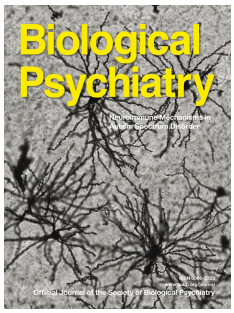


Changes in Model-Based and Model-Free Control Prospectively Predict Drinking Trajectories in Young Men

Hao Chen, Sören Kuitunen-Paul, Maria Garbusow, Marina Lukezic, Quentin J.M. Huys, Michael A. Rapp, Andreas Heinz, Michael N. Smolka

PII: S0006-3223(25)01302-2

DOI: https://doi.org/10.1016/j.biopsych.2025.06.028

Reference: BPS 15863

To appear in: *Biological Psychiatry*

Received Date: 26 August 2024

Revised Date: 20 June 2025

Accepted Date: 21 June 2025

Please cite this article as: Chen H., Kuitunen-Paul S., Garbusow M., Lukezic M., Huys Q.J.M., Rapp M.A., Heinz A. & Smolka M.N., Changes in Model-Based and Model-Free Control Prospectively Predict Drinking Trajectories in Young Men, *Biological Psychiatry* (2025), doi: https://doi.org/10.1016/j.biopsych.2025.06.028.

Manuscript for Biological Psychiatry July 4, 2025

Changes in Model-Based and Model-Free Control Prospectively Predict Drinking Trajectories in Young Men

Short Title: Model-based/Model-free Control Predict Future Drinking

Hao Chen[1], Sören Kuitunen-Paul[2,3,4], Maria Garbusow[5,6], Marina Lukezic[1], Quentin J.M. Huys[7], Michael A. Rapp[8], Andreas Heinz[4] & Michael N. Smolka[1]

1 Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Dresden, Germany

2 Chair of Clinical Child and Adolescent Psychology and Psychotherapy, Technische Universität Chemnitz, Chemnitz, Germany

3 Chair of Clinical Psychology and Psychotherapy, Technische Universität Chemnitz, Chemnitz, Germany

4 Department of Child and Adolescent Psychiatry, Technische Universität Dresden, Dresden, Germany

5 Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Campus Charité Mitte, Berlin, Germany

6 Department of Psychology, MSB Medical School Berlin DE

7 Applied Computational Psychiatry Lab, Mental Health Neuroscience Department, Division of Psychiatry and Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Queen Square Institute of Neurology, University College London

8 Area of Excellence Cognitive Sciences, University of Potsdam, Potsdam, Germany

Corresponding author:

Michael N. Smolka, MD

Section of Systems Neuroscience

Department of Psychiatry and Psychotherapy

Technische Universität Dresden

Würzburger Str. 35

01187 Dresden, Germany

Phone: +49 351 463 42201

Fax: +49 351 463 42202

E-Mail: michael.smolka@tu-dresden.de

Keywords: Model-based control, Model-free control, Risky alcohol use, Binge drinking, Reward prediction error, Alcohol drinking trajectory

Abstract

Background: Expanding our previous findings that model-based/model-free (MB/MF) control—often conceptualized as goal-directed and habitual behavior—at age 18 is associated with alcohol drinking trajectories over three years, this study investigates whether changes in MB/MF control from ages 18 to 21 *i)* stem from alcohol exposure and *ii)* predict drinking patterns up to age 24.

Methods: We followed a community sample of 124 18-year-old young men for six years. At ages 18 and 21, participants performed a two-step task assessing MB and MF control while undergoing functional magnetic resonance imaging (91 neural datasets). Drinking behavior was assessed using annual interviews complemented by questionnaires every six months. Correlation coefficients assessed the effect of cumulative alcohol exposure from age 18 to 21 on changes in MB/MF parameters. Latent growth curve models evaluated associations between MB/MF changes and drinking trajectories from ages 21 to 24.

Results: Alcohol exposure from ages 18 to 21 showed no significant effect on changes of MB/MF control. An increased MB behavioral score was protective for binge drinking, while an increased MF behavioral score predicted higher binge drinking at age 21, but not its future development. Changes in MF ventral striatum signals were associated with escalated consumption score development from ages 21 to 24, whereas MF ventromedial prefrontal signals exhibited a protective effect.

Conclusions: Preceding changes in behavioral and neural MB and MF control were linked to future drinking patterns, suggesting that interventions aimed at modulating MB/ MF controls could help mitigate subsequent risky drinking behaviors.

Introduction (685 words)

Alcohol Use Disorder (AUD) poses significant health risks and societal challenges, making understanding its underlying mechanisms a public health priority (1). The progression of AUD is marked by the transition from initially controlled, primarily goal-directed alcohol use to more habitual consumption, with some researchers suggesting this may involve elements of automaticity (2, 3). This transition underscores the necessity of dissecting the intertwined causes and consequences of AUD to develop preventions or interventions that directly target the underlying processes.

The two-step task, introduced by Daw, Gershman (4), is a well-established tool for exploring the interplay between goal-directed and habitual behaviors. Grounded in reinforcement learning, it distinguishes between model-based (MB) and model-free (MF) control systems. The MF system calculates the value of actions based on past rewards, with reward prediction error (RPE) signal predominantly originating in the midbrain (5). While these phasic signals align with MF predictions, evidence suggests that midbrain dopamine neurons also contribute to associative learning and outcome-specific predictive learning, rather than being strictly MF (6, 7). This allows task structure and future outcomes to influence reinforcement learning, integrating elements of MB computations alongside MF learning (4, 8). In contrast, the MB system depends on interactions between ventral striatum (VS) dopamine and lateral prefrontal cortex activation (9), engages in forward-planning decision-making and is sensitive to environmental structure (10). While the MF system is efficient in stable environments, its inflexibility becomes apparent in more complex settings. The MB system, though more adaptive

and future-oriented, incurs higher computational costs, potentially leading to inefficiency. Typically, human decision-making in the two-step task reflects a blend of both strategies, balancing efficiency with flexibility (4).

Some evidence suggests an association between reduced MB control and AUD severity. No significant behavioral differences were reported between persons with AUD and healthy volunteers by Voon, Derbyshire (11) and Sebold, Nebe (12), the latter in contrast to an initially significant finding (13). However, (12) observed that individuals who relapsed exhibited lower MB neural responses in the medial prefrontal cortex. In non-clinical populations, Doñamayor, Strelchuk (14) observed reduced MB control among adults with severe binge drinking. In a large online study, Gillan, Kosinski (15) found a link between lower MB control and higher scores on Alcohol Use Disorder Identification Test (AUDIT) (16), Conversely, Patzelt, Kool (17) did not detect this association using an online study with a modified version of the two-step task (18).

The mixed evidence underscores the intricate and multifaceted nature of the relationship between MB/MF control and AUD. The varying findings could be, at least partly, attributed to the complex interplay between inherent predispositions and the consequences of alcohol use, making it challenging to delineate clear cause-and-effect patterns. In this context, we followed young adults aged 18 for 6 years until age 24, which is a critical period when risky alcohol use and distinctive drinking patterns develop (19, 20). At age 18, no association was detected between MB/MF control and drinking behaviors (21). However, our findings indicate that more MB behavioral control at age 18 was associated with a reduction in the development of binge drinking over the following three years. Conversely, more MF RPE in the VS and ventromedial prefrontal cortex (vmPFC) were associated with an increase in the development of higher consumption

scores. These findings support the role of MB and MF control as predisposing factors (22). We now obtained data on (1) MB/MF decision making at age 21 and (2) annual/biannual alcohol use from age 21 to 24. We want to assess whether changes in the balance between MB/MF controls at ages 18 to 21 are associated with a) alcohol use during this time period, and b) with the future development of risky drinking behaviors from ages 21 to 24. This will allow us to examine the temporal direction of associations between changes in MB/MF control and the development of drinking behavior. Following our previous finding that MB behavioral control and MF neural responses at age 18 predict future risky drinking development (22), we now expect changes in MB/MF parameters to be associated with both the development of future drinking trajectories and cumulative drinking up to the age 24.

Materials and Methods

Participants & Procedure (282 words)

At baseline, 201 eighteen-year-old men recruited from local registration offices in Berlin and Dresden took part in our study. Participants needed to have normal or corrected-to-normal vision, be right-handed and eligible for MRI, and have had at least two drinking occasions during the last three months. Individuals with a history or current diagnosis of mental disorder or substance dependence (excluding nicotine), as assessed through structured clinical interviews based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; 23), were excluded, while those who met criteria for alcohol abuse were still included.

We recruited only males because we believed the higher prevalence of hazardous drinking in males compared to females would increase the statistical power to detect associations in our longitudinal study. Additionally, sex differences in MB and MF control (later found in (15)) could introduce interactions that reduce statistical power, also making it more difficult to detect associations. In retrospect, a sample of both females and males would have been preferable.

Participants performed the two-step sequential decision-making task (4) during functional magnetic resonance imaging (fMRI) at ages 18 and 21. Following quality control, 188 behavioral and 146 imaging datasets from baseline (age 18) were included in the final analysis (21, 22). At age 21, 124 behavioural datasets remained after exclusions. Imaging data were preprocessed identically to baseline, with 91 participants included in the final longitudinal analysis (see S-1 in the supplementary material).

Drinking behavior was assessed annually (ages 18–24) using the Munich Composite International Diagnostic Interview (M-CIDI; 24) and Alcohol Use Disorder Identification Test (AUDIT; 16) questionnaire biannually starting at age 18.5 years. More details about drinking behavior assessments are in Supplementary Material S-2.

Drinking Behaviour (95 words)

To assess cumulative alcohol consumption, two key variables were considered: total alcohol consumption (in kg) and the total number of binge drinking occasions from ages 18 to 21 (25), both derived from M-CIDI assessments. The details of this calculation can be found in the S-2 in the supplementary material.

Regarding the analysis of how the changes in MB/MF control predispose future drinking trajectories, we used the gram/occasion variable (binge drinking score) and the AUDIT-C score from ages 21 to 24. These two variables were selected to maintain consistency with our previous study (22).

Two-step Task (13 words)

Details of the two-step task (4) are described in Figure 1.

Figure 1: Two-step Paradigm

---------- Figure 1 about here ----------

Figure 1: Schematic of the Two-Step Decision-Making Task. In the depicted two-step paradigm (4), participants begin each trial by choosing one of two gray boxes within a two-second limit. For example, selecting the left box leads to a 'common' transition to a green pair of stimuli with a 70% probability, or a 'rare' transition to a yellow pair with a 30% likelihood. If the right box is chosen, these transition probabilities to the second-stage stimuli are reversed. Upon entering the second stage, participants were required to select one of the two second-stage stimuli within a two-second time frame. Below each second-stage stimulus are fluctuating reward probability charts, illustrating the chance (ranging from 25% to 75%) of earning a monetary reward throughout the task, according to a Gaussian random walk algorithm. Monetary rewards are given based on this probability, as depicted by the coin image at the bottom. Participants received €10 for each hour of their participation, in addition to a bonus determined by their performances on the two-step task. The payouts for this bonus ranged from €3.80 to €8, based on randomly selected one-third of the trials.

Behavioral parameters: As established in Daw, Gershman (4), the MF agent tends to repeat the first-stage choice following a reward, while the MB agent also considers transition structures and this results in a reward-by-transition

interaction. These scores, derived from the participant's first-stage decision across all trials, quantify the extent to which their behavior aligns with pure MF and MB agents. Specifically, MF score is formulated as the main effect of reward on the decision probability P: P(rewarded common) + P(rewarded rare) - P(unrewarded common) - P(unrewarded rare). Meanwhile, the MB score captures the interaction of reward and transition: P(rewarded common) - P(rewarded rare) - P(unrewarded common) + P(unrewarded rare).

Neural parameters: At the neural level, we analyzed the imaging data from age 21 using the same first-level model as outlined in our baseline report (21). Our primary regressors of interest in the fMRI model were the MF and MB RPEs. These RPEs, modeled as two parametric regressors, corresponded to the onset of the second-stage cue and the outcome presentation. They were computed with the same computational model detailed in Nebe, Kroemer (21). Our regions of interest were the bilateral vmPFC and the VS. From these regions, we extracted the MF and MB RPEs. These MB and MF RPEs obtained from the VS and vmPFC were then used to predict future drinking trajectories.

MB and MF control parameters (141 words)

To maintain consistency with our previous study (22), we employed the same behavioral and neural predictors in the current study. At the behavioral level, we calculated the MF and MB scores, based on whether participants repeated their first-stage choice in subsequent trials (4). Neural MB/MF RPEs were extracted from the VS and vmPFC at age 21 using the same first-level model as in our baseline report (21). See Figure 1 for detailed specifications of MB/MF parameters.

After extracting the two behavioral and four neural parameters, we assessed the stability of the six parameters across three years. Specifically, we reported the intraclass correlation (ICC) and Spearman's rho, as the Shapiro-Wilk tests confirmed non-normal distributions. This assessment provides insights into individual changes over time. Additionally, to test for changes in the overall mean, we performed Wilcoxon signed-rank tests.

Association between alcohol exposure and changes in MB/MF control (97 words)

The objective of this analysis was to assess the association between cumulative alcohol use and changes in MB/MF control. Given that both total alcohol consumption and total binge drinking occasions are not normally distributed (see Figure S1 in the supplementary material), we computed Spearman's correlation coefficients between the two drinking variables and the six two-step measures. Associations between alcohol consumption and two-step behavioural measures were assessed using data from 124 participants, while the associations between alcohol consumption and neural measures were analysed based on data from 91 participants.

LGCM Analysis: Changes in MB/MF control in association with future drinking trajectories (354 words)

The aim of this analysis was to examine whether changes in MB and MF control from ages 18 to 21 were associated with drinking trajectories over the subsequent three years (from ages 21 to 24), controlling for the values of two-step predictors at age 18. We have previously published findings on the association between baseline MB/MF control and the three-year drinking trajectory; here, we focus specifically on the impact of changes during follow-up. Controlling for baseline MB/MF control is essential to isolate the effect of these changes and avoid biases such as regression to the mean (26). This was achieved by fitting latent growth curve model (LGCM) using the lavaan package in R (27). The missing data can be handled using the full information maximum likelihood method, which has been demonstrated to be unbiased when the data are missing at random (28). The testing of the missing pattern in the drinking data

supported the assumption that the missings are at random (details in the supplementary material S-3).

We first confirmed that the development of drinking trajectories from ages 21 to 24 followed a linear rather than quadratic pattern, and therefore constructed the LGCM models for gram/occasion and AUDIT-C score based on the conceptual model (details in Figure S2). We included both baseline and change scores two-step predictors in our models, regressing them against the latent intercepts and slopes. Unlike in our previous work (22), we separated the behavioural and neural models to enhance the robustness of our analysis by retaining more observations, given that only 91 participants had complete data for both types of assessment. This resulted in four models: two behavioural (MB and MF scores as predictors) and two neural (RPE signals in the VS and vmPFC for MB and MF control), each considering the trajectories for both gram/occasion and AUDIT-C score. Additional analyses examining MB/MF indices at age 21 and their associations with drinking trajectories are in Supplementary Material S-4 and Table S2. We also investigated potential associations and interactions between alcohol expectancy scores and MB/MF control (Supplementary Material S-5).

Results

Drinking Behaviour (97 words)

From ages 18 to 24, participants consumed an average of 57g of alcohol per occasion, with six drinking occasions per month. Binge drinking occurred approximately 12 times per year, and total alcohol consumption was 4.4 kg per year (SD=4.7). The mean AUDIT-C score remained stable

at around 4.3 (SD=2.0) over six years. Compared to the general German population (29), our sample showed higher at-risk drinking behaviors. Further details are provided in supplementary material S-2.

We also plotted and described the drinking trajectories of gram/occasion and AUDIT-C from ages 21 to 24 in Figure 2.

Figure 2: Drinking Trajectories

---------- Figure 2 about here ----------

Figure 2: Integrated Observed and Predicted Individual Drinking Trajectories Using Unconditional Latent Growth Curve Models (LGCM). The graph presents individual growth trajectories for key measures of alcohol consumption: alcohol use disorder identification test consumption (AUDIT-C) scores (left panel) and grams of alcohol per occasion (right panel), across the ages of 21 to 24. The linear trajectories were modeled using unconditional LGCM that allow for the estimation of initial status (intercept) and change over time (slope) for each individual's drinking behavior. By fitting the LGCM without external predictors, the models provide a 'pure' view of each participant's developmental pattern, based on the observed data across the specified time points. Each colored line depicts an individual's predicted trajectory. Colored dots represent the actual observed data, while the open circles indicate the model's predicted values for each time point. The thick green line represents the mean of the model's predicted values over time. The blue dashed line represents the threshold for risky drinking, which is 60g for the gram/occasion variable (30) and 4 for the AUDIT-C score (31). At the group level (see thick green line in both plots), the mean gram/occasion exhibited a slight increase while the AUDIT-C score remained stable. The individual trajectories illustrated here exhibited a combination of increases and decreases.

Development of MB/MF control from ages 18 to 21 (187 words)

The descriptive statistics for MB and MF behavioral scores and neural responses in the VS and vmPFC are presented in Table 1. At the group level, no significant changes in MB and MF control or their neural underpinnings were found from ages 18 to 21 (Wilcoxon signed-rank test: all $p \geq 0.138$; all $r \leq 0.17$). However, this stability at the group level does not preclude changes at the level of individuals, as visualized in Figure S4 in the supplementary material.

To assess the temporal stability, Spearman's correlation and ICC were calculated over three years. The MB score exhibited moderate temporal stability (rho=0.46, p<.001; ICC[3,1]=0.47, 95% CI: [-0.32, 0.59]), while stability of the MF score was minimal (rho=-0.01, p=.946; ICC[3,1]=0.01, 95% CI: [-0.16, 0.19]). Neural MF responses in the VS and vmPFC exhibited modest stability, while the MB signals in these regions exhibited relatively lower stability. Further details are presented in Table 1. These findings indicate that there were changes in these predictors over time, which is to be expected given that there are three years between the initial and final assessments.

260 **Table 1: Descriptive Statistics for the Two-step Measures**

261

		Age 18				***Age 21***				***Age 18 vs. Age 21***			
	MB/MF	Mean	Median	Range	SD	Mean	Median	Range	SD	Change (Wilcoxon rank sum test)		Temporal Stability	
										W (p)	Effect Size r	Spearman's rho (p)	ICC* [3,1][95% CI]
Behavioral (N=124)	MB	0.29	0.23	-0.34 - 1.21	0.33	0.30	0.29	-0.22 - 1.13	0.29	7242 (.430)	0.08	**0.46 (<.001)**	0.47 [-0.32, 0.59]
	MF	0.09	0.09	-0.38 - 0.63	0.18	0.09	0.08	-0.55 - 1.15	0.22	7865 (.755)	0.03	-0.01 (.946)	0.01 [-0.16, 0.19]
Neural RPE Signals (N=91)	MB VS	0.34	0.48	-2.28 - 2.48	0.84	0.33	0.30	-3.23 - 3.42	0.99	4140 (1.00)	0.01	0.08 (.434)	0.09 [-0.12, 0.29]
	MB vmPFC	0.39	0.46	-4.13 - 4.12	1.14	0.11	0.23	-4.29 - 2.83	1.26	4668 (.138)	0.17	0.10 (.346)	0.17 [-0.03, 0.37]
	MF VS	0.28	0.22	-0.76 - 1.09	0.35	0.26	0.22	-0.74 - 1.41	0.36	4227 (.809)	0.08	**0.31 (.002)**	0.27 [0.01, 0.45]
	MF vmPFC	0.08	0.11	-1.29 - 1.08	0.43	0.05	0.05	-1.55 - 1.48	0.44	4429 (.418)	0.05	**0.24 (.024)**	0.18 [-0.03, 0.37]

Note: p-values smaller than 0.05 are marked in bold

**ICC are calculated as two-way mixed effects, consistency, single-measurement*

262

Association between alcohol exposure and changes in MB/MF control (110 words)

We examined the association between alcohol exposure, measured by total alcohol consumption and total number of binge drinking occasions from ages 18 to 21, and changes in two-step parameters over this period. No significant associations were found (all p≥.175; Table 2), indicating that alcohol exposure was not substantially associated with the changes in MB and MF control. Similarly, alcohol exposure showed no significant associations with MB/MF outcomes at age 21 (Table S4, supplementary material). To provide a comprehensive overview, we also examined the association between the cumulative AUDIT-C score and the MB/MF control changes, with results presented in supplementary material S-6.

Table 2: Associations between Alcohol Exposure and Development of MB/MF Control

Age 21 - Age 18	Total alcohol consumption (in kg)		Total number of binge drinking occasions	
	rho	P	rho	P
Δ MB Score	0.07	0.433	-0.02	0.856
Δ MF Score	0.10	0.279	0.11	0.213
Δ MB VS	0.02	0.841	-0.03	0.796
Δ MB vmPFC	-0.05	0.620	-0.06	0.591
Δ MF VS	-0.14	0.175	-0.04	0.727
Δ MF vmPFC	-0.13	0.205	-0.10	0.388

Changes in MB/MF control in association with future drinking trajectories (383 words)

Having established that alcohol exposure from ages 18 to 21 was not significantly associated with MB and MF control changes, we now examine whether these changes may predispose individuals to different drinking trajectories from ages 21 to 24. Using LGCM models, we assessed the association between the MB/MF control changes (ages 18-21) and subsequent

drinking trajectories (ages 21-24). Three of the four models demonstrated moderate to good model fit (binge drinking score with behavioural predictors, AUDIT-C models with behavioural and neural predictors; Table 3). The binge drinking score model with neural predictors showed poor fit and is reported in the supplementary material S-7.

For the binge drinking score measure (Figure 3A), which assesses the alcohol consumption per drinking occasion, we observed a negative association between MB score changes and the slope of the gram/occasion trajectory (Beta=-14.07, standard error [SE]=5.80, p=.015). This suggests that participants with stronger increases in MB behavioural scores exhibited a greater decrease in binge drinking score development. Conversely, MF score increase was associated with higher binge drinking scores at age 21, as evidenced by its positive association with the intercept (Beta=41.72, SE=15.82, p=.008).

The AUDIT-C trajectory, evaluating changes in drinking frequency and quantity, showed that higher MB behavioural score at age 18 to be associated with a higher AUDIT-C intercept (Beta=1.88, SE=0.70, p=.007), indicating that individuals with higher MB control at baseline tended to have higher AUDIT-C score at age 21 (Figure 3B). Conversely, changes in the behavioral score did not significantly predict the AUDIT-C trajectories.

In the AUDIT-C model with neural RPE signals (Figure 3C), we observed that a higher MB signal in the VS at age 18 was associated with lower AUDIT-C intercept (Beta=-1.38, SE=0.54, p =.010). This means that individuals with stronger MB signals in the VS at baseline tend to have lower AUDIT-C scores at age 21. Regarding the changes in the neural responses, we found that an increase in the MF RPE signal in the VS was associated with a more pronounced upward trend in the AUDIT-C trajectory from ages 21 to 24 (Beta=0.24, SE=0.12, p =.041). Conversely, changes in

301 the MF vmPFC signals showed an inverse effect, whereby an increase was associated with a more

302 pronounced decline in AUDIT-C development (Beta=-0.22, SE=0.09, p =.016).

303 **Table 3: LGCM Results**

Behavioral /Neural	MF/MB	Path	unstandardized Estimate	SE	Z	P	standardized estimate
gram/occasion							
Behavioral	MF Age 18	Behavioral score -> intercept	40.244	24.340	1.653	0.098	0.310
		Behavioral score -> slope	-9.004	9.858	-0.913	0.361	-0.171
	Δ MF	Δ Behavioral score -> intercept	41.718	15.817	2.638	**0.008**	0.498
		Δ Behavioral score -> slope	-6.274	6.480	-0.968	0.333	-0.184
	MB Age 18	Behavioral score -> intercept	5.641	13.046	0.432	0.665	0.079
		Behavioral score -> slope	3.570	5.777	0.618	0.537	0.122
	Δ MB	Δ Behavioral score -> intercept	19.928	13.798	1.444	0.149	0.269
		Δ Behavioral score -> slope	-14.067	5.801	-2.425	**0.015**	-0.468
	Model fit: χ2=23.85, df=11, p=.013, CFI=0.946, RMSEA=0.097, SRMR=0.054						
Neural	*Model fit: χ2=81.19, df=31, p<.001, CFI=0.894, RMSEA=0.133, SRMR=0.143*						
AUDIT-C							
Behavioral	MF Age 18	Behavioral score -> intercept	0.767	1.322	0.580	0.562	0.073
		Behavioral score -> slope	0.158	0.195	0.810	0.418	0.134
	Δ MF	Δ Behavioral score -> intercept	0.855	0.850	1.006	0.315	0.124
		Δ Behavioral score -> slope	0.075	0.125	0.600	0.548	0.096
	MB Age 18	Behavioral score -> intercept	1.878	0.702	2.676	**0.007**	0.318
		Behavioral score -> slope	0.066	0.104	0.633	0.527	0.100
	Δ MB	Δ Behavioral score -> intercept	0.075	0.125	0.600	0.548	0.096
		Δ Behavioral score -> slope	-0.130	0.106	-1.219	0.223	-0.193
	Model fit: χ2=67.09, df=40, p=.005, CFI=0.968, RMSEA=0.074, SRMR=0.080						
Neural	MF Age 18	VS -> intercept	-0.933	0.994	-0.939	0.348	-0.158
		vmPFC -> intercept	1.602	0.827	1.938	0.053	0.330
		VS -> slope	-0.039	0.134	-0.290	0.772	-0.060
		vmPFC -> slope	-0.035	0.111	-0.312	0.755	-0.065
	Δ MF	Δ VS -> intercept	-1.275	0.815	-1.563	0.118	-0.263
		Δ vmPFC -> intercept	1.007	0.641	1.570	0.116	0.273
		Δ VS -> slope	0.236	0.115	2.046	**0.041**	0.445
		Δ vmPFC -> slope	-0.219	0.091	-2.401	**0.016**	-0.543
	MB Age 18	VS -> intercept	-1.376	0.535	-2.571	**0.010**	-0.558
		vmPFC -> intercept	0.634	0.390	1.626	0.104	0.351
		VS -> slope	0.033	0.073	0.450	0.652	0.122
		vmPFC -> slope	0.035	0.054	0.641	0.522	0.176
	Δ MB	Δ VS -> intercept	-0.297	0.307	-0.968	0.333	-0.179
		Δ vmPFC -> intercept	0.138	0.243	0.566	0.571	0.105
		Δ VS -> slope	0.024	0.042	0.568	0.570	0.130
		Δ vmPFC -> slope	0.009	0.034	0.272	0.785	0.065
	Model fit: χ2=130.19, df=70, p<.001, CFI=0.940, RMSEA=0.097, SRMR=0.117						

304 *Note: P-values smaller than 0.05 are marked in Bold*

Figure 3: LGCM model results

---------- Figure 3 about here ----------

Figure 3: Significant Pathways in Latent Growth Curve Model (LGCM) Results. This figure illustrates the results from the LGCMs: the binge drinking score model measured in grams of alcohol consumed per occasion (A) and alcohol use disorder identification test consumption (AUDIT-C) scores (B and C). Green paths indicate significant positive associations, whereas red paths indicate significant negative associations. In the grams/occasion behavioral model, a negative association was found between the change in MB score and the development of the binge drinking trajectory (slope) from ages 21 to 24. Conversely, a positive association was observed between the change in MF behavioral score and the binge drinking score at age 21 (intercept). In the AUDIT-C behavioral model, we found a positive association between the MB behavioral score at age 18 and the intercept. To maintain clarity, only significant path estimates are displayed for the AUDIT-C neural model; comprehensive details are provided in Table 3. The change in MF VS signal is positively associated with the rate of change (slope) in the AUDIT-C trajectory, while the change in the MF vmPFC signal is negatively associated with this rate of change. Additionally, the MB RPE signal in the VS is negatively associated with the drinking behavior at age 21 (intercept).

Exploratory Mediation and Moderation Analysis (122 words)

The observed results indicated that the changes in MF RPE signals in the VS and vmPFC have opposite roles when predicting the trajectory of AUDIT-C from ages 21 to 24. This divergence suggests a potentially intricate relationship between the RPE signals in these two regions concerning their influence on future drinking behaviours. To better understand this dynamic, we tested whether vmPFC RPE signals moderate or mediate the relationship between VS RPE signals and AUDIT-C trajectory. Moderation analysis examined whether vmPFC altered this relationship,

328 while mediation analysis tested whether vmPFC explained part of the effect of VS RPE on drinking
329 behavior. Results suggest competitive mediation rather than moderation. See Figure 4 and
330 supplementary material S-8 for more details.

Figure 4: Mediation Analysis Results

---------- Figure 4 about here ----------

Figure 4: Mediation Analysis Results. We utilized the R mediation package (32) for the mediation analysis. The direct effect of Δ model-free ventral striatum (MF VS) on alcohol use disorder identification test consumption score (AUDIT-C) drinking trajectory is significant (Estimate=0.114, p=.028), implying that changes in Δ MF VS are associated with an increase in the slope of the drinking trajectory. The mediation effect, as represented by the average causal mediation effect of Δ MF ventromedial prefrontal cortex (vmPFC) is significant (Estimate=-0.090, p=.014) and operates in the opposite direction to the direct effect, thereby indicating competitive mediation (33). This suggests that the influence of Δ MF VS on AUDIT-C scores is partly offset by the mediating role of Δ MF vmPFC. Additionally, the path from the independent variable to the mediator (Δ MF VS to Δ MF vmPFC) is significant (Estimate=0.895, p<.001). The total effect of Δ MF VS on the drinking trajectory is non-significant (Estimate= 0.024, p = .510), which is consistent with the competitive mediation, where the mediator's effect contrasts with the direct effect. This competitive dynamic suggests that while the MF VS RPE changes are associated with an increase in AUDIT-C scores, the vmPFC signal changes offset this effect, leading to a nuanced interplay between the neural correlates and the progression of alcohol use behavior. The bold p-values highlight the statistical significance of the relationships between the variables.

Discussion (976 words)

In our longitudinal study, we tracked a community sample of 18- to 24-year-old men for six years, and found that changes in MB and MF control during young adulthood are predisposing factors for subsequently observed drinking trajectories. Notably, an increase in MB behavioral control from ages 18 to 21 was found to be protective and associated with a stronger decrease of binge drinking scores over the subsequent three years. Furthermore, an increase in the MF RPE signal in the VS preceded an escalation in consumption scores. The influence of changes in the VS RPE signals on future drinking behavior was found to be competitively mediated by changes in the MF RPE signal in the vmPFC, indicating that the latter signal may serve as a protective factor against increasing drinking behavior. Conversely, our analysis does not support the hypothesis that moderate alcohol consumption during young adulthood alters MB and MF control. These findings suggest that MB/MF development may play a role in the progression of alcohol use in young adults, potentially informing the development of targeted early intervention strategies.

Overall, our findings align well with our previous research (22). Earlier, we found that high MB behavioral control at age 18 protects against binge drinking score development from age 18 to 21. The current study extends this understanding by showing that an additional increase in MB behavioral control during early adulthood is associated with stunted progression of binge drinking score development over the subsequent three years, i.e. after age 21. This finding emphasizes the protective role of MB behavioral control for the binge drinking trajectory. Further, we previously observed that the MF RPE signal in the VS at age 18 positively correlated with the development of consumption scores in the following three years, i.e. was a risk factor (22). We extend this by demonstrating that changes in this MF signal during early adulthood may be linked

371 to excessive alcohol use. Taken together, these findings provide additional evidence that not only
372 MB/MF control at one time point, but also their development may be associated with future
373 drinking trajectories. This indicates that MB/MF control and drinking trajectories are co-
374 developed in a dynamic manner. Importantly, these associations were identified after the initial
375 levels of MB/MF control at age 18 were included as predictors, allowing us to test whether
376 MB/MF control at age 18 is also associated with future drinking trajectories. In summary,
377 consistent with our hypothesis, these findings delineate MB behavioral control as protective and
378 MF processes as detrimental in shaping alcohol use trajectories.

379 However, not all findings align neatly. Upon initial examination, the negative association
380 between changes in the MF signal in the vmPFC and consumption score development did not
381 align with the hypothesis that increasing MF signals are a risk factor. This unexpected result
382 prompted the hypothesis that a moderation or mediation effect may be present. Exploratory
383 analyses indeed suggest that changes of the vmPFC RPE signal act as a competitive mediator (33).
384 We speculate that the vmPFC signal might be involved in action inhibition during the development
385 of addiction (34), counteracting the heightened MF RPE signals from the VS, thus providing a
386 protective mechanism against future risky drinking patterns. This idea aligns with the broader
387 literature, which suggests that the vmPFC is crucial for integrating various signals and guiding
388 decision-making based on the expected value of an action (35). Additionally, the competitive
389 dynamic between MF vmPFC and VS signals may indicate distinct roles in MF processing: VS
390 signals likely reflect habitual, reward-driven tendencies, whereas the vmPFC MF signal may
391 encode more nuanced feedback about the broader consequences of behavior, mitigating the
392 influences of heightened VS activity. The differential maturation of the VS and vmPFC during this

period may also underlie these contrasting roles, with the vmPFC's later development enhancing its ability to regulate behaviors (36). However, MB and MF RPE signals in the two-step task may not be entirely distinctive (37), suggesting that our measurements might reflect general RPE signals rather than distinct MF or MB RPE signals having a direct influence on choices during the task. Overall, these findings highlight the complex interaction between neural signal changes and future drinking behavior, emphasizing the significant role of the vmPFC in this dynamic.

Complementary to the predisposing effects observed, our study is the first to investigate whether alcohol consumption alters MB/MF control in humans. Overall, We found no evidence that moderate levels of alcohol consumption (on average one standard drink per day; (38)) or binge drinking in young adults are associated with changes in MB and MF control. While research in this area is limited, Groman, Massi (39) did find both MB and MF control were reduced in rats following self-administered methamphetamine use. Our findings do not rule out the possibility that alcohol consumption may alter MB/MF control; rather, the lack of observed changes may be attributable to moderate alcohol use in our study population during early adulthood. Future research is required to determine whether higher levels of alcohol consumption and/or longer durations of alcohol exposure impact MB/MF control over time.

Our findings on the predisposing side underscore the importance of the development of decision-making mechanisms during early adulthood, which in turn influence future drinking behaviors. This highlights a critical opportunity for preventative measures. One promising direction is evaluating existing neuropsychological interventions, such as those reviewed by Verdejo-García, Alcázar-Córcoles (40). For instance, goal management training has been proposed as a means of enhancing goal-directed behaviors by training techniques such as mindfulness

practices, response inhibition, goal-setting, self-monitoring and decision-making strategies (41). Additionally, the ongoing study by Karl, Wieland (42) explores interventions like chess-based cognitive remediation and habit-modifying training in smokers, aimed at balancing goal-directed and habitual behavior. These approaches could be adapted to prevent risky drinking, highlighting a promising research avenue on the impact of such training or intervention on improving MB/MF decision-making and thus mitigating risky alcohol use.

Limitations (307 words)

Although we found evidence that MB/MF control predisposes future drinking behavior, this should not simply be interpreted as a dichotomy between goal-directed and habitual control (43, 44). The complexities underlying these constructs suggest that our findings might reflect broader cognitive processes rather than a straightforward binary categorization. Additionally, the stability of our measurements was modest. This could be attributed to two factors: significant changes and the fact that consistency measurement represents only the lower bound of real stability. Measurement errors could also contribute to the low ICC or correlation coefficients observed, emphasizing the need for further research to disentangle stable traits from the state-dependent aspects of unbalanced MB/MF control, which may provide a more profound understanding of their impact on drinking behavior. Additionally, the substantial amount of missing data at age 24 represents a limitation, as it required imputation methods; future studies should aim for larger initial sample sizes to more efficiently address attrition during critical developmental stages. The neural model for binge drinking scores demonstrated a suboptimal fit, requiring cautious interpretation; future studies should aim to increase the sample size and the

number of measurement time points to improve model robustness. Finally, our findings, derived from a male-only sample aged 18 to 24, limits the generalization of the results to other age groups, developmental stages, or to female populations. Having identified these associations in males, future research should examine whether they hold in females and more diverse samples. Additionally, the exclusion of participants with prior mental illness—intended to minimize variance and ensure task homogeneity—may have omitted particularly at-risk individuals, given the high comorbidity between mental illness and substance use. Moreover, these findings may not be applicable to other drinking cultures or countries with differing regulations regarding alcohol use, availability, and marketing, indicating a need for broader demographic and cultural representation in future research.

Conclusions (136 words)

Building upon our previous research, this study further elucidates the crucial role of MB and MF control in shaping drinking behaviors during young adulthood in non-dependent social drinking men. We found that increases of MB behavioral control act as a protective factor against the development of future binge drinking. Furthermore, changes in MF RPE signals in the VS and vmPFC both significantly impact future drinking behaviors. The VS signal appears to predispose individuals to future alcohol consumption, while a vmPFC signal may have a protective effect. Our study is the first to address both the predisposing factors and consequences of risky drinking behavior on MB/MF control. These findings highlight mechanisms that could potentially inform interventions during this pivotal developmental period, offering valuable insights for developing preventive strategies against risky drinking within this crucial age bracket.

Acknowledgments and Disclosures

This study was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) (Grant Nos. 186318919 [FOR1617], 178833530 [SFB 940], 402170461 [TRR 265] and 454245598 [IRTG 2773]). We extend our gratitude to Dr. Stephan Nebe for his significant contributions to the data collection and quality control, as well as for his valuable work in processing and developing analytical scripts for the baseline dataset.

QJMH acknowledges support by the NIHR UCLH BRC. QJMH has obtained fees and options for consultancies for Aya Technologies and Alto Neuroscience. All other authors report no biomedical financial interests or potential conflicts of interest.

Supplement Description:
Supplement Methods, Results, Figures S1-S4, Tables S1-S5

References

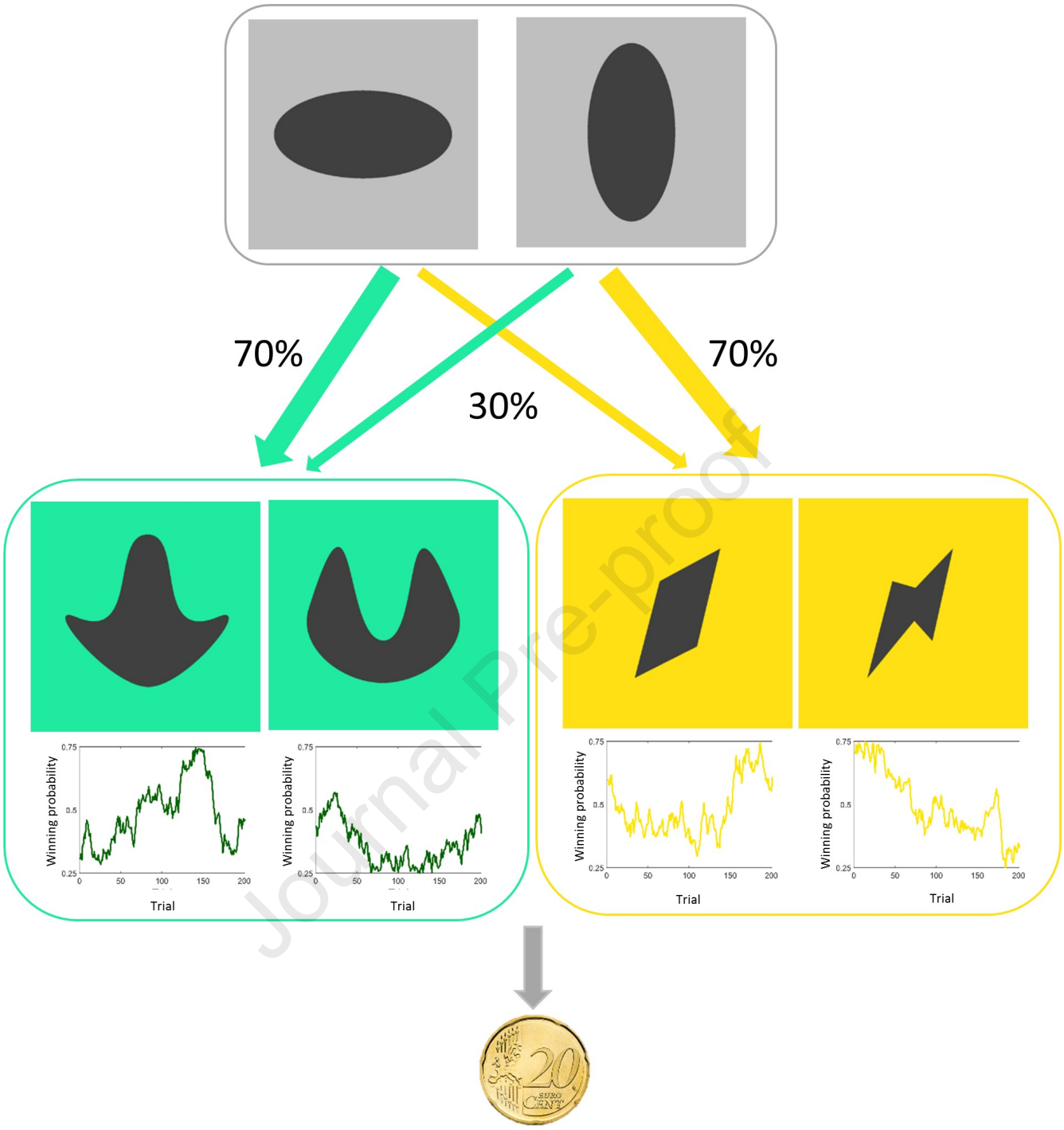
1. World Health Organization (2019): *Global status report on alcohol and health 2018*. World Health Organization.
2. Everitt BJ, Robbins TW (2016): Drug addiction: updating actions to habits to compulsions ten years on. *Annu Rev Psychol*. 67:23-50.
3. Heinz A, Gutwinski S, Bahr NS, Spanagel R, Di Chiara G (2024): Does compulsion explain addiction? *Addiction Biology*. 29:e13379.
4. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011): Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 69:1204-1215.
5. Schultz W (1998): Predictive reward signal of dopamine neurons. *Journal of neurophysiology*. 80:1-27.
6. Sharpe MJ, Batchelor HM, Mueller LE, Yun Chang C, Maes EJ, Niv Y, et al. (2020): Dopamine transients do not act as model-free prediction errors during associative learning. *Nature communications*. 11:106.
7. Keiflin R, Pribut HJ, Shah NB, Janak PH (2019): Ventral tegmental dopamine neurons participate in reward identity predictions. *Current Biology*. 29:93-103. e103.
8. Sadacca BF, Jones JL, Schoenbaum G (2016): Midbrain dopamine neurons compute inferred and cached value prediction errors in a common framework. *elife*. 5:e13665.

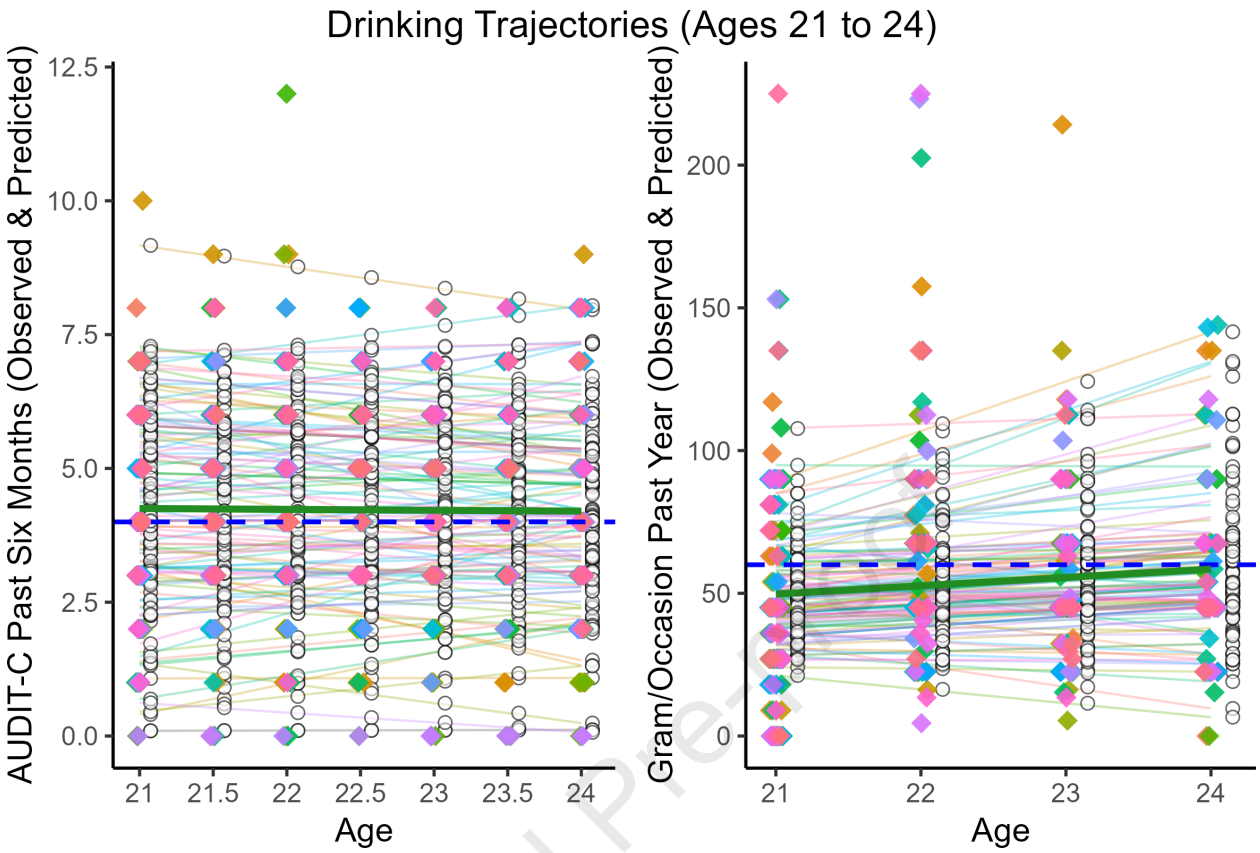
9. Deserno L, Huys QJ, Boehme R, Buchert R, Heinze H-J, Grace AA, et al. (2015): Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences*. 112:1595-1600.
10. Daw ND, Niv Y, Dayan P (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 8:1704-1711.
11. Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, et al. (2015): Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry*. 20:345.
12. Sebold M, Nebe S, Garbusow M, Guggenmos M, Schad DJ, Beck A, et al. (2017): When habits are dangerous: alcohol expectancies and habitual decision making predict relapse in alcohol dependence. *Biological psychiatry*. 82:847-856.
13. Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hägele C, et al. (2014): Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology*. 70:122–131.
14. Doñamayor N, Strelchuk D, Baek K, Banca P, Voon V (2018): The involuntary nature of binge drinking: goal directedness and awareness of intention. *Addiction biology*. 23:515-526.
15. Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND (2016): Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *elife*. 5:e11305.
16. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG, Organization WH (2001): AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care. World Health Organization.
17. Patzelt EH, Kool W, Millner AJ, Gershman SJ (2019): Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biol Psychiatry*. 85:425-433.
18. Kool W, Gershman SJ, Cushman FA (2017): Cost-benefit arbitration between multiple reinforcement-learning systems. *Psychol Sci*. 28:1321-1333.
19. Jackson KM, Sher KJ, Schulenberg JE (2008): Conjoint developmental trajectories of young adult substance use. *Alcoholism: Clinical and Experimental Research*. 32:723-737.
20. Tucker JS, Orlando M, Ellickson PL (2003): Patterns and correlates of binge drinking trajectories from early adolescence to young adulthood. *Health Psychol*. 22:79.
21. Nebe S, Kroemer NB, Schad DJ, Bernhardt N, Sebold M, Muller DK, et al. (2018): No association of goal-directed and habitual control with alcohol consumption in young adults. *Addict Biol*. 23:379-393.
22. Chen H, Mojtahedzadeh N, Belanger MJ, Nebe S, Kuitunen-Paul S, Sebold M, et al. (2021): Model-based and model-free control predicts alcohol consumption developmental trajectory in young adults: A 3-year prospective study. *Biol Psychiatry*. 89:980-989.
23. Saß H, Wittchen H-U, Zaudig M, Houben I (2003): DSM-IV-TR–Diagnostisches und Statistisches Manual Psychischer Störungen–Textrevision. *Hogrefe, Göttingen*.
24. Wittchen H-U, Pfister H (1997): DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft.
25. Hentschel A, Petzold J, Chen H, Heinz A, Smolka MN Higher alcohol use is associated with increased risk seeking towards gains in young men: A longitudinal cohort study. . *Unpublished manuscript under revision in Alcohol: Clinical and Experimental Research*.
26. Clifton L, Clifton DA (2019): The correlation between baseline score and post-intervention score, and its implications for statistical analysis. *Trials*. 20:1-6.
27. Rosseel Y (2012): Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of statistical software*. 48:1-36.
28. Arbuckle JL, Marcoulides GA, Schumacker RE (1996): Full information estimation in the presence of incomplete data. *Advanced Structural Equation Modeling: Issues and Techniques*. 243:277.
29. Orth B, Merkel C (2022): Der substanzkonsum jugendlicher und junger erwachsener in deutschland. Ergebnisse des alkoholsurveys 2021 zu alkohol, rauchen, cannabis und trends. :: Bundeszentrale für gesundheitliche Aufklärung.

30. Organization WH (2000): *International guide for monitoring alcohol consumption and related harm*. World Health Organization.
31. Dawson DA, Grant BF, Stinson FS, Zhou Y (2005): Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT‐C) in screening for alcohol use disorders and risk drinking in the US general population. *Alcoholism: Clinical and Experimental Research*. 29:844-854.
32. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K (2014): Mediation: R package for causal mediation analysis.
33. Zhao X, Lynch Jr JG, Chen Q (2010): Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of consumer research*. 37:197-206.
34. Koob GF, Volkow ND (2016): Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry*. 3:760-773.
35. Levy DJ, Glimcher PW (2012): The root of all value: a neural common currency for choice. *Curr Opin Neurobiol*. 22:1027-1038.
36. Christakou A, Gershman SJ, Niv Y, Simmons A, Brammer M, Rubia K (2013): Neural and psychological maturation of decision-making in adolescence and young adulthood. *Journal of cognitive neuroscience*. 25:1807-1823.
37. Feher da Silva C, Lombardi G, Edelson M, Hare TA (2023): Rethinking model-based and model-free influences on mental effort and striatal prediction errors. *Nature Human Behaviour*.1-14.
38. Dufour MC (1999): What is moderate drinking?: Defining “drinks” and drinking levels. *Alcohol Research & Health*. 23:5.
39. Groman SM, Massi B, Mathias SR, Lee D, Taylor JR (2019): Model-free and model-based influences in addiction-related behaviors. *Biol Psychiatry*. 85:936-945.
40. Verdejo-García A, Alcázar-Córcoles MA, Albein-Urios N (2019): Neuropsychological interventions for decision-making in addiction: A systematic review. *Neuropsychol Rev*. 29:79-92.
41. Valls-Serrano C, Caracuel A, Verdejo-Garcia A (2016): Goal Management Training and Mindfulness Meditation improve executive functions and transfer to ecological tasks of daily life in polysubstance users enrolled in therapeutic community treatment. *Drug Alcohol Depend*. 165:9-14.
42. Karl D, Wieland A, Shevchenko Y, Grundinger N, Machunze N, Gerhardt S, et al. (2023): Using computer-based habit versus chess-based cognitive remediation training as add-on therapy to modify the imbalance between habitual behavior and cognitive control in tobacco use disorder: protocol of a randomized controlled, fMRI study. *BMC psychology*. 11:24.
43. Miller KJ, Shenhav A, Ludvig EA (2019): Habits without values. *Psychol Rev*. 126:292.
44. Collins AG, Cockburn J (2020): Beyond dichotomies in reinforcement learning. *Nature Reviews Neuroscience*. 21:576-586.

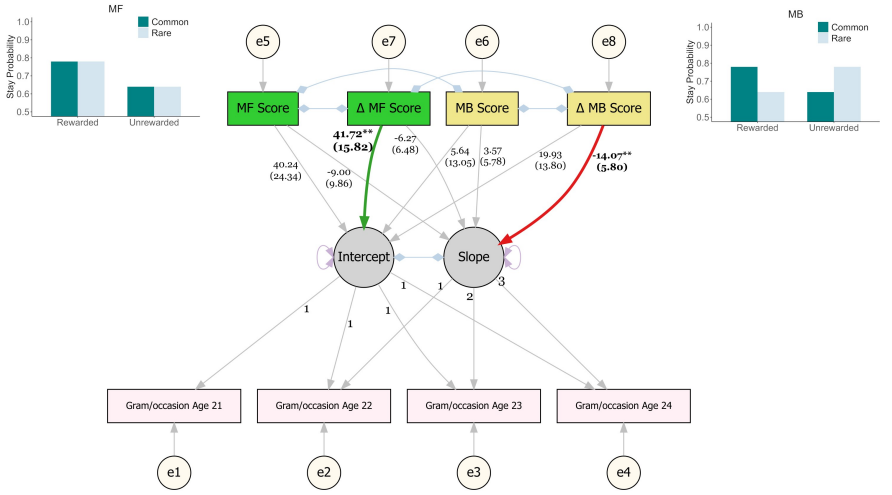
KEY RESOURCES TABLE

Resource Type	Specific Reagent or Resource	Source or Reference	Identifiers	Additional Information
Add additional rows as needed for each resource type	Include species and sex when applicable.	Include name of manufacturer, company, repository, individual, or research lab. Include PMID or DOI for references; use “this paper” if new.	Include catalog numbers, stock numbers, database IDs or accession numbers, and/or RRIDs. RRIDs are highly encouraged; search for RRIDs at https://scicrunch.org/resources.	Include any additional information or notes if necessary.
Software; Algorithm	R project for statistical computing	R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/.	RRID:SCR_001905	
Software; Algorithm	R lavaan package	Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/	N/A	
Software; Algorithm	R ggplot2 package	Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org	RRID:SCR_014601	
Software; Algorithm	R DiagrammeR package	https://CRAN.R-project.org/package=DiagrammeR	N/A	
Software; Algorithm	Matlab v2018b	Mathworks	RRID:SCR_001622	
Software; Algorithm	Statistical Parametric Mapping (SPM) software package v12	Wellcome Trust Centre for Neuroimaging, London, UK)	RRID:SCR_007037	

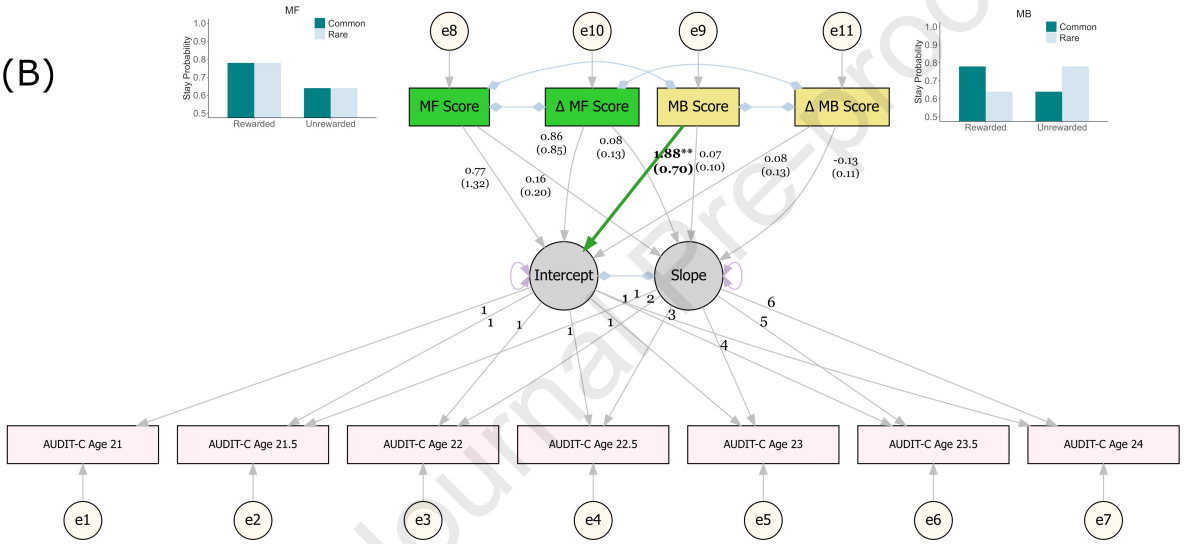




(A)



(B)



(C)