Forward planning under uncertainty in a population-based alcohol use disorder sample

Pascale C. Fischbach[1+], Lorenz Gönner[1+], Johannes Steffen[1+], Stefan J. Kiebel[2], Michael N. Smolka[1*]

[1]Department of Psychiatry, Technische Universität Dresden

[2]Department of Psychology, Technische Universität Dresden

**Corresponding author*

[+]Shared first authorship

Corresponding author: Michael N. Smolka

Full postal address: Section of Systems Neuroscience, Department of Psychiatry, Technische Universität Dresden, Würzburger Str. 35, 01187 Dresden, Germany

Telephone number: +49 351 463-42201

Fax number: +49 351 463-42202

Email address: michael.smolka@tu-dresden.de

Funding This research was funded by the German Research Foundation (project number 402170461 - TRR 265).

ABSTRACT

Altered decision-making is a defining component of addiction, but there is little evidence whether these alterations affect multi-step planning in individuals with alcohol use disorder (AUD). We used a recently developed planning task in a cross-sectional approach to test the planning performance of 30 individuals diagnosed with AUD relative to 32 healthy control subjects, both sampled from the general population. To gain insight into the factors underlying behavioral performance, we inferred the parameters of a reinforcement learning agent performing rational planning, using a Bayesian inference scheme. This approach allowed us to differentiate between separate factors determining planning performance, e.g., planning depth, decision noise, and bias for certain action choices. Contrary to our hypothesis, we did not observe reduced planning depth in AUD subjects. Instead, our results show a small effect in the opposite direction: Healthy controls were slightly less efficient in the planning task. Importantly, subjects in the control group allocated less time for planning than AUD subjects, potentially indicative of motivational differences between groups. The group difference in planning depth persisted when controlling for both reaction times and general cognitive performance, albeit at a lesser magnitude. Altogether, our results do not favor the view that mild-to-moderate alcohol use disorder in general-population individuals generally involves impairments in cognitive tasks requiring forward planning across multiple steps.

INTRODUCTION

Substance use disorders often result in serious somatic, mental health as well as social and economic consequences. Especially the harmful use of alcohol is one of the most important risk factors for population health worldwide (World Health Organization, 2018). Despite these detrimental effects, our understanding of the mechanisms and trajectories in the development of alcohol use disorders (AUD) remains limited. According to the dominating “transition-to-habit” theory (Everitt and Robbins, 2005), the process of developing an AUD begins with adverse goal-directed decision-making leading to repetitive consumption behavior which later becomes habitual and compulsive. The exact principles of this adverse decision-making of certain individuals are still an unresolved issue. One option to approach these decisions is model-based forward planning, which rests upon knowledge about the consequences and a weighing of their negative and positive values (Dolan and Dayan, 2013; Sutton and Barto, 2018). One obvious challenge of forward planning is the question of how deep one should plan ahead (Steffen et al., 2022). Deeper forward planning leads to more accurate valuations of consequences but is also computationally more costly. Regarding decisions on alcohol consumption, negative consequences usually manifest sequentially and at a later stage than rewarding effects. Hence, high planning depth (PD) could be a crucial parameter to prevent individuals from developing increasingly habitual consumption behavior. In this study, we therefore aimed at investigating how PD differs between individuals fulfilling an AUD diagnosis and healthy controls (HC).

So far, no study has examined the depth of forward planning in these subjects. However, related evidence comes from two branches of research: First, planning capabilities have classically been

studied using sequential problem-solving tasks like the Tower of London Task (ToL; Shallice, 1982). While these tasks do not directly assess PD, they provide indicators of forward planning performance, where lower performance usually means more moves needed to reach a target configuration. In a meta-analysis of 39 studies by Stephan et al. (2017), AUD samples showed a significantly lower planning performance according to a planning and problem-solving composite score. In contrast, in a general population sample study by Black et al. (2017), performance in the ToL was not associated with heavy episodic drinking the two weeks later. Secondly, the two-stage Markov task by Daw et al. (2011) has been used to study the balance between model-based and model-free control in sequential decision-making, where the former is equivalent to a simple form of forward planning. Studies using this approach in clinical AUD samples either found no difference in model-based control (Sebold et al., 2017; Voon et al., 2015) or the effect did not survive control for performance in a neuropsychological test (Sebold et al., 2014). However, Voon et al. (2015) found a positive association of the model-based weight parameter with the number of weeks subjects were abstinent indicating that forward planning does play a role in the control of consumption behavior. Studies focusing on general-population samples also showed ambiguous results. While Gillan et al. (2016) found a significant negative association of the number of self-reported AUD symptoms with model-based control in a big general-population sample, no association was found between model-based control and characteristics of recent alcohol use in 18-year-old social drinkers (Nebe et al., 2018) or the effect could not be replicated during a second experiment (Byrne et al., 2019). However, a longitudinal assessment of the sample of Nebe et al. (2018) revealed that model-based control predicted binge drinking trajectories for the subsequent years (Chen et al., 2021). Taken together, previous research provides an indistinct picture of the relationship of forward planning and the development of problematic alcohol use. One reason for this could be that previous tasks did not differentiate computational parameters of forward planning, such as PD. Moreover, all reported studies on goal-directed decision-making in AUD were either based on clinical samples of detoxified alcohol-dependent participants, which most commonly have had a long history of heavy drinking and severe AUD, or healthy adults without an AUD diagnosis and rare harmful alcohol use. To shed more light on the role of goal-directed planning in the development of AUD, we therefore focused the current study on a population-based, non-treatment-seeking AUD sample.

To assess PD in the described sample in comparison to HC, we used our sequential decision-making task, the Space Adventure Task (SAT), as described already by Steffen et al. (2022). Participants' choices in this task were modelled with a reinforcement learning agent model with PD as a free parameter that could be inferred with Bayesian Inference (see Methods section). Planning was incentivized by higher rewards for choices that were based on deeper planning. Moreover, the task contained two different

levels of uncertainty in state transitions to investigate how participants might adjust planning when consequences of actions become less predictable.

Furthermore, forward planning is influenced by several cognitive abilities like working memory, processing speed and general crystalline intelligence (Otto et al., 2014, 2013; Sebold et al., 2014). These have been shown to be impaired in AUD samples (Chanraud et al., 2007; Sebold et al., 2014; Stephan et al., 2017). We therefore considered measures of these constructs as covariates in our analysis to investigate to what extent they might explain differences in PD.

Based on previous research, we hypothesized that in the SAT, participants fulfilling an AUD diagnosis compared to HC should demonstrate reduced forward planning capabilities indicated by lower scores and a lower inferred PD. Regarding outcome certainty, we presumed that a less predictable environment should lead to a reduction in PD. Finally, we expected performance measures for the three assessed general cognitive abilities to be positively associated with PD while not fully explaining group differences in PD.

MATERIALS AND METHODS

Participants and procedure

Potential candidates were recruited from the Dresden city area by distribution of flyers and online advertisement. Interested candidates were prescreened via telephone for our inclusion and exclusion criteria. We included individuals aged from 16 to 65 years. Criteria of AUD according to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V; American Psychiatric Association, 2013) were diagnosed with a structured clinical interview (SCID-5-CV; First, 2016). Individuals were excluded if they reported any physical withdrawal symptoms from alcohol, a lifetime history of mania or psychosis, an acute severe major depression episode or suicidal thoughts, a severe illness of the central nervous system or any previous or current substance dependencies other than alcohol or tobacco. Female participants were additionally excluded in case of pregnancy or if they were currently nursing an infant. On site, participants were only assessed if they had a measured breath alcohol value of 0‰ and no positive urine test result for any substance other than cannabis. After giving written consent, participants filled out a sociodemographic questionnaire, worked on the SAT, the neuropsychological tests and were additionally interviewed about alcohol use characteristics with the AUDIT (Saunders et al., 1993). At the end, participants reporting 2 or more AUD symptoms were seen by our study physician and offered contact mediation with the outpatient clinic or addiction counseling centers if wanted. The overall procedure took approximately 2.5-3 hours and participants were compensated with 25-35€ depending on their performance in the SAT. The ethics committee of TU Dresden gave ethical approval for this work (EK514122018).

Out of 70 invited individuals, 65 met our inclusion criteria and managed to complete the assessment. Three subjects had to be excluded due to close to random behavior in the main task of the experiment. The remaining *N*=62 participants were either assigned to the AUD group (*n* = 30) fulfilling 2 or more AUD criteria or to the HC group (*n* = 32) fulfilling less than 2 AUD criteria. According to the DSM-5 degrees of AUD severity, the great majority (90%) of the AUD group showed a mild to moderate form of AUD (see Figure 1). For information on substance use behavior and additional demographic data for both groups see Table 1.

Table 1. Substance use variables and additional demographic data.

		AUD	HC
Gender	n_m/n_f	18/12	13/19
Age	*M (SD)*	28.3 (8.6)	29.8 (10.8)
AUDIT score	*M (SD)*	14.72 (5.07)	6.78 (4.94)
DSM-V AUD criteria	*M (SD)*	3.57 (1.48)	0.25 (0.44)
Higher Education Entrance Qualification	*%*	83.3	78.1
Employed	*%*	83.33	70.97

Note. AUDIT: Alcohol Use Disorder Identification Test. DSM-V: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. AUDIT data missing for 1 AUD subject because of technical failure. AUD: Alcohol Use Disorder. n_m/n_f: Number of males / females.

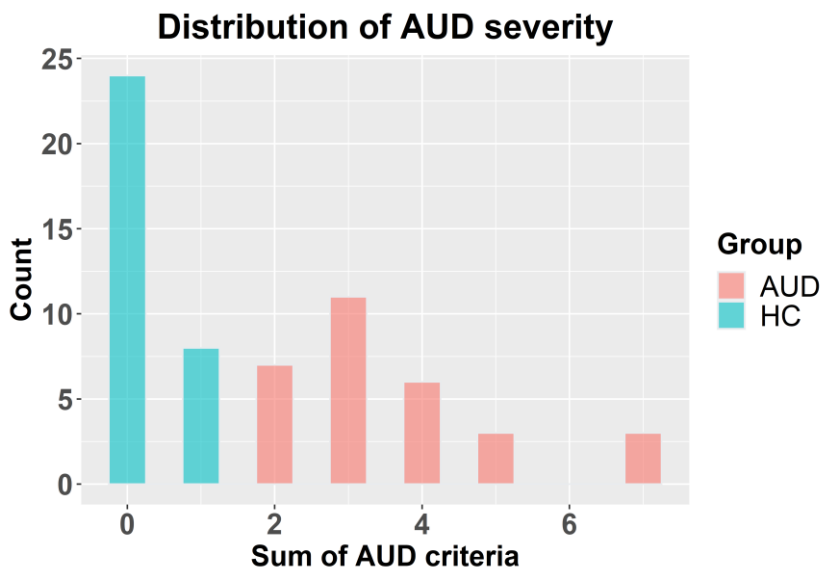


Figure 1: Distribution of the number of AUD criteria fulfilled per participant.

Neuropsychological Tests

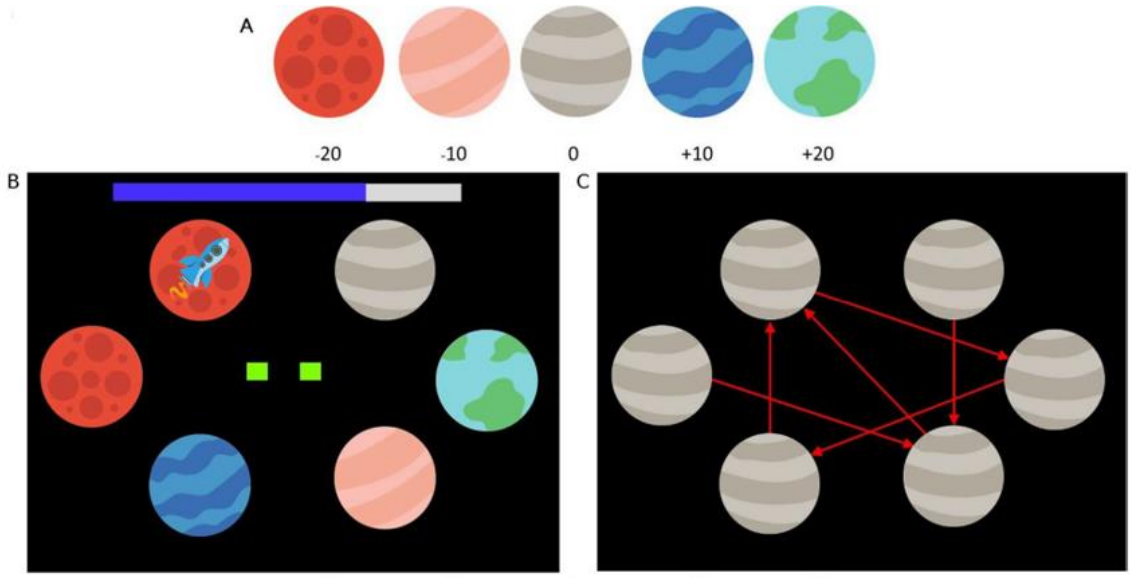
We used the first subtask of the computerized *Spatial Working Memory* (SWM) task (Nagel et al., 2008) as a measure for spatial processing abilities. In this task, participants are presented with a 4 x 4 grid of circles in which a series of dots is displayed consecutively in specific locations of the grid. At the end of each sequence one circle is marked and participants have to indicate whether a dot was presented at that position or not. To assess verbal intelligence as part of crystallized intelligence, we used a German vocabulary test, the Spot-A-Word (SAW) test as used in the Berlin Aging Study (Lindenberger et al., 1993). Here, participants indicate which word out of five is not a pseudoword but the real one as quickly and accurately as possible. As a measure for processing speed, participants conducted the *Identical Pictures Task* (IDP; Lindenberger et al., 1993). In this task, participants are presented with five symbols and have to find the symbol matching another one at the top as quickly and correctly as possible. Finally, as a measure of fluid intelligence, we used a computerized implementation of the short-form of *Raven's Advanced Progressive Matrices* (APM-short; Bors and Stokes, 1998). For the SWM, the IDP and the APM-short, we acquired reaction time (RT) and accuracy and used accuracy as a measure of performance.

Space Adventure Task

The SAT is a computerized sequential decision-making task which requires forward planning to perform effectively. During a fixed sequence of 100 mini-blocks, participants had to accumulate as much fuel as possible (indicated by a fuel bar at the top) by visiting planets with their spaceship. Figure 2A shows the planet types with their respective fuel rewards. Figure 2*B* depicts an example mini-block with three steps left (green squares) and low noise (black background). The mini-blocks were designed in a way that participants had to plan forward to find the route leading to maximum fuel gain with either two or three steps. At each step, participants could choose between two actions: (i) move to the next planet in a clockwise fashion by using 2 fuel points or (ii) jump to a non-neighboring planet as determined by a given travel pattern (Figure 2*C*) which consumed 5 fuel points. While moving clockwise was deterministic, jumping was assigned a certain probability for the transition to fail causing the spaceship to land on one of the neighboring planets of the target planet, each with equal probability. In low noise mini-blocks, jumping failed with a 10% probability, while in high noise mini-blocks (indicated by asteroids in the background) it failed with a 50% probability. Prior to the main set of mini-blocks, participants were explicitly instructed about all aspects of the tasks, implicitly learned the jumping failure probability ("noise") by trial and error, memorized the travel pattern and familiarized with the procedure during 20 training mini-blocks. The SAT was implemented in MathWorks MATLAB versions 2017 and 2018 and run on a standard PC. Participants entered their responses via keyboard with key

'S' indicating the *jump* action and the right arrow key indicating the *move* action. For an exhaustive description of the SAT see Steffen et al. (2022).

Figure 2. Schematic of the Space Adventure Task



Note. (A) Five planet types with different amounts of fuel to win or lose when landing on them. (B) Bar with fuel on top of the screen, green squares in the middle of the screen showing how many actions are left, no asteroids indicating high jumping success probability and planet configuration with rocket on starting planet. (C) Jump pattern.

Computational model

We modelled participants' action choices with a mixture model of three single model-based RL agent models with PD of 1, 2 and 3, respectively. Each agent had an optimal model of the environment, i.e. it was completely informed about the rules of the task. This environment model entailed the set of available actions $A = \{'move', 'jump'\}$ and states S (the planet configuration of the mini-block), the transition probabilities $p(s_{t+1}|s_t, a_t)$ for reaching a subsequent state s_{t+1} from a given state s_t with action a_t as well as the immediate reward of reaching each state $r(\mathrm{s})$ indicating the planet types of the current configuration. To choose the optimal action for a specific state in a specific mini-block, the agents computed the expected cumulative reward for executing each action with an optimal forward planning algorithm (value iteration) only limited by their PD.

Value iteration outputs the expected cumulative reward for executing each action a in the current state s with planning horizon d, which is called the state-action- or Q-values. These Q-values were the essential values for action selection. The higher the relative value of an action was, the higher should have been the probability of selecting that action. Action selection was therefore as usual modelled

probabilistically with a softmax function (Sutton and Barto, 2018). For our case of two available actions, this corresponded to a sigmoid transformation $\sigma(x)$ of the difference between the Q-values, $\Delta Q(s_t, d)$. Choice probabilities were thus defined as:

$$p(a_t = 'jump' | s_t, d) = \sigma(\beta * \Delta Q(s_t, d) + \theta), \tag{1}$$

where

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

$$\Delta Q(s_t, d) = Q(a_t = 'jump', s_t, d) - Q(a_t = 'move', s_t, d). \tag{3}$$

Here, the inverse temperature beta (β) controlled the extent to which differences in Q-values affected action selection. If beta $= 0$, actions are selected with a constant probability independent of outcomes, while higher values of beta represented higher probability to select the action with the highest Q-value. The parameter theta (θ) denoted an a priori response bias, where positive values imply a bias towards choosing 'jump'.

Since the state transition probabilities for the jump action $p(s_{t+1} | s_t, a_t = 'jump')$ were not given explicitly during training, we assumed an experience-based learning process of the state transition probabilities. The belief about the probability that a jump will be successful $\rho = p(s_{t+1} = 'target' | s_t, a_t = 'jump')$ was updated using the temporal difference rule:

$$\rho_{t+1} = \rho + \alpha(o_t - \rho_t), \tag{4}$$

depending on the experienced success $(o_t = 1)$ or failure $(o_t = 0)$ of a jump. The learning rate parameter alpha (α) modelled how fast participants changed their beliefs about the probability of transition success. Larger values of alpha could also be interpreted as faster forgetting of prior experience and stronger reliance on recent outcomes.

Planning depth and parameter inference

To infer the four described free model parameters, inverse temperature beta, response bias theta, learning rate alpha and PD (d), from participants' choices, we used a hierarchical probabilistic model of free parameters and approximate Bayesian inference scheme. The approximate inference can be imagined as a two-step procedure. Firstly, approximate posteriors of beta, theta and alpha were computed for each of the three RL agent models with PD $d \in \{1,2,3\}$. Secondly, these agents with their respective inferred parameter distributions for beta, theta and alpha were used to infer the posterior over PD d. Note that the response likelihood is a function of a specific PD d (compare Equation (5)). Therefore, the response likelihood of each participant corresponds to a mixture model of these three possible, PD dependent, response likelihoods. Hence, we can express the mixture model as:

$$p(a_b|s_b) = \sum_{d=1}^{3} p(d_b = d)p(a_b|s_b, d_b = d) \quad (5)$$

Importantly, we assumed that forward planning should mostly happen before the first action of each mini-block. Hence, for the model inversion (parameter inference), we have constrained behavioral data only to the first choice in each mini-block.

In Equation (5), $p(d_b = d)$ denotes the prior probability each PD has in the mixture model. To account for the limited amount of behavioral data, and for expected within group similarities in participants' responses, we designed the probabilistic generative model in a hierarchical fashion with group-level, subject-level and mini-block-level dependence of parameter values. The parameters beta, theta and alpha were modelled on the subject-level, while PD d was modelled on the mini-block-level. As an analytical solution for the posteriors of the parameters was intractable, we instead used the stochastic variational inference scheme from the probabilistic programming library Pyro v1.5.2 (Bingham et al., 2019) to approximate the posterior distributions. For more details on the computational model and inference procedure see Steffen et al. (2022).

Statistical Analysis

Statistical Analysis was performed using R, version 4.1.2 (https://cran.r-project.org/). Linear mixed models and estimated marginal means were computed with the lme4 package (version 1.1-27.1) and the emmeans package (version 1.7.2).

For all relevant variables (including task outcome, inferred model parameters, sociodemographic, and cognitive covariates), we tested for normal distribution via the Shapiro-Wilk test, and for variance homogeneity across groups via Levene's test. If non-normality was a concern, we used the Wilcoxon test for group comparisons. In case of variance inhomogeneity, we compared groups using a permutation test with 1000 samples. Otherwise, we used two-sided *t*-tests. For ease of reporting, we report the results of t-tests whenever their result was highly similar to other tests used.

For analysis of the SAT, we first checked each participant for random behavior in the SAT. For this, we simulated 1000 random agents with β=0 and compared loss of fuel points over the course of the paradigm. We excluded three participants which scored below 500 points, equivalent to the mean gain obtained by an agent with PD 1 with β = 1. We also considered RTs of the first action and β-values as criteria for exclusion from further analysis. RTs should not be below the common timeframe of 200ms for perceptual and motor processes (Whelan, 2008) and the mean of β should have at least a value of one to be far enough from random behavior. No participant had any reaction time below 200ms in the first action, and values of β < 1 occurred only in those three participants with scores below 500 points.

For the analysis, mean PD and RTs were averaged for each participant and condition (i.e., step- and uncertainty-condition).

We evaluated the inferred mean PD using a linear mixed-effects model, with fixed effects for group, step condition, and noise condition, and random intercept per subject (Model 1.1).

To further test whether group differences in mean PD could be explained by performance in the assessed neuropsychological variables and by reaction times, we used another linear mixed effects model in which spatial working memory task performance and reaction times were added, and other effects were the same as in Model 1.1.

RESULTS

We report descriptive statistics in Table 2. When analyzing for outlier data based on RTs below 200 ms, no SAT mini-blocks had to be excluded. In the SWM task, one trial with an RT < 200 ms had to be excluded for three participants.

Table 2. Descriptive Statistics and group comparison of task outcomes and model parameters.

		AUD (N=30)	**HC** (N=32)	**Statistic**	***p***	***Cohen's d***
Sample characteristics						
Age	*M (SD)*	28.3 (8.6)	29.8 (10.8)	$t(58.4)=-.64^{a}$	.53	
Gender (female/male/NA)		12/18/0	19/10/1	$\chi^2(1)=57.07$	<.001	
Task performance						
Total points (final score)	*M (SD)*	1373.5 (196.9)	1164.8 (304.7)	$t(219)=6.45^{b}$	<.001	
Model parameters						
Mean PD	*M (SD)*	2.3 (.1)	2.1 (.3)	$t(45.9)=2.74$	.009	
Beta	*M (SD)*	2.3 (.6)	1.8 (.9)	$t(56.5)=2.57^{b}$	.013	
Theta	*M (SD)*	-.4 (.4)	-.1 (.4)	$t(58.7)=-2.1$	.04	-.53

Alpha	*M* *(SD)*	.0003 (.0001)	.04 (.05)	*t*(31)=-4.7	<.001	-1.15
Cognitive covariates						
IDP	*M* *(SD)*	94.5 (2.8)	94.6 (2.5)	*t*(58.5)=-.08	.97	
APM-short	*M* *(SD)*	67.8 (18.7)	59.4 (21.5)	*t*(59.7)=1.6[b]	.11	
SAW	*M* *(SD)*	64.5 (8.9)	61.2 (15.4)	*t*(50.4)=1.03	.31	
SWM, spatial memory accuracy (%)	*M* *(SD)*	94.4 (4.8)	90.7 (6.4)	*t*(57.0)=2.5	.015	.63
Reaction times [sec]						
SAT mean RT (first action)	*M* *(SD)*	9.9 (5.0)	7.2 (3.9)	*t*(54.6)=2.39	.021	.61

Note: PD: planning depth; IDP: Identical Pictures task; APM-short: Ravens Advanced Progressive Matrices, short form, SWM: Spatial Working Memory task; SAW: Spot-A-Word task; RT: reaction time; SAT: Space Adventure Task. [a] Because of non-normality, an additional Wilcoxon test was performed which yielded comparable results. [b] Because of variance inhomogeneity, an additional permutation test was performed with 1000 shuffles which yielded similar results.

Performance and planning depth

To our surprise, we found that HC subjects scored less points than AUD subjects ($t(219)=208.7$, $p=.001$), which was associated with a PD about 0.2 steps lower than the AUD groups PD (see Table 2). A linear mixed-effects model (Model 1.1: PD ~ Group * steps + Group * noise, with random intercept per subject) revealed a main effect of steps ($t(182)=32.291$, $p < .001$) and a group x steps interaction ($t(182)$, $p < .001$; other p's > .013), but no main effect of group ($t(239.8)=1.485$, $p= .14$; see Table 3). A post-hoc analysis based on estimated marginal means showed that PD was lower for HC than AUD subjects specifically in the 3-step condition (estimate = 0.223, $t(71.6)=3.667$, $p<.001$). For distributions of PD across groups and step condition see Figure 3. Consistent with previous findings (Steffen et al., 2022), we observed neither a significant effect of the noise condition ($t(182)=-1.253$, $p=.21$) nor a significant interaction effect of group x noise ($t(182)=1.086$, $p=.28$).

We observed significantly faster RTs for HC subjects (t(54.6)=2.39, p=.021). Further, across groups and conditions, we observed that RTs were positively associated both with points gained (r=.57, p < .001) and with PD (r=.49, p < .001, see Figure 4).

Turning to the inferred parameters for the planning model, we observed significantly lower beta values for HC subjects (t(56.5)=2.57, p=.013), indicating higher decision noise. Further, AUD subjects showed lower values of the bias parameter theta (t(58.7)=-2.1, p=.04). For the learning rate parameter alpha, HC subjects showed higher values than AUD subjects (t(31)=-4.67, p<.001), indicating a larger influence of recently observed transitions on the estimated jump success probability for HC. For parameter distributions across groups, see Figure 5.

Table 3. Estimates of Linear Mixed Effects Model for Planning Depth.

	b	*SE b*	*p*
Intercept	.199*		.011
Group	.160	.108	.139
steps	.826***	.026	< .001
noise	-.032	.026	.211
Group x steps	-.134***	.036	<.001
Group x noise	.039	.036	.278

Cognitive covariates

In terms of cognitive covariates, we did not observe performance differences between groups for the IDP task (t(58.5)=-.08, p=.97), the APM-short form (t(59.7)=1.6, p=.11), and the SAW task (t(50.4)=1.03, p=.63). However, we observed significantly lower performance in HC participants in the SWM task (spatial memory accuracy: t(57.0)=2.5, p=.015). For the serial memory component of the SWM task, most participants showed low numbers in serial memory responses overall (less than 50%). In our view, this makes an interpretation of serial memory performance and its possible effect on the task performance difficult, thus not providing substantial additional explanation. We therefore do not report the SWM serial memory accuracy here.

To control for the potentially confounding effect of this variable on the group comparison of PD, we included the SWM task and RTs in a linear mixed-effects model of PD (Model 1.2: PD ~ Group * steps + Group * noise + SWM spatial accuracy + RT, with random intercept per subject, see Table 5). This improved the model fit relative to model 1.1 (χ^2(2)=16.93, p < .001.). In model 1.2, we observed significant main effects for group (t(234.2) = 2.95, p=.003) and steps (t(180.5)=31.62a, p < .001), as well as a significant group x steps interaction (t(179.3)=-3.68, p <.001). While the effect of SWM spatial

accuracy was significant ($t(58.0)=-3.39$, $p < .001$), the effect of RTs was marginal ($t(238.5)=1.71$, $p=.09$). A post-hoc analysis based on estimated marginal means showed that for the 3-step condition, the contrast estimate for PD between AUD subjects and controls remained significant, but was reduced in magnitude relative to model 1.1 (estimate=.15, $t(74)=2.67$, $p=.009$). This shows that the reduced PD for control subjects may in part be explained by group differences in cognitive performance.

Table 4: Estimates of Linear Mixed Effects Model for Planning Depth Including Cognitive Covariates.

	b	SE *b*	*p*
Intercept	.247**	.084	.003
Group	.230*	.106	.032
steps	.821***	.026	.000
noise	-.008	.029	.774
SAT reaction time	.005	.003	.089
SWM accuracy (%)	-.016**	.005	.001
Group x steps	-.132***	.036	.000
Group x noise	.027	.036	.454

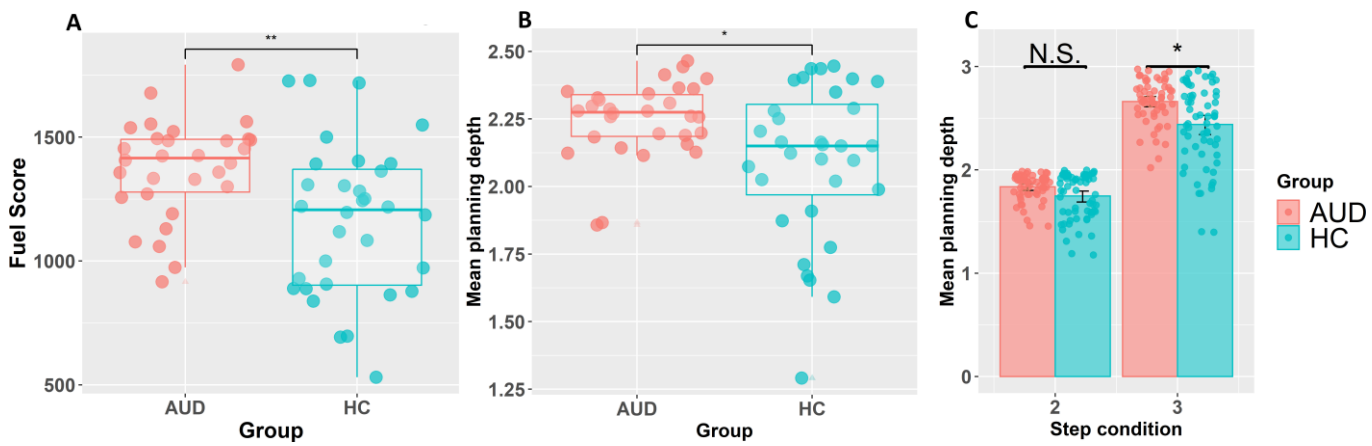


Figure 3: SAT Behavioral performance and planning depth. A: Fuel score (points gained at the end of the paradigm). B: Mean planning depth averaged across conditions. C: Mean planning depth per step condition.

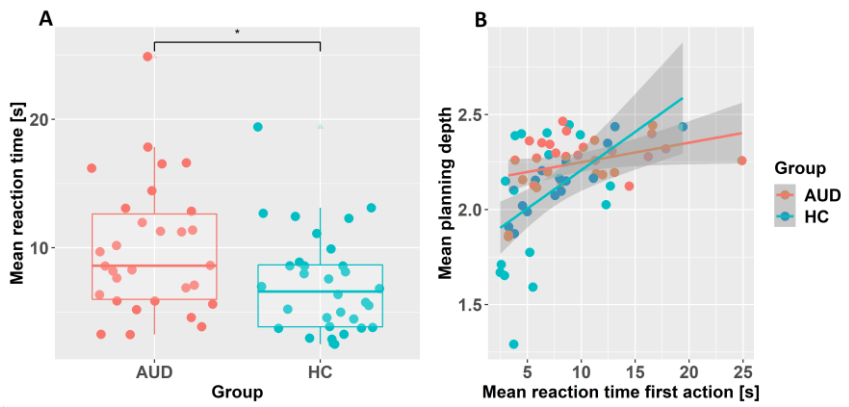


Figure 4: SAT Reaction times. A: Mean reaction times for the first action. Mean across conditions per group. B: Mean planning depth (first action, across conditions) versus SAT reaction times.

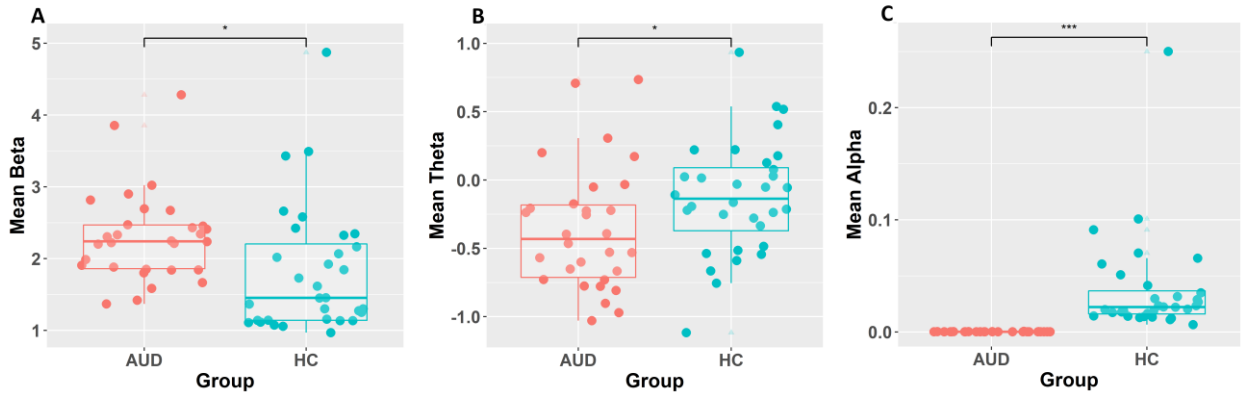


Figure 5: Estimated parameters for the planning agent model. Mean across 100 samples from the posterior distribution. A: Inverse temperature (decision noise) parameter beta. B: Action bias parameter theta. C: Learning rate parameter alpha.

DISCUSSION

The current study was aimed at assessing forward PD in a population-based AUD sample in comparison to a HC group. Contrary to our expectations, we did not find any evidence of impaired forward planning capacity in individuals fulfilling an AUD diagnosis. Indeed, we even found small effects in the opposite direction: the AUD group showed a slightly higher PD in the more complex 3-step condition, spent more time planning and – according to model parameters – showed less response noise and a smaller learning rate indicating less distractibility after unexpected transitions. Even after controlling for group differences in general cognitive performance measures, the slightly higher PD in the AUD group remained.

Although we expected the AUD group to show impaired forward planning, our findings are not very surprising considering the ambiguous previous evidence. As discussed in prior studies on model-based decision-making, forward planning may as well just be impaired in severe cases of longstanding problematic alcohol consumption and therefore be one aspect of cognitive symptomatology of alcohol-induced brain damage (Hogarth, 2020). However, other studies did find predictive power of model-based decision-making capacities for consumption behavior or the success of abstinence decisions (Chen et al., 2021; Voon et al., 2015). The cross-sectional nature of the current study may not be suitable to capture these effects. Future longitudinal investigations should shed more light on the exact dynamical interplay between forward planning and trajectories of AUD development.

Another explanation of our findings could lie in the type of AUD sample we recruited. As previous studies only focused on detoxified patients with long records of alcohol problems or healthy samples without current problems, the intermediate stage during the development of AUD which our study focused on is not well characterized yet. We believe that this stage is a central switch point determining which individuals will more likely transgress into more severe forms of AUD into more severe forms of AUD. We suspect that during this stage, boosted forward planning could act as a compensatory mechanism in individuals experiencing first problems but still maintaining normal levels of functioning (Scheurich, 2005). There might even be a characteristic selection bias such that resilient individuals who are coping well with minor problems in controlling their alcohol consumption might be more likely to participate in a study as ours. This group of people might also be particularly interested in getting accurate feedback on their current state of health and therefore be distinctively motivated to show good performance in the study whereas HC subjects are mainly motivated by the monetary compensation.

A third explanation comes from the current suggestion that compulsive drug-seeking behavior may not be habitual but may instead be excessive goal-directed behavior for a persistent goal (Hogarth, 2020). This means that repeated drug intake might even strengthen goal-directed forward planning, especially when the goal is to organize consumption. The aspect of automaticity in compulsion is instead suggested to occur on the higher hierarchical level of goal-selection. Evidence for this comes from a study by Black et al. (2017) for example who found that in participants with strong intentions to drink, higher planning performance and longer planning time in the ToL predicted more frequent heavy episodic drinking. This supports the idea that forward planning facilitates hazardous drinking when the goal to do so is already set.

Limitations

A limitation of our work is that the task design turned out to be suboptimal to evaluate the effect of the noise condition, which therefor did not allow a stringent interpretation. Further, the outbreak of

the Covid-19 pandemic and the subsequent lockdown interrupted the assessment. In particular, while most control subjects were assessed before the lockdown, most AUD subjects were assessed after the first lockdown, which may cause an additional confound in our data with regard to potential effects of the lockdown on mental well-being and additional AUD symptoms.

Conclusion

Our findings suggest that planning capacity, a key aspect underlying goal-directed behavior, is not reduced in the early stage of AUD. In relation to more severe cases of the disease, associated with deteriorated cognitive functioning in the long term, the question remains when the turning point between these opposite ends is approached. Further work is needed to address this question, preferably involving a combined longitudinal and cross-sectional design, and including a sample of individuals with moderate and severe AUD but who do not seek treatment.

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Fifth edition.). American Psychiatric Association.

Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P., & Goodman, N. D. (2019). Pyro: Deep Universal Probabilistic Programming. *The Journal of Machine Learning Research*, *20*(1), 973–978.

Black, N., Mullan, B., & Sharpe, L. (2017). Predicting heavy episodic drinking using an extended temporal self-regulation theory. *Addictive Behaviors*, *73*, 111–118. https://doi.org/10.1016/j.addbeh.2017.04.017

Bors, D. A., & Stokes, T. L. (1998). Raven's Advanced Progressive Matrices: Norms for First-Year University Students and the Development of a Short Form. *Educational and Psychological Measurement*, *58*(3), 382–398. https://doi.org/10.1177/0013164498058003002

Byrne, K. A., Otto, A. R., Pang, B., Patrick, C. J., & Worthy, D. A. (2019). Substance use is associated with reduced devaluation sensitivity. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(1), 40–55. https://doi.org/10.3758/s13415-018-0638-9

Chanraud, S., Martelli, C., Delain, F., Kostogianni, N., Douaud, G., Aubin, H.-J., Reynaud, M., & Martinot, J.-L. (2007). Brain Morphometry and Cognitive Performance in Detoxified Alcohol-Dependents with Preserved Psychosocial Functioning. *Neuropsychopharmacology*, *32*(2), Art. 2. https://doi.org/10.1038/sj.npp.1301219

Chen, H., Mojtahedzadeh, N., Belanger, M. J., Nebe, S., Kuitunen-Paul, S., Sebold, M., Garbusow, M., Huys, Q. J. M., Heinz, A., Rapp, M. A., & Smolka, M. N. (2021). Model-Based and Model-Free Control Predicts Alcohol Consumption Developmental Trajectory in Young Adults: A 3-Year Prospective Study. *Biological Psychiatry*, *89*(10), 980–989. https://doi.org/10.1016/j.biopsych.2021.01.009

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312–325. https://doi.org/10/f5ft22

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*(11), Art. 11. https://doi.org/10.1038/nn1579

First, M. B. (2016). *SCID-5-CV: Structured clinical interview for DSM-5 disorders, clinician version*. American Psychiatric Association.

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, *5*, e11305. https://doi.org/10.7554/eLife.11305

Hogarth, L. (2020). Addiction is driven by excessive goal-directed drug choice under negative affect: Translational critique of habit and compulsion theory. *Neuropsychopharmacology*, *45*(5), Art. 5. https://doi.org/10.1038/s41386-020-0600-8

Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging*, *8*(2), 207–220. https://doi.org/10.1037//0882-7974.8.2.207

Nagel, I. E., Chicherio, C., Li, S.-C., Von Oertzen, T., Sander, T., Villringer, A., Heekeren, H. R., Bäckman, L., & Lindenberger, U. (2008). Human aging magnifies genetic effects on executive functioning and working memory. *Frontiers in Human Neuroscience*, *2*. https://doi.org/10.3389/neuro.09.001.2008

Nebe, S., Kroemer, N. B., Schad, D. J., Bernhardt, N., Sebold, M., Müller, D. K., Scholl, L., Kuitunen-Paul, S., Heinz, A., Rapp, M. A., Huys, Q. J. M., & Smolka, M. N. (2018). No association of goal-directed and habitual control with alcohol consumption in young adults. *Addiction Biology*, *23*(1), 379–393. https://doi.org/10.1111/adb.12490

Otto, A. R., Gershman, S. J., Markman, A. B., & Daw, N. D. (2013). The curse of planning: Dissecting multiple reinforcement-learning systems by taxing the central executive. *Psychological science*, *24*(5), 751–761. https://doi.org/10.1177/0956797612463080

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2014). Cognitive Control Predicts Use of Model-based Reinforcement Learning. *Journal of Cognitive Neuroscience*, *27*(2), 319–333. https://doi.org/10/ggf8fw

Saunders, J. B., Aasland, O. G., Babor, T. F., De La Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, *88*(6), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Scheurich, A. (2005). Neuropsychological functioning and alcohol dependence. *Current Opinion in Psychiatry*, *18*(3), 319–323. https://doi.org/10.1097/01.yco.0000165602.36671.de

Sebold, M., Deserno, L., Nebe, S., Schad, D. J., Garbusow, M., Hägele, C., Keller, J., Jünger, E., Kathmann, N., Smolka, M., Rapp, M. A., Schlagenhauf, F., Heinz, A., & Huys, Q. J. M. (2014). Model-Based and Model-Free Decisions in Alcohol Dependence. *Neuropsychobiology*, *70*(2), 122–131. https://doi.org/10.1159/000362840

Sebold, M., Nebe, S., Garbusow, M., Guggenmos, M., Schad, D. J., Beck, A., Kuitunen-Paul, S., Sommer, C., Frank, R., Neu, P., Zimmermann, U. S., Rapp, M. A., Smolka, M. N., Huys, Q. J. M., Schlagenhauf, F., & Heinz, A. (2017). When Habits Are Dangerous: Alcohol Expectancies and Habitual Decision Making Predict Relapse in Alcohol Dependence. *Biological Psychiatry*, *82*(11), 847–856. https://doi.org/10.1016/j.biopsych.2017.04.019

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *298*(1089), 199–209. https://doi.org/10.1098/rstb.1982.0082

Steffen, J., Markovic, D., Glöckner, F., Neukam, P., Kiebel, S., Li, S.-C., & Smolka, M. (2022). *Shorter Planning Depth and Higher Response Noise During Sequential Decision-Making in Old Ag*. Research Square. https://doi.org/10.21203/rs.3.rs-2095779/v1

Stephan, R. A., Alhassoon, O. M., Allen, K. E., Wollman, S. C., Hall, M., Thomas, W. J., Gamboa, J. M., Kimmel, C., Stern, M., Sari, C., Dalenberg, C. J., Sorg, S. F., & Grant, I. (2017). Meta-analyses of clinical neuropsychological tests of executive dysfunction and impulsivity in alcohol use disorder. *The American Journal of Drug and Alcohol Abuse*, *43*(1), 24–43. https://doi.org/10.1080/00952990.2016.1206113

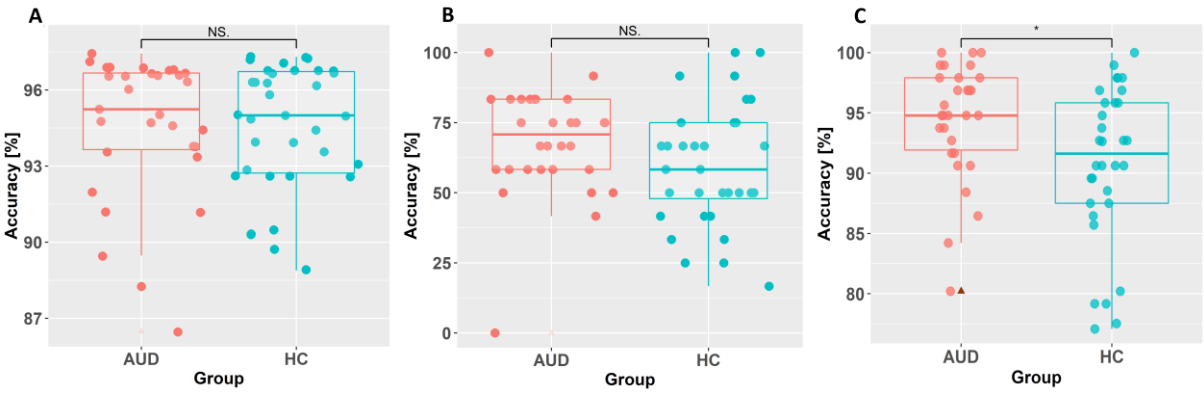
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2.). MIT press.

Voon, V., Derbyshire, K., Rück, C., Irvine, M. A., Worbe, Y., Enander, J., Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., & Bullmore, E. T. (2015). Disorders of compulsivity: A common bias towards learning habits. *Molecular Psychiatry*, *20*(3), Art. 3. https://doi.org/10.1038/mp.2014.44

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record, 58*, 475–482.

World Health Organization. (2018). *Global Status Report on Alcohol and Health 2018*. Poznyak, V. and Rekve, D. (Eds.) World Health Organization. https://www.who.int/publications/i/item/9789241565639

SUPPLEMENTAL INFORMATION



Supplemental Figure 1: Cognitive covariates. A: Identical Pictures Task. B: Raven task. C: Spatial working memory task, spatial memory accuracy.