploration was associated with larger reductions in methamphetamine use). These results were not explained by task engagement, with no correlation between log median reaction time and methamphetamine use ($r = .02$, 95%CI $[-0.21, 0.25]$, $p = 0.87$, $BF_{10} = 0.15$).

We then conducted a post hoc hierarchical regression to see if our best fitting model of VSE predictors (directed exploration bonus, φ ; consistency, C , and inverse decay, Δ) had stronger predictive utility than a previous approach in this sample that used working memory to predict future methamphetamine use.⁴¹ The first block of this regression that retained the original study's variables (age, sex, intelligence, sociodemographic status, depression, severity of dependence to methamphetamine, severity of dependence to cannabis, working

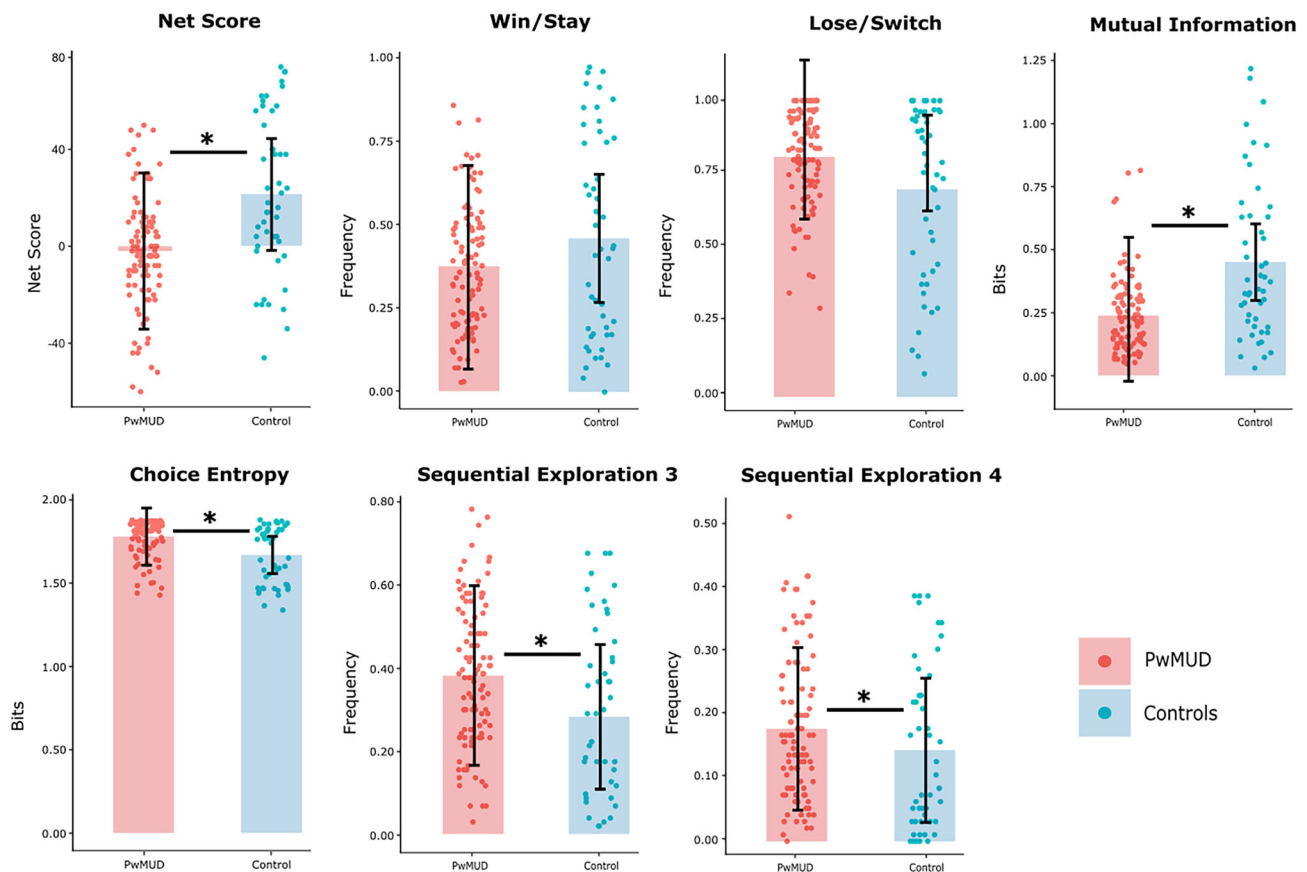


FIGURE 3 Behaviour indices of choice behaviour in PwMUD (red) and controls (blue). *Note.* Individual dots represent each participant in each group; * = $p < .05$. While sequential exploration 4 was significant, $BF_{10} = 1.75$. Higher values of mutual information highlight indicate that choices on subsequent trials are better predicted by the current choice. Higher values of choice entropy indicate that participants evenly explored all decks; lower values of choice entropy indicate selecting from only a few decks/one deck. Sequential exploration 3/4 refers to how often participants chose three/four different decks across three/four subsequent trials

memory and delay discounting k score) did not significantly predict reductions in TLFB scores, $F(9,61) = 1.51$, $p = 0.17$ and explained 6% of the variance. In comparison, adding in the three best VSE parameters in the second block made the model significant, $F(12,58) = 2.61$, $p = 0.008$ and explained 22% of the variance in PwMUD's reduction in TLFB scores.

4 | DISCUSSION

We revealed the mechanisms driving maladaptive explore/exploit decisions amongst PwMUD and identified how some of these mechanisms can predict future changes in methamphetamine use. At treatment onset, PwMUD exhibited limited exploitive computational processes, which seemed to produce a persistent form of 'random' behavioural exploration. This was evident by PwMUD having a reduced sensitivity to overall reinforcement (reward sensitivity, θ); a likely preference for using fewer, more recent trials to guide decision-making (inverse decay, Δ), and a behavioural pattern of consistently switching their choices across the task without improving their performance. In addition, those PwMUD who showed greater reductions of

methamphetamine use during early recovery demonstrated, at baseline, greater directed exploration (directed exploration bonus, φ), reduced choice stochasticity (consistency, C) and greater incorporation of previous outcomes in their decisions (inverse decay, Δ). These findings highlight that PwMUD's presentation to treatment is characterised by dysfunction in exploitive computational processes, while both exploration and exploitive computational processes impact treatment success. We provide a summary reminder of the VSE parameters and their key findings in Table 4 to aid the reader.

Currently, the prevailing view is that substance addiction represents a bias away from exploration and towards excessive exploitation.^{2,15,16} Our finding that PwMUD demonstrate weakened exploitive mechanisms challenges this perspective. In fact, previous behavioural,⁶¹ computational⁶² and neuroimaging studies⁶³ have shown that PwMUD have a reduced sensitivity to positive reinforcement and decreased dopaminergic functioning in key regions related to exploitive reinforcement learning (e.g., the striatum).⁶⁴ Furthermore, our finding of a likely decreased inverse decay amongst PwMUD is similar to another computational study of decision-making amongst PwMUD.⁶² Together, these lines of evidence, in conjunction with lack of between-group differences in the directed exploration

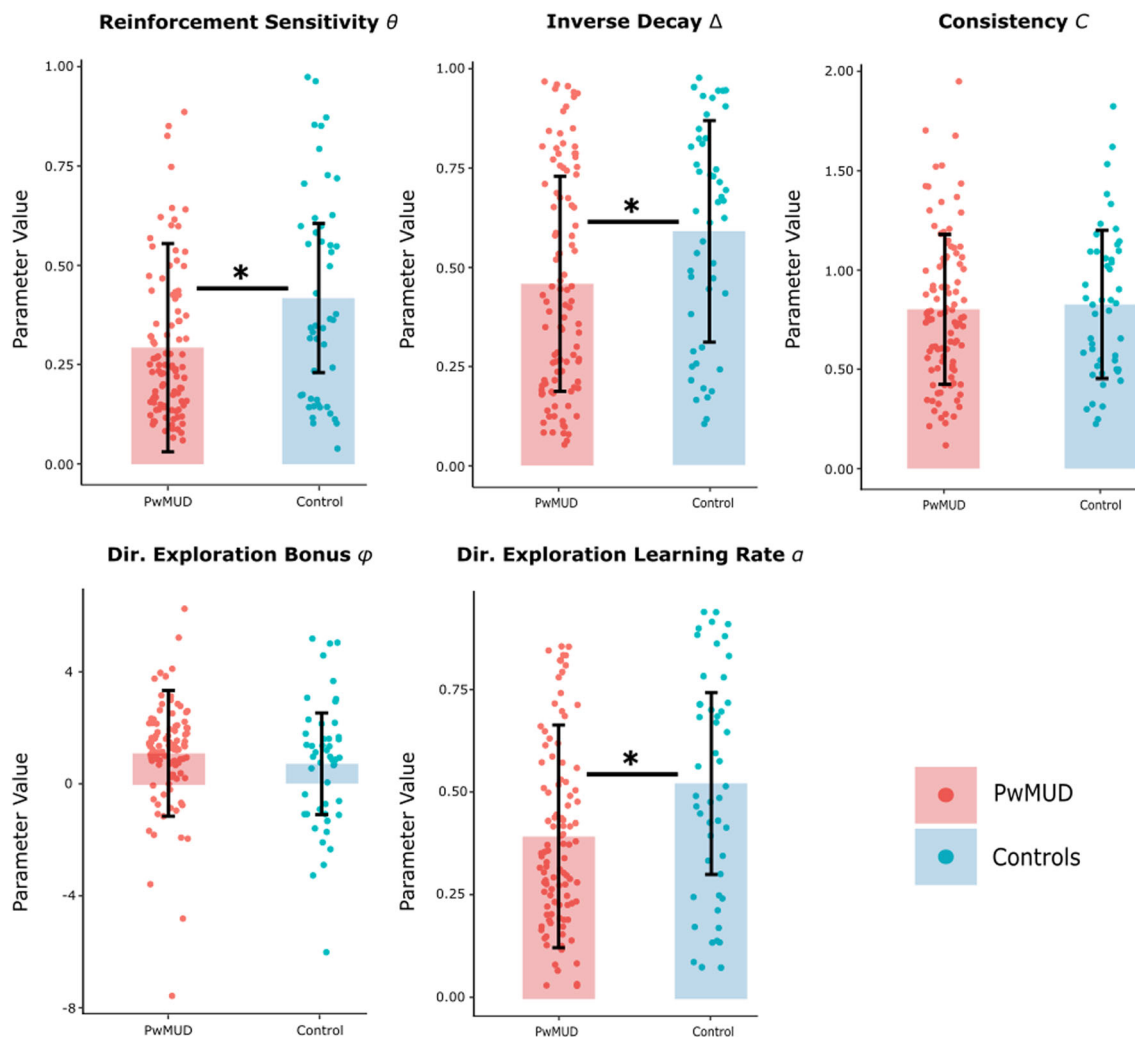


FIGURE 4 VSE parameter estimates of PwMUD (red) and controls (blue). Note. Individual dots represent each participant in each group; $*p < .05$. While inverse decay was significant, $BF_{10} = 2.48$. Higher values of reinforcement sensitivity indicate greater sensitivity for reward/punishment. Higher values of inverse decay indicate use of more previous outcomes. Higher values of consistency reflect less choice stochasticity. Higher values of exploration bonus indicate a greater maximum value of directed exploration. Higher values of exploration learning rate indicate a greater frequency of reaching the maximum directed exploration value

TABLE 3 Statistics of best fitting model using the VSE parameters to predict PwMUDs reduction of use between baseline and follow-up

Predictor	Unstandardised coefficient	Standard error	Standardised Coef. (β)	β 95% CI	t	p	BF_{inc}
Intercept	33.31	4.67			7.14	<0.001	1.00
Inverse decay, Δ	-16.12	5.05	-0.39	[-0.63, -0.14]	-3.19	0.002	16.29
Exploration bonus, φ	1.35	0.66	0.22	[0.04, 0.44]	2.04	0.045	4.36
Consistency, C	-11.94	3.71	-0.39	[-0.64, -0.15]	-3.22	0.002	14.94

bonus parameter (φ), indicate that PwMUD have fundamental and specific weaknesses in identifying and exploiting prior instances of reinforcement, which then lead to a behavioural presentation of 'random exploration'.

In contrast, PwMUD who showed the greatest reductions of methamphetamine use had greater directed exploration, showed greater consistency in their decisions and incorporated more past

outcomes into their choices. There are at least two non-mutually exclusive possibilities for this finding. First, PwMUD with greater directed exploration may be more open to engage with alternative, non-drug rewards. As such, these participants may have been more likely to identify rewarding aspects of non-drug related actions (i.e., social connectedness and personal meaning), which is protective towards recovery.^{65,66} Second, because these participants

TABLE 4 Summary of VSE parameters, their interpretation, meanings and findings in between-group and predictive analysis

Parameter name	Range	Interpretation	Between-group differences	Predicting changes in meth. use
Exploitative parameters				
Reinforcement sensitivity, θ	$\theta \in [0, 1]$	Influences the strength of rewards/losses on value calculations. Values near 0 reflect a weaker sensitivity to value of rewards/losses equally. Values near 1 reflect a greater sensitivity to value of rewards/losses equally.	PwMUD had less sensitivity to reinforcement than controls.	Was not included in the best fitting model.
Inverse decay, Δ	$\Delta \in [0, 1]$	Identifies how many previous trial outcomes the participant is using to guide their current decision. At 0, only the previous trials outcome guides the next decision. At 1, all the previous trials outcomes guide the next decision.	PwMUD likely used fewer previous trial outcomes to guide their current decision, compared to controls.	Was in the best fitting model. Using fewer previous choices were associated with a smaller reduction in methamphetamine use.
Explorative parameters				
Directed exploration bonus, φ	$\varphi \in [-\infty, \infty]$	Sets a maximal 'threshold' of how biased the participant is to select decks that have not been explored in recent trials. Negative values indicate an overall preference to keep selecting familiar decks (exploitation). Positive values indicate an overall preference to explore recently unselected decks (directed exploration).	A likely null difference between groups. Neither group had a higher threshold or bias towards directed exploration.	Was in the best fitting model. Having a greater bias towards directed exploration was associated with a greater reduction in methamphetamine use.
Directed exploration learning rate, α	$\alpha \in [0, 1]$	How quickly participants return to their maximal value of exploration again, after having recently explored. Values near 0 reflect a slow return to the maximal explore value. Values near 1 reflect a quick return to the maximal explore value.	PwMUD reached their upper value of directed exploration slower than controls. Please note, this does not necessarily mean their exploration value was lower than exploitation.	Was not included in the best fitting model.
Other parameters				
Consistency, C	$C \in [0, \infty]$	Reflects stochastic behaviour (i.e., whether the behaviour of the participant is consistent or not with the equations of the VSE model). Greater values reflect behaviour with greater consistency to the VSE model. Smaller values reflect more unpredictable behaviour to the VSE model.	A likely null difference between groups. This means both groups appeared to act equally consistently within the parameters of the VSE model.	Was in the best fitting model. Having greater choice stochasticity was associated with a smaller reduction in methamphetamine use.

incorporated more past choices into their decision-making, they may be more mindful about the outcomes of methamphetamine-related choices during recovery. For example, such PwMUD may have been able to incorporate greater amounts of evidence towards recovery (or against relapse) and thus have greater biases towards treatment goals.⁶⁶

Our study highlights how computational modelling of the explore/exploit framework can help predict the clinical prognosis of PwMUD. This is particularly important, as the search for neuropsychological predictors of PwMUD's treatment outcomes is still relatively new⁶⁷ and, as such, provides an additional data source for clinicians to identify those at risk of poor treatment outcomes.⁶⁸ Furthermore, the VSE parameters identified appear to show a significant improvement over previous modelling attempts with this dataset, when predicting changes of methamphetamine use over the early stages of recovery. Finally, we show how the VSE model and detailed behavioural outcomes can dissect critical decision-making mechanisms that characterise explore/exploit behaviour at treatment onset.

However, these findings must be considered in the context of the limitations of our study. For example, the original protocol, on which our data are based, lost ~25% of PwMUD to follow-up.³⁴ Even though this retention rate is satisfactory relative to other studies,⁶⁹ our results may be skewed towards those who can consistently engage with research programmes. Furthermore, our PwMUD group presented with higher scores of depressive symptomatology and inattention/disinhibition and lower scores of full-scale IQ, all of which may impact task performance. Still, we accounted for these potential covariates in our between-group analyses and found no change to our overall findings. Finally, it is important to note that performance on a rapid, small-scale and structured computer task may not necessarily generalise to slower, larger-scale and less structured decisions a person makes in their day-to-day lives during recovery. As such, replication of our findings in more ecologically relevant paradigms appears necessary.

To conclude, we identified specific dysfunction in explore/exploit mechanisms amongst PwMUD and then used these parameters to predict changes in methamphetamine use during early recovery. At treatment onset, PwMUD exhibited a behavioural pattern of random exploration, driven by deficits in identifying and exploiting the outcomes of non-drug actions. In comparison, we found that the greatest reductions in methamphetamine use were associated with greater directed exploration, more consistent choices and a greater use of previous outcomes. Our findings highlight the importance of identifying the underlying computational mechanisms of explore/exploit behaviour in addiction, while also providing a useful, methodological approach that can be applied across other addiction-related cohorts.

ACKNOWLEDGMENTS

We sincerely thank Adam Rubenis and Rebecca Fitzpatrick for their research efforts which provided the data for this paper. We also sincerely thank all those who participated in the original study. AHR was supported by an Australian Government Research Training Program Scholarship. TT-JC was supported by the Australian Research Council

(DP 180102383 and DE 180100389). AVG was supported by an Australian Medical Research Future Fund, Next Generation of Clinical Researchers, CDF-2 Fellowship (MRF1141214). Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

AUTHOR CONTRIBUTIONS

AHR and AVG designed the study. AHR ran the analyses and interpreted the results, with support from TT-JC. AHR led the manuscript writing process with TT-JC and AVG providing significant feedback and revisions. All authors critically reviewed the above content and approved this final version for publication.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available upon request from the corresponding author. It is not publicly available due to privacy and ethical restrictions.

ORCID

Alex H. Robinson  <https://orcid.org/0000-0002-8180-7716>

Trevor T.-J. Chong  <https://orcid.org/0000-0001-7764-3811>

Antonio Verdejo-Garcia  <https://orcid.org/0000-0001-8874-9339>

REFERENCES

1. Substance-related and addictive disorders. In: *Diagnostic and Statistical Manual of Mental Disorders Fifth Edition*. American Psychiatric Association; 2013. doi:10.1176/appi.books.9780890425596.dsm16
2. Addicott MA, Pearson JM, Sweitzer MM, Barack DL, Platt ML. A primer on foraging and the explore/exploit trade-off for psychiatry research. *Neuropsychopharmacology*. 2017;42(10):1931-1939. doi:10.1038/npp.2017.108
3. Axelsson M, Snell TL. Breaking the ice: narratives of recovery from crystal methamphetamine. *Aust Psychol*. 2021;56(1):81-92. doi:10.1080/00050067.2021.1893600
4. Henry BL, Minassian A, Perry W. Effect of methamphetamine dependence on everyday functional ability. *Addict Behav*. 2010;35(6):593-598. doi:10.1016/j.addbeh.2010.01.013
5. Maxwell JC. A new survey of methamphetamine users in treatment: who they are, why they like "meth," and why they need additional services. *Subst Use Misuse*. 2014;49(6):639-644. doi:10.3109/10826084.2013.841244
6. Gonzales R, Ang A, Glik DC, Rawson RA, Lee S, Iguchi MY. Quality of life among treatment seeking methamphetamine-dependent individuals. *Am J Addict*. 2011;20(4):366-372. doi:10.1111/j.1521-0391.2011.00142.x
7. Gonzales R, Ang A, Marinelli-Casey P, Glik DC, Iguchi MY, Rawson RA. Health-related quality of life trajectories of methamphetamine-dependent individuals as a function of treatment completion and continued care over a 1-year period. *J Subst Abuse Treat*. 2009;37(4):353-361. doi:10.1016/j.jsat.2009.04.001
8. Grigg J, Manning V, Arunogiri S, et al. *Methamphetamine Treatment Guidelines: Practice Guidelines for Health Professionals (Second Edition)*. Turning Point; 2018. <http://s3-ap-southeast-2.amazonaws.com/turning-point-website-prod/drupal/2019-05/Turning-Point-Methamphetamine-Treatment-Guidelines.pdf>
9. Brecht M-L, Herbeck D. Time to relapse following treatment for methamphetamine use: A long-term perspective on patterns and predictors. *Drug Alcohol Depend*. 2014;139:18-25. doi:10.1016/j.drugaldep.2014.02.702

10. McKetin R, Kothe A, Baker AL, Lee NK, Ross J, Lubman DI. Predicting abstinence from methamphetamine use after residential rehabilitation: Findings from the Methamphetamine Treatment Evaluation Study. *Drug Alcohol Rev.* 2018;37(1):70-78. doi:10.1111/dar.12528
11. Hills TT, Todd PM, Lazer D, Redish AD, Couzin ID. Exploration versus exploitation in space, mind, and society. *Trends Cogn Sci.* 2015;19(1):46-54. doi:10.1016/j.tics.2014.10.004
12. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature.* 2006;441(7095):876-879. doi:10.1038/nature04766
13. Ligneul R. Sequential exploration in the Iowa Gambling Task: Validation of a new computational model in a large dataset of young and old healthy participants. *PLoS Comput Biol.* 2019;15(6):e1006989. doi:10.1371/journal.pcbi.1006989
14. Wilson RC, Geana A, White JM, Ludvig EA, Cohen JD. Humans use directed and random exploration to solve the explore-exploit dilemma. *J Exp Psychol Gen.* 2014;143(6):2074-2081. doi:10.1037/a0038199
15. Morris LS, Baek K, Kundu P, Harrison NA, Frank MJ, Voon V. Biases in the explore-exploit tradeoff in addictions: The role of avoidance of uncertainty. *Neuropsychopharmacology.* 2016;41(4):940-948. doi:10.1038/npp.2015.208
16. Addicott MA, Pearson JM, Wilson J, Platt ML, McClernon FJ. Smoking and the bandit: A preliminary study of smoker and non-smoker differences in exploratory behavior measured with a multi-armed bandit task. *Exp Clin Psychopharmacol.* 2013;21(1):66-73. doi:10.1037/a0030843
17. Myers CE, Sheynin J, Balsdon T, et al. Probabilistic reward- and punishment-based learning in opioid addiction: Experimental and computational data. *Behav Brain Res.* 2016;296:240-248. doi:10.1016/j.bbr.2015.09.018
18. Smith R, Schwartenbeck P, Stewart JL, Kuplicki R, Ekhtiari H, Paulus MP. Imprecise action selection in substance use disorder: evidence for active learning impairments when solving the explore-exploit dilemma. *Drug Alcohol Depend.* 2020;215:108208. doi:10.1016/j.drugalcdep.2020.108208
19. Gueguen MC, Schweitzer EM, Konova AB. Computational theory-driven studies of reinforcement learning and decision-making in addiction: what have we learned? *Curr Opin Behav Sci.* 2021;38:40-48. doi:10.1016/j.cobeha.2020.08.007
20. Bechara A, Dolan S, Denburg N, Hindes A, Anderson SW, Nathan PE. Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia.* 2001;39(4):376-389. doi:10.1016/S0028-3932(00)00136-6
21. Gonzalez R, Bechara A, Martin EM. Executive functions among individuals with methamphetamine or alcohol as drugs of choice: preliminary observations. *J Clin Exp Neuropsychol.* 2007;29(2):155-159. doi:10.1080/13803390600582446
22. van der Plas EAA, Crone EA, van den Wildenberg WPM, Tranel D, Bechara A. Executive control deficits in substance-dependent individuals: A comparison of alcohol, cocaine, and methamphetamine and of men and women. *J Clin Exp Neuropsychol.* 2009;31(6):706-719. doi:10.1080/13803390802484797
23. Duarte NA, Woods SP, Rooney A, Atkinson JH, Grant I. Working memory deficits affect risky decision-making in methamphetamine users with attention-deficit/hyperactivity disorder. *J Psychiatr Res.* 2012;46(4):492-499. doi:10.1016/j.jpsychires.2012.01.006
24. Fitzpatrick RE, Rubenis AJ, Lubman DI, Verdejo-Garcia A. Cognitive deficits in methamphetamine addiction: independent contributions of dependence and intelligence. *Drug Alcohol Depend.* 2020;209:107891. doi:10.1016/j.drugalcdep.2020.107891
25. Wang G, Shi J, Chen N, et al. Effects of length of abstinence on decision-making and craving in methamphetamine abusers. *PLoS ONE.* 2013;8(7):e68791. doi:10.1371/journal.pone.0068791
26. Lake MT, Shoptaw S, Ipser JC, et al. Decision-making by patients with methamphetamine use disorder receiving contingency management treatment: magnitude and frequency effects. *Front Psych.* 2020;11:22. doi:10.3389/fpsy.2020.00022
27. Zhu Y, Jiang H, Su H, et al. A newly designed mobile-based computerized cognitive addiction therapy app for the improvement of cognitive impairments and risk decision making in methamphetamine use disorder: randomized controlled trial. *JMIR Mhealth Uhealth.* 2018;6(6):e10292. doi:10.2196/10292
28. Ahn WY, Vasilev G, Lee SH, et al. Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. *Front Psychol.* 2014;5:849. doi:10.3389/fpsyg.2014.00849
29. Stout JC, Bussemeyer JR, Lin A, Grant SJ, Bonson KR. Cognitive modeling analysis of decision-making processes in cocaine abusers. *Psychon Bull Rev.* 2004;11(4):742-747. doi:10.3758/BF03196629
30. Ahn W-Y, Bussemeyer JR, Wagenmakers E-J, Stout JC. Comparison of decision learning models using the generalization criterion method. *Cognit Sci.* 2008;32(8):1376-1402. doi:10.1080/03640210802352992
31. Worthy DA, Pang B, Byrne KA. Decomposing the roles of perseveration and expected value representation in models of the Iowa Gambling Task. *Front Psychol.* 2013;4:640. doi:10.3389/fpsyg.2013.00640
32. Haines N, Vassileva J, Ahn W-Y. The Outcome-Representation Learning Model: a novel reinforcement learning model of the Iowa Gambling Task. *Cognit Sci.* 2018;42(8):2534-2561. doi:10.1111/cogs.12688
33. Zorick T, Nestor L, Miotto K, et al. Withdrawal symptoms in abstinent methamphetamine-dependent subjects. *Addiction.* 2010;105(10):1809-1818. doi:10.1111/j.1360-0443.2010.03066.x
34. Fitzpatrick RE, Robinson AH, Rubenis AJ, Lubman DI, Verdejo-Garcia A. Lack of longitudinal changes in cognition in individuals with methamphetamine use disorder during the first 6 weeks after commencing treatment. *Am J Drug Alcohol Abuse.* 2021;47(3):383-392. doi:10.1080/00952990.2020.1869243
35. First MB, Spitzer RL, Gibbon M, Williams JB. *User's Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders SCID-IV: Clinician Version.* American Psychiatric Pub; 1997.
36. Wechsler D. *Wechsler Abbreviated Scale of Intelligence-Second Edition (WASI-II).* NCS Pearson; 2011.
37. Topp L, Mattick RP. Choosing a cut-off on the Severity of Dependence Scale (SDS) for amphetamine users. *Addiction.* 1997;92(7):839-845. doi:10.1111/j.1360-0443.1997.tb02953.x
38. Lawrinson P, Copeland J, Gerber S, Gilmour S. Determining a cut-off on the Severity of Dependence Scale (SDS) for alcohol dependence. *Addict Behav.* 2007;32(7):1474-1479. doi:10.1016/j.addbeh.2006.09.005
39. van der Pol P, Liebrechts N, de Graaf R, Korff DJ, van den Brink W, van Laar M. Reliability and validity of the Severity of Dependence Scale for detecting cannabis dependence in frequent cannabis users. *Int J Methods Psychiatr Res.* 2013;22(2):138-143. doi:10.1002/mpr.1385
40. Rubenis AJ, Fitzpatrick RE, Lubman DI, Verdejo-Garcia A. Impulsivity predicts poorer improvement in quality of life during early treatment for people with methamphetamine dependence. *Addiction.* 2017;113(4):668-676. doi:10.1111/add.14058
41. Rubenis AJ, Fitzpatrick RE, Lubman DI, Verdejo-Garcia A. Working memory predicts methamphetamine hair concentration over the course of treatment: moderating effect of impulsivity and implications for dual-systems model. *Addict Biol.* 2017;24(1):145-153. doi:10.1111/adb.12575
42. Donovan DM, Bigelow GE, Brigham GS, et al. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. *Addiction.* 2012;107(4):694-708. doi:10.1111/j.1360-0443.2011.03473.x

43. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measur.* 1977;1(3): 385-401. doi:10.1177/014662167700100306
44. Australian Bureau of Statistics. Socio-Economic Indexes for Areas (SEIFA), Australia, 2011; 2011. <https://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2033.0.55.001~2011~Main%20Features~SEIFA%20Basics~3>
45. Verdejo-García AJ, López-Torrecillas F, de Arcos Aguilar F, Pérez-García M. Differential effects of MDMA, cocaine, and cannabis use severity on distinctive components of the executive functions in polysubstance users: a multiple regression analysis. *Addict Behav.* 2005;30(1):89-101. doi:10.1016/j.addbeh.2004.04.015
46. Gossop M, Darke S. The Severity of Dependence Scale (SDS): psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. *Addiction.* 1995;90(5): 607-614. doi:10.1046/j.1360-0443.1995.9056072.x
47. Verdejo-García A, Pérez-García M. Profile of executive deficits in cocaine and heroin polysubstance users: common and differential effects on separate executive components. *Psychopharmacology (Berl).* 2007;190(4):517-530. doi:10.1007/s00213-006-0632-8
48. Baayen RH, Milin P. Analyzing reaction times. *Int J of Psych Res.* 2010;3(2):12-28. doi:10.21500/20112084.80728
49. Seli P, Cheyne JA, Smilek D. Attention failures versus misplaced diligence: separating attention lapses from speed-accuracy trade-offs. *Conscious Cogn.* 2012;21(1):277-291. doi:10.1016/j.concog.2011.09.017
50. Sobell LC, Sobell MB. *Timeline Followback Users Guide: A Calendar Method for Assessing Alcohol and Drug Use.* Addiction Research Foundation; 1996.
51. JASP Team. JASP (Version 0.13.1) [Computer Software]; 2020. <https://jasp-stats.org/>
52. Keyesers C, Gazzola V, Wagenmakers E-J. Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nat Neurosci.* 2020;23(7):788-799. doi:10.1038/s41593-020-0660-4
53. MATLAB. Version 2017a. MathWorks; 2017.
54. Tabachnick BG, Fidell LS. *Using Multivariate Statistics: Pearson New International Edition.* Pearson Education Limited; 2013.
55. Daunizeau J, Vincent A, Rigoux L. VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput Biol.* 2014;10(1):e1003441. doi:10.1371/journal.pcbi.1003441
56. Wagenmakers E-J, Farrell S. AIC model selection using Akaike weights. *Psychon Bull Rev.* 2004;11(1):192-196. doi:10.3758/BF03206482
57. Potvin S, Pelletier J, Grot S, Hébert C, Barr AM, Lecomte T. Cognitive deficits in individuals with methamphetamine use disorder: a meta-analysis. *Addict Behav.* 2018;80:154-160. doi:10.1016/j.addbeh.2018.01.021
58. Strauss E, Sherman EM, Spreen O. *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* American Chemical Society; 2006.
59. Venables WN, Ripley BD. *Modern Applied Statistics with S.* Springer; 2002. <http://www.stats.ox.ac.uk/pub/MASS4/> 10.1007/978-0-387-21706-2
60. R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/>. 2020
61. Robinson AH, Perales JC, Volpe I, Chong TT-J, Verdejo-García A. Are methamphetamine users compulsive? Faulty reinforcement learning, not inflexibility, underlies decision making in people with methamphetamine use disorder. *Addict Biol.* 2021;26(4):e12999. doi:10.1111/adb.12999
62. Harlé KM, Zhang S, Schiff M, Mackey S, Paulus MP, Yu AJ. Altered statistical learning and decision-making in methamphetamine dependence: evidence from a two-armed bandit task. *Front Psychol.* 2015;6: 1910. doi:10.3389/fpsyg.2015.01910
63. Bischoff-Grethe A, Connolly CG, Jordan SJ, et al. Altered reward expectancy in individuals with recent methamphetamine dependence. *J Psychopharmacol.* 2017;31(1):17-30. doi:10.1177/0269881116668590
64. Chang L, Alicata D, Ernst T, Volkow N. Structural and metabolic brain changes in the striatum associated with methamphetamine abuse. *Addiction.* 2007;102:16-32. doi:10.1111/j.1360-0443.2006.01782.x
65. McKay JR. Making the hard work of recovery more attractive for those with substance use disorders. *Addiction.* 2017;112(5):751-757. doi:10.1111/add.13502
66. Field M, Heather N, Murphy JG, Stafford T, Tucker JA, Witkiewitz K. Recovery from addiction: behavioral economics and value-based decision making. *Psychol Addict Behav.* 2020;34(1):182-193. doi:10.1037/adb0000518
67. Paulus MP, Stewart JL. Neurobiology, clinical presentation, and treatment of methamphetamine use disorder: a review. *JAMA Psychiat.* 2020;77(9):959-966. doi:10.1001/jamapsychiatry.2020.0246
68. Symons M, Feeney GFX, Gallagher MR, Young RM, Connor JP. Predicting alcohol dependence treatment outcomes: a prospective comparative study of clinical psychologists versus 'trained' machine learning models. *Addiction.* 2020;115(11):2164-2175. doi:10.1111/add.15038
69. Lappan SN, Brown AW, Hendricks PS. Dropout rates of in-person psychosocial substance use disorder treatments: a systematic review and meta-analysis. *Addiction.* 2020;115(2):201-217. doi:10.1111/add.14793

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Robinson AH, Chong TT-J, Verdejo-García A. Computational models of exploration and exploitation characterise onset and efficacy of treatment in methamphetamine use disorder. *Addiction Biology.* 2022;27(3): e13172. doi:10.1111/adb.13172