The association of non-drug-related Pavlovian-to-instrumental transfer effect in nucleus accumbens with relapse in alcohol dependence: a replication

Ke Chen[a], Florian Schlagenhauf[a], Miriam Sebold[a,b], Sören Kuitunen-Paul[c,d], Hao Chen[e], Quentin J.M. Huys[f,g], Andreas Heinz[a], Michael N. Smolka[e], Ulrich S. Zimmermann[h,i], Maria Garbusow[a]

[a] Department of Psychiatry and Psychotherapy, Charité Campus Mitte, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[b] Department for Social and Preventive Medicine, University of Potsdam, Potsdam, Germany

[c] Department of Child and Adolescent Psychiatry, Technische Universität Dresden, Dresden, Germany

[d] Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany

[e] Department of Psychiatry, Technische Universität Dresden, Dresden, Germany

[f] Division of Psychiatry, University College London, London, United Kingdom

[g] Max Planck UCL Centre for Computational Psychiatry and Ageing Research, University College London, London, United Kingdom

[h] University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

[i] Department of Addiction Medicine and Psychotherapy, kbo Isar-Amper-Klinikum Region München, Germany

Corresponding author: Ke Chen; ke.chen@charite.de

Accepted Manuscript

Running title: PIT and relapse in alcohol dependence

Keywords: alcohol use disorder, Pavlovian-to-instrumental transfer, human cue-induced behavior, nucleus accumbens, relapse prediction, replication

Abstract

Background: The Pavlovian-to-instrumental transfer (PIT) paradigm measures the effects of Pavlovian conditioned cues on instrumental behavior in the laboratory. A previous study in our research group observed activity in the left nucleus accumbens (NAcc) elicited by a non-drug-related PIT task across alcohol-dependent (AD) patients and healthy controls, and the left NAcc PIT effect differentiated patients who subsequently relapsed from who remained abstinent. In the present study, we aimed to examine whether such effects were present in a larger subsequently collected sample.

Methods: A total of 129 recently detoxified AD patients (21 females) and 74 healthy, age- and sex-matched controls (12 females) performing a PIT task during functional magnetic resonance imaging (fMRI) were examined. After task assessments, patients were followed up for 6 months. Forty-seven patients relapsed and 37 remained abstinent.

Results: We found a significant behavioral non-drug-related PIT effect and PIT-related activity in the NAcc across all participants. Moreover, subsequent relapsers showed stronger behavioral and left NAcc PIT effects compared to abstainers. These findings are consistent with the previous findings.

Conclusions: Behavioral non-drug-related PIT and neural PIT correlates are associated with prospective relapse risk in alcohol dependence. This study replicated previous findings and provide evidence for the clinical relevance of PIT mechanisms with the treatment outcome in alcohol dependence. The observed difference between prospective relapsers and abstainers in NAcc PIT effect in our study is overall small. Future studies are needed to further elucidate the mechanisms and the possible modulators of neural PIT in relapse in alcohol dependence.

Accepted Manuscript

Introduction

Alcohol dependence is a prevalent disorder characterized by a high relapse rate (1, 2). The impact of cues on drug-seeking and drug-taking behavior has been hypothesized to be an important mechanism underlying relapse (3). According to the incentive salience sensitization theory of addiction, alcohol can induce the sensitization of incentive salience attribution to alcohol-predictive cues, promoting alcohol seeking and consumption despite one's intention to remain abstinent (4). The Pavlovian-to-instrumental transfer (PIT) paradigm has been established to experimentally measure the effects of reward-predictive cues on instrumental behaviors (e.g., 5, 6-8). In a PIT task, Pavlovian and instrumental training are first conducted separately, and then instrumental performance is assessed in the presence of Pavlovian conditioned stimuli (CSs) (9). The PIT effect refers to the promotion or inhibition effect of Pavlovian CSs on instrumental behavior (10). Studies further found that PIT effects come in two neurobiologically distinct forms (11-13). In the outcome-specific PIT, the Pavlovian CS associated with a reward enhances instrumental behavior leading to the same reward, whereas the general PIT refers to the situation when a CS enhances instrumental behavior regardless of the identity of the reward (11, 13).

Accepted Manuscript

Rodent studies found that drug-related cues enhance PIT effects in the drug-treated group (5, 14). For example, ethanol-related cues promoted not only ethanol seeking but also non-ethanol reward seeking (5). In addition to drug-related PIT, PIT tasks applying non-drug-related cues allows for studying a more general impact of drug use or addiction on cue-guided behavior. Enhanced non-drug-related PIT effects were observed in cocaine-exposed rats (15-18). Similarly, mice under chronic ethanol exposure showed enhanced non-ethanol-related PIT effects (19). These findings suggest a general alteration in motivational processes in drug-exposed animals. Comparably, our research group previously observed a more pronounced behavioral non-drug-related PIT effect in detoxified alcohol-dependent (AD) patients

compared to healthy controls (HCs) (20-22). Furthermore, the behavioral non-drug-related PIT effect was predictive of future relapse, with the evidence that prospective relapsers were less able to inhibit instrumental approach behavior when positive Pavlovian CSs were present (23).

Using neural imaging techniques, the nucleus accumbens (NAcc) has been identified as an essential neural substrate for PIT (18, 24-27). The NAcc is a part of the ventral striatum and is a core area of the human brain reward system (28). In the NAcc, a low availability of dopamine D_2 receptors has been associated with alcohol dependence (29). Few studies have thus far explored the neural mechanism of PIT in AD patients. To our best knowledge, Garbusow et al. (21) was the first study investigating the neural non-drug-related PIT in a clinical sample of AD patients after detoxification. In that study, functional activation of the left NAcc ($NAcc_L$) elicited by PIT was observed across AD patients and HCs. Moreover, the $NAcc_L$ PIT effect was stronger in subsequent relapsers (n =13) compared to abstainers (n = 11) (21).

These findings indicate a role of the non-drug-related PIT in predicting treatment outcomes in AD patients after detoxification. However, human studies showed heterogeneous findings on this account. In a recent study using a different PIT paradigm, researchers did not find differences between relapsing and abstaining AD patients regarding both the behavioral and neural PIT effects (30). Another human study also did not find significant differences in non-drug-related PIT between treatment-seeking drug users and controls (31). Besides, an animal study reported no categorical difference in non-drug-related PIT between cocaine addicted and non-addicted rats, but an association between the strength of the PIT effect and the amount of drug intake (32). These inconsistent findings underscore a need for further studies to elucidate the clinical relevance of PIT in alcohol dependence.

Accepted Manuscript

The aim of the present study is to examine whether the initial findings of non-drug-related PIT-induced activity in the NAcc, and a stronger NAcc PIT effect in prospective relapsers compared to abstainers from Garbusow et al. (21) can be replicated in a subsequently acquired larger sample. We analyzed a replication sample of 129 recently detoxified AD patients and 74 HCs who performed the PIT task in a functional magnetic resonance imaging (fMRI) scanner. On the behavioral level, following the finding of a group difference involving relapse in behavioral PIT in Sommer et al. (23), we asked if there is a stronger behavioral PIT effect in relapsers compared to abstainers with a 6-month follow-up. On the neural level, we hypothesized a stronger NAcc PIT effect in relapsers compared to abstainers. In addition to the replication sample, we conducted the analyses again with the full sample containing all the assessed participants (regardless of whether the data has been reported in Garbusow et al. (21)) in Supplement S12 and S13 to check if the findings are comparable in the full sample.

Accepted Manuscript

Methods and materials

Participants

The data were collected as part of the LeAD study (https://ssl.psych.tu-dresden.de/lead/; clinical trial numbers: NCT01679145 and NCT02615977). The replication sample described in the present paper is a subsample that was assessed after the initial sample reported in Garbusow et al. (21). The study was conducted in Berlin and Dresden, Germany, with approval by local ethics committees of Charité Universitätsmedizin Berlin (EA/1/157/11 and EA1/268/14) and Technische Universität Dresden (EK 228072012 and EK 300082014).

Participants performed the PIT task and other tasks (see clinical trial registration). The present study focused on the replication of the non-drug-related PIT task. Other data of the study are reported elsewhere (see https://ssl.psych.tu-dresden.de/lead/node/7 for an overview

of already reported data until 2019). Nevertheless, for comparison, we also report results of a drug (alcohol) versus water cue experiment in the supplement (S15 and S16). A total of 129 AD patients who fulfilled the criteria of alcohol dependence according to Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR) (33, 34), and 74 HCs were included into the final analyses of neural PIT as a replication sample after data cleaning (see Supplement S1 for study inclusion and exclusion procedure). Sample characteristics are presented in Table 1. AD patients were followed up for 6 months and had six or seven times of in-person or telephone interviews. Alcohol use was assessed using the timeline follow-back (TLFB) in each follow-up interview (35). Relapse was defined as ≥ five or four standard drinks (e.g., one standard drink = 0.33 L beer) that were consumed on one drinking occasion for males and females respectively. Forty-seven patients relapsed, while 37 remained abstinent. The other 45 patients had missing follow-up information and cannot be classified as an abstainer or a relapser. Sample characteristics of patients with known versus unknown relapse status, as well as participants had successful versus unsuccessful Pavlovian conditioning (see Supplement S2 and S7) were reported in Supplement S3.

[Insert Table 1]

Accepted Manuscript

PIT paradigm

The PIT paradigm has been described in the first study (21) and other previous publications (20, 22, 23). Participants conducted an instrumental task (i.e., collecting "good" shells via repeated button pressing or leaving "bad" shells via omitting a reaction) while monetary Pavlovian CSs were presented in the background. In addition, participants also performed trials in which alcohol or water cues instead of Pavlovian CSs were used (results are reported in Supplement S15 and S16). For a detailed description of the task, see Supplement S2.

MRI acquisition

At both study centers, scanning was performed using a Siemens Trio 3 Tesla MRI scanner. Details of MRI acquisition are reported in Supplement S4.

Data analysis

Data were analyzed using Matlab R2020b (41) and the R System for Statistical Computing Version 4.0.3 (42). FMRI data were analyzed using Statistical Parametric Mapping 12 (SPM12) (43).

Behavioral analyses

AD patients (n = 129) and HCs (n = 74) were included into analyses as a replication sample, with some of the participants (56 AD patients and 50 HCs) has been reported in another study regarding the behavioral non-drug-related PIT in Sommer et al. (23) using the measurement of accuracy. In the present paper, we analyzed the behavioral data with the measurement of number of button presses following Garbusow et al. (21) to compute the behavioral PIT effects in our analyzed sample.

We established Poisson distributed generalized linear mixed-effects models (GLMMs) with predictors of the associated monetary value of Pavlovian CSs (Pavlovian CS value: -2€, -1€, 0€, +1€, +2€) and the trial type of the instrumental condition (instrumental condition: +0.5 = go trial vs. -0.5 = no-go trial) to predict the trial-by-trial number of button presses in the transfer part. Participant IDs, instrumental stimuli (shells), and Pavlovian CSs (fractals combined with pure tones) were taken for random intercept effects in order to control for potential subject and item effects. For group comparison between AD patients and HCs, a group factor (+0.5 = AD patient vs. -0.5 = HC) as well as its interaction with other predictors were included as fixed effects. For the three-group comparisons between abstainers, relapsers and HCs, another GLMM was conducted with a three-level group factor. The analysis method currently used is different from that used in the first study, where the individual PIT effects

Accepted Manuscript

were first calculated by regressing the number of button presses on Pavlovian CS value and then subjected to group comparisons (21). We used this new analysis method because it is more sensitive for detecting small effects. We additionally explored whether the behavioral PIT effect was correlated with the severity of alcohol dependence, current alcohol craving, and family history of alcohol dependence (see Supplement S8).

Imaging analyses

Details of the imaging data pre-processing are reported in Supplement S5. After pre-processing, individual general linear models (GLM) were established for single-subject analyses (see Supplement Figure S3 for the GLM design matrix). Non-drug-related PIT trials were modelled as one condition with onset as the main regressor. Three parametric modulators for the main regressor were used: (1) Pavlovian CS value (-2€, -1€, 0€, +1€, +2€), (2) the number of button presses (log transformed calculated with $\ln(original\ number\ of\ button\ presses + e)$), and (3) the PIT parameter (transformed number of button presses × Pavlovian CS value). A higher number of button presses to a higher CS value would then lead to a higher numerical value in the PIT parameter. This GLM was adapted from the original one used in Garbusow et al. (21), where button presses corresponding to each Pavlovian CS were put into separate regressors. The new model congregated the PIT parameters corresponding to all Pavlovian CSs into one regressor so that the model would not easily fail in case there was no behavioral response variability.

Individual contrasts for the parametric modulator of non-drug-related PIT were calculated. To measure the neural PIT effect across participants, individual contrast images were subjected to a one-sample t-test in the second-level analysis in SPM, with participants' age, sex, study center, and Pavlovian training version (early vs. later version, see Supplement S2) as covariates. Following the previous study, a ROI analysis was conducted with a priori-defined compound ROI comprising the bilateral NAcc ($NAcc_L$ and $NAcc_R$) derived from the

Accepted Manuscript

Wake Forest University (WFU) PickAtlas software; http://www.fmri.wfubmc.edu/download.htm). In addition, explorative whole-brain analyses for the neural PIT effect on a significance level of $p_{unc} < .001$ and with $k \geq 20$ activated voxels per cluster was performed (see Supplement S9). Moreover, we examined whether the behavioral PIT effect correlated with the neural PIT effect in the NAcc by adding the extracted individual behavioral PIT slopes from another GLMM without the group factor as an additional covariate in the second-level analysis in SPM. For group comparisons, the mean beta values in the predefined NAcc ROI were extracted separately for $NAcc_L$ and $NAcc_R$. Wilcoxon rank sum test was conducted for two group comparisons and Kruskal–Wallis test for three-group comparisons for non-normally distributed data. We further explored the effects of excitatory and inhibitory Pavlovian CSs separately in Supplement S14. Analyses were also conducted for drug-related PIT trials (see Supplement S15 and S16).

In addition, we explored whether the neural PIT effect was correlated with alcohol dependence severity, current alcohol craving, and family history of alcohol dependence using spearman correlation test or Wilcoxon rank sum test (see Supplement S10). Finally, similar to the initial study (21), we conducted a logistic regression that measures if the behavioral and neural PIT effects were associated with relapse status in patients when controlling variables of alcohol dependence severity, alcohol craving and smoking status (see Supplement S11).

Accepted Manuscript

Results

Behavioral PIT

A significant behavioral PIT effect was present across groups, indicated by more button presses in trials with higher valued Pavlovian CSs (Pavlovian CS value: estimate = 0.28, z = 108.27, p < .001; see Table 2). AD patients displayed a stronger PIT effect compared to HCs (Pavlovian CS value × group interaction: estimate = 0.03, z = 5.21, p < .001).

[Insert Table 2]

When comparing the behavioral PIT effect among abstainers, relapsers and HCs (see Table 3), we observed a significant interaction of group and Pavlovian CS value (Chi-squared = 434.32, $p < .001$; Type II Wald chi-square tests for the GLMM). Post-hoc analyses showed that the PIT effect was strongest in relapsers, followed by HCs, and smallest in abstainers (relapsers > abstainers: estimate = 0.15, z = 20.24, $p < .001$; relapsers > HCs: estimate = 0.10, z = 15.39, $p < .001$; HCs > abstainers: estimate = .05, z = 7.24, $p < .001$; see Figure 1).

[Insert Table 3]

[Insert Figure 1]

Accepted Manuscript

NAcc BOLD signal elicited by PIT

Collapsing across groups, we observed a significant neural PIT effect in $NAcc_L$ (x = −10, y = 6, z = −8, $t_{(198)} = 3.34$, $p_{FWE\text{-}SVC} = .009$, voxel-based analysis; see Figure 2) while trend-wise significant in $NAcc_R$ (x = 6, y = 10, z = -10, $t_{(198)} = 2.66$, $p_{FWE\text{-}SVC} = .058$, voxel-based analysis). Furthermore, when including the individual behavioral PIT slopes into the second level GLM analysis, we found a significant association between the behavioral PIT and PIT-related activation in $NAcc_L$ (x = −8, y = 8, z = −12, $t_{(197)} = 3.60$, $p_{FWE\text{-}SVC} = .004$) but not in $NAcc_R$ (x = 6, y = 10, z = -10, $t_{(197)} = 2.44$, $p_{FWE\text{-}SVC} = .095$).

[Insert Figure 2]

The individually extracted mean beta values in the predefined ROI of $NAcc_L$ were then subjected to group comparisons. AD patients did not show different $NAcc_L$ PIT effect compared to HCs (Wilcoxon rank sum test; W = 4813, $p = .922$). When comparing abstainers, relapsers and HCs, a significant effect of group was observed (Kruskal-Wallis rank sum test; chi-squared = 6.27, $p = .044$, $\eta^2[H] = 0.03$). Post-hoc multiple comparisons showed a stronger effect in relapsers compared to abstainers (Dunn test with Bonferroni correction: z = 2.50, p

= .037), while no difference between abstainers and HCs (z = -1.65, p = .299), nor between relapsers and HCs (z = 1.17, p = .731), see Figure 3. In the full sample, there was no significant group difference when comparing three groups including HCs. However, the additionally conducted Wilcoxon rank sum test that replicated the two-group comparison analysis strategy in Garbusow et al. (21) still showed significantly stronger $NAcc_L$ PIT effect in relapsers compared to abstainers (see Supplement S13).

[Insert Figure 3]

Discussion

We observed a stronger behavioral non-drug-related PIT effect in prospective relapsers compared to abstainers and controls, which is consistent with Sommer et al. (23). More importantly, with a larger sample respective to the first study (21), we observed a neural PIT effect in the NAcc across participants in the present study, and $NAcc_L$ PIT effect was stronger in subsequent relapsers compared to abstainers. Overall, this study basically replicated the neural PIT findings in Garbusow et al. (21).

Given the abundant evidence implying an altered NAcc functioning in alcohol dependence, which could be due to drug effects on monoaminergic neurotransmission in this brain area (29), one would also expect different PIT-related NAcc activation induced by excessive alcohol intake. Indeed, it was found that young male, non-clinical adults with a high-risk drinking pattern displayed increased neural responses to PIT in the ventral striatum on a trend level (44). However, with clinical AD patients, we did not observe a different NAcc PIT effect compared to controls in both the first study (21) and the current one, unless we distinguished between prospective relapsers and abstainers. This indicates that instead of being a marker of alcohol dependence, the NAcc PIT effect perhaps has more importance in predicting an aspect of clinical severity, i.e., the propensity to relapse. The underlying mechanism could be that patients who are prone to the influence of environmental cues (i.e., a

Accepted Manuscript

stronger PIT effect) may be inclined to alcohol intake in environments associated with alcohol intake or certain mood states (45). Indeed, an animal study observed that mice with higher food-related behavioral PIT effects had stronger subsequent cue-induced reinstatement of alcohol seeking (46). Consistently, the behavioral PIT effect differentiated relapsers and abstainers in Sommer et al. (23) and in the present study. Accordingly, and in line with Garbusow et al. (21), we observed a higher $NAcc_L$ PIT effect in relapsers compared to abstainers across different follow-up periods (3 months in the first study and 6 months in the current one). Surprisingly, the group difference in neural PIT effect was not as stable as expected when examining the effect in the full sample (see Supplement S13). The additionally conducted two-group comparison, however, still indicated a stronger $NAcc_L$ PIT effect in relapsers compared to abstainers. We suspect that the insignificant three-group difference in the full sample may be partly due to a sample effect in controls in PIT performance. Indeed, the variance among controls in behavioral PIT differs, with stronger behavioral PIT effects in controls in the replication sample compared to controls reported previously (Supplement S12).

Accepted Manuscript

The difference in functional activation elicited by non-drug-related PIT between relapsers and abstainers consistently showed in $NAcc_L$, both in the first study and the present one. Previous research suggested lateralized dopamine release to US and CS in the NAcc, with dopamine release in the $NAcc_L$ mostly reflects alcohol intake (US, intoxication), while dopamine release in the $NAcc_R$ mostly reflects the drink related CSs (i.e., beer flavor) in male heavy drinkers (47). The finding in our study may underline the significance of $NAcc_L$ in relapse to alcohol intake (21). Nevertheless, our study was not designed to investigate the hemispheric difference of NAcc. In drug-related PIT, group differences between relapsers and abstainers were conversely found in $NAcc_R$ rather than $NAcc_L$ (48; and Supplement S16). Further research is needed to elucidate the roles of $NAcc_L$ and $NAcc_R$ in different PIT tasks.

Several earlier findings did not support an association between PIT and addiction (e.g., 30, 31, 32, 49). The inconsistent findings question categorical differences in cue reactivity. Therefore, replication studies are important and should include assessments of clinical severity that could reflect differences in drug effects on the ventral and dorsal striatum (50). The inconsistent findings in human studies may be explained by differences in terms of samples characteristics, PIT manipulations, and sample sizes. For example, findings in one study were based on social drinkers rather than AD patients and with different types of reward (49). There, no association between hazardous drinking and PIT was observed using beer points as the outcome (cover story, no beer was provided after the PIT session) (49). Another study investigating diagnosed AD patients found no association of behavioral or neural PIT neither with alcohol dependence status nor with treatment outcome (30). That study used food outcomes (participants were allowed to eat the earned snack at the end of the experiment) rather than monetary outcomes as in our task. We speculate that the type of reward used for the conditioning and the approach of providing the reward has an impact on the resulting experimental behavior.

Accepted Manuscript

The observed group difference in the activation of NAcc associated with our non-drug-related PIT task is overall small (η^2[H] = 0.03 for three-group comparison in the replication sample and r = 0.21 for relapers versus abstainers comparison in the full sample), indicating that neural PIT cannot thoroughly elucidate the mechanisms of relapse in alcohol dependence. Indeed, relapse in alcohol dependence has multifactorial causes that vary from person to person and within each individual (51), and other mechanisms might interact with PIT process and relapse (52). It is also worth noting that behavioral PIT could be more efficient in relapse prediction than neural PIT in our study, as the logistic regression comprising multiple predictors of $NAcc_L$ PIT, behavioral PIT, alcohol craving, alcohol dependence severity, and smoking status yield only significant effect of behavioral PIT in

predicting relapse (Supplement S11). Future studies are warranted to further elucidate the mechanisms and possible modulators of neural PIT in relapse in alcohol dependence and to translate neurobiological findings to treatment of alcohol dependence.

There are limitations in this study. First, although we used the identical procedure as in the first study, the present replication study was not preregistered except clinical trials registration. Second, patients compared to controls, and relapsers compared to abstainers, showed less changed button pressing responses to instrumental go versus no-go trials along the instrumental training (Supplement S6) in our study. The group difference in behavioral PIT may partly be explained by differences in learned instrumental response – outcome contingency, as previous research observed a larger PIT effect when the instrumental response – outcome contingency was less reliable (53). However, adding the instrumental learning slope as a covariate in the behavioral PIT GLMM model does not change the significance of group differences in PIT effects (Supplement S6), indicating that the observed group differences in PIT cannot be fully explained by the difference in instrumental learning. Third, a part of participants (60 patients and 20 HCs) in our study conducted a cognitive bias modification training after the PIT task reported here, which we hypothesized to be effective to reduce relapse risk. However, we did not observe such an effect (results will be reported elsewhere). In fact, the proportion of relapsers did not differ significantly between patients who underwent verum training (14/21; 67%), placebo training (9/19; 47%), or no training (24/44; 55%) in the currently analyzed sample ($\chi^2 = 1.58$, $p = .453$; Cramer's $V = 0.14$). Therefore, we believe the training did not confound our findings. Fourth, we slightly changed the design of Pavlovian training (an interstimulus interval between the presentation of CS and US was used in the first but not the second version of training; see Supplement S2) during the study. Although it can lead to different training efficacy, we argue that it does not impact the PIT findings, because participants who did not successfully learn the association of CSs and

Accepted Manuscript

USs (11% AD patients and 9% HCs) were not included in the analyses. Moreover, the version of Pavlovian training was included as a covariate in the neural PIT analysis to eliminate potential confounding effect. However, patients who had unsuccessful Pavlovian conditioning also had more life-time drinking than those had successful Pavlovian conditioning (Supplement S3), indicating an altered associative learning induced by prolonged alcohol intake (54, 55). Excluding patients with unsuccessful Pavlovian conditioning potentially limits the PIT assessment on patients with less alcohol intake in our study. Fifth, the non-drug-related PIT task in our study could be contributed by both outcome-specific and general PIT, as monetary outcomes were used in both instrumental and Pavlovian training, but with different values. Research suggested that outcome-specific and general PIT effects depend on NAcc shell and core respectively (12). Due to limitations in spatial resolution and smoothing, we cannot distinguish between NAcc core and shell in this study. Last, this study was conducted by the same research group as the first study. Replication studies conducted by independent investigators and institutions are needed to further reduce potential biases.

In conclusion, this study replicated previous findings of a stronger behavioral non-drug-related PIT and PIT-related activation in $NAcc_L$ in relapsing AD patients compared to abstaining patients. The findings suggest that behavioral and $NAcc_L$ PIT may be related to the vulnerability to relapse in AD patients after detoxification. Future studies are needed to further elucidate the mechanisms and the possible modulators of neural PIT in relapse in alcohol dependence. In addition, to confirm clinical relevance, in addition to replication studies, further research is required to generalize the finding across variations of samples and measurement conditions (56).

Accepted Manuscript

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (project numbers 402170461 [TRR 265] to FS, MS, AH, MNS and MG); 186318919 [FOR 1617] to FS, SKP, QJMH, AH, MNS, USZ and MG; 178833530 [SFB 940] to MNS) and the China Scholarship Council (CSC grant 201806750014 to KC).

Accepted Manuscript

Disclosures

QJMH has received consultancy fees and options from Aya Health and a research grant from Koa Health. Other coauthors (KC, FS, MS, SKP, HC, AH, MNS, USZ and MG) report no biomedical financial interests or potential conflicts of interest.

Accepted Manuscript

Reference

1. Boothby LA, Doering PL (2005): Acamprosate for the treatment of alcohol dependence. *Clin Ther*. 27:695-714.

2. Walter M, Dürsteler KM, Petitjean SA, Wiesbeck GA, Euler S, Sollberger D, et al. (2015): [Psychosocial Treatment of Addictive Disorders--An Overview of Psychotherapeutic Options and their Efficacy]. *Fortschr Neurol Psychiatr*. 83:201-210.

3. Perry CJ, Zbukvic I, Kim JH, Lawrence AJ (2014): Role of cues and contexts on drug-seeking behaviour. *Br J Pharmacol*. 171:4636-4672.

4. Cofresí RU, Bartholow BD, Piasecki TM (2019): Evidence for incentive salience sensitization as a pathway to alcohol use disorder. *Neurosci Biobehav Rev*. 107:897-926.

5. Corbit LH, Janak PH (2007): Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcohol Clin Exp Res*. 31:766-774.

6. Glasner SV, Overmier JB, Balleine BW (2005): The role of Pavlovian cues in alcohol seeking in dependent and nondependent rats. *J Stud Alcohol*. 66:53-61.

7. Everitt BJ, Robbins TW (2005): Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 8:1481-1489.

8. Corbit LH, Janak PH (2007): Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *J Neurosci*. 27:13977-13981.

9. Cartoni E, Balleine B, Baldassarre G (2016): Appetitive Pavlovian-instrumental Transfer: A review. *Neurosci Biobehav Rev*. 71:829-848.

10. Huys QJ, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, et al. (2011): Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. *PLoS Comput Biol*. 7:e1002028.

Accepted Manuscript

11. Corbit LH, Balleine BW (2005): Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci*. 25:962-970.

12. Corbit LH, Balleine BW (2011): The general and outcome-specific forms of Pavlovian-instrumental transfer are differentially mediated by the nucleus accumbens core and shell. *J Neurosci*. 31:11786-11794.

13. Cartoni E, Puglisi-Allegra S, Baldassarre G (2013): The three principles of action: a Pavlovian-instrumental transfer hypothesis. *Front Behav Neurosci*. 7:153.

14. LeBlanc KH, Ostlund SB, Maidment NT (2012): Pavlovian-to-instrumental transfer in cocaine seeking rats. *Behav Neurosci*. 126:681-689.

15. LeBlanc KH, Maidment NT, Ostlund SB (2013): Repeated cocaine exposure facilitates the expression of incentive motivation and induces habitual control in rats. *PLoS One*. 8:e61355.

16. LeBlanc KH, Maidment NT, Ostlund SB (2014): Impact of repeated intravenous cocaine administration on incentive motivation depends on mode of drug delivery. *Addict Biol*. 19:965-971.

17. Ostlund SB, LeBlanc KH, Kosheleff AR, Wassum KM, Maidment NT (2014): Phasic mesolimbic dopamine signaling encodes the facilitation of incentive motivation produced by repeated cocaine exposure. *Neuropsychopharmacology*. 39:2441-2449.

18. Saddoris MP, Stamatakis A, Carelli RM (2011): Neural correlates of Pavlovian-to-instrumental transfer in the nucleus accumbens shell are selectively potentiated following cocaine self-administration. *Eur J Neurosci*. 33:2274-2287.

19. Shields CN, Gremel CM (2021): Prior chronic alcohol exposure enhances Pavlovian-to-instrumental transfer. *Alcohol*. 96:83-92.

Accepted Manuscript

20. Garbusow M, Schad DJ, Sommer C, Jünger E, Sebold M, Friedel E, et al. (2014): Pavlovian-to-instrumental transfer in alcohol dependence: a pilot study. *Neuropsychobiology*. 70:111-121.

21. Garbusow M, Schad DJ, Sebold M, Friedel E, Bernhardt N, Koch SP, et al. (2016): Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addict Biol*. 21:719-731.

22. Sommer C, Garbusow M, Jünger E, Pooseh S, Bernhardt N, Birkenstock J, et al. (2017): Strong seduction: impulsivity and the impact of contextual cues on instrumental behavior in alcohol dependence. *Transl Psychiatry*. 7:e1183.

23. Sommer C, Birkenstock J, Garbusow M, Obst E, Schad DJ, Bernhardt N, et al. (2020): Dysfunctional approach behavior triggered by alcohol-unrelated Pavlovian cues predicts long-term relapse in alcohol dependence. *Addict Biol*. 25:e12703.

24. Corbit LH, Muir JL, Balleine BW (2001): The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J Neurosci*. 21:3251-3260.

25. Peciña S, Berridge KC (2013): Dopamine or opioid stimulation of nucleus accumbens similarly amplify cue-triggered 'wanting' for reward: entire core and medial shell mapped as substrates for PIT enhancement. *Eur J Neurosci*. 37:1529-1540.

26. Talmi D, Seymour B, Dayan P, Dolan RJ (2008): Human pavlovian-instrumental transfer. *J Neurosci*. 28:360-368.

27. Geurts DE, Huys QJ, den Ouden HE, Cools R (2013): Aversive Pavlovian control of instrumental behavior in humans. *J Cogn Neurosci*. 25:1428-1441.

28. Breiter HC, Aharon I, Kahneman D, Dale A, Shizgal P (2001): Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*. 30:619-639.

Accepted Manuscript

29. Heinz A, Siessmeier T, Wrase J, Hermann D, Klein S, Grüsser SM, et al. (2004): Correlation between dopamine D(2) receptors in the ventral striatum and central processing of alcohol cues and craving. *Am J Psychiatry*. 161:1783-1789.

30. van Timmeren T, Quail SL, Balleine BW, Geurts DEM, Goudriaan AE, van Holst RJ (2020): Intact corticostriatal control of goal-directed action in Alcohol Use Disorder: a Pavlovian-to-instrumental transfer and outcome-devaluation study. *Sci Rep*. 10:4949.

31. Hogarth L, Lam-Cassettari C, Pacitti H, Currah T, Mahlberg J, Hartley L, et al. (2019): Intact goal-directed control in treatment-seeking drug users indexed by outcome-devaluation and Pavlovian to instrumental transfer: critique of habit theory. *Eur J Neurosci*. 50:2513-2525.

32. Takahashi TT, Vengeliene V, Enkel T, Reithofer S, Spanagel R (2019): Pavlovian to Instrumental Transfer Responses Do Not Correlate With Addiction-Like Behavior in Rats. *Front Behav Neurosci*. 13:129.

33. Wittchen H-U, Pfister H (1997): DIA-X-Interviews: Manual für Screening-Verfahren und Interview; Interviewheft.

34. Jacobi F, Mack S, Gerschler A, Scholl L, Höfler M, Siegert J, et al. (2013): The design and methods of the mental health module in the German Health Interview and Examination Survey for Adults (DEGS1-MH). *Int J Methods Psychiatr Res*. 22:83-99.

35. Sobell L, Sobell M (1992): Timeline follow-back measuring alcohol consumption. *Center for Psychological Studies, Nova Southeastern University*.41-72.

36. Skinner HA, Horn JL (1984): *Alcohol dependence scale (ADS): User's guide*. Addiction Research Foundation.

37. Rice JP, Reich T, Bucholz KK, Neuman RJ, Fishman R, Rochberg N, et al. (1995): Comparison of direct interview and family history diagnoses of alcohol dependence. *Alcohol Clin Exp Res*. 19:1018-1023.

Accepted Manuscript

38. Meule A, Vögele C, Kübler A (2011): Psychometrische evaluation der deutschen Barratt impulsiveness scale–Kurzversion (BIS-15). *Diagnostica*.

39. Mann K, Ackermann K (2000): Die OCDS-G: Psychometrische Kennwerte der deutschen Version der obsessive compulsive drinking scale. *Sucht*. 46:90-100.

40. Herrmann-Lingen C, Buss U, Snaith R (1995): Hospital anxiety and depression scale–Deutsche Version. *Testdokumentation und Handanweisung Verlag Hans Huber, Bern*.1-41.

41. MATLAB (2020): version 9.9. Natick, Massachusetts: The MathWorks Inc.

42. R Core Team (2020): R: A language and environment for statistical computing. . Vienna, Austria.: R Foundation for Statistical Computing.

43. Ashburner J, Barnes G, Chen C-C, Daunizeau J, Flandin G, Friston K, et al. (2014): SPM12 manual. *Wellcome Trust Centre for Neuroimaging, London, UK*. 2464:4.

44. Chen H, Nebe S, Mojtahedzadeh N, Kuitunen-Paul S, Garbusow M, Schad DJ, et al. (2020): Susceptibility to interference between Pavlovian and instrumental control is associated with early hazardous alcohol use. *Addict Biol*.e12983.

45. Heinz A, Löber S, Georgi A, Wrase J, Hermann D, Rey ER, et al. (2003): Reward craving and withdrawal relief craving: assessment of different motivational pathways to alcohol intake. *Alcohol Alcohol*. 38:35-39.

46. Barker JM, Torregrossa MM, Taylor JR (2012): Low prefrontal PSA-NCAM confers risk for alcoholism-related behavior. *Nat Neurosci*. 15:1356-1358.

47. Oberlin BG, Dzemidzic M, Tran SM, Soeurt CM, O'Connor SJ, Yoder KK, et al. (2015): Beer self-administration provokes lateralized nucleus accumbens dopamine release in male heavy drinkers. *Psychopharmacology (Berl)*. 232:861-870.

48. Schad DJ, Garbusow M, Friedel E, Sommer C, Sebold M, Hägele C, et al. (2019): Neural correlates of instrumental responding in the context of alcohol-related cues index disorder severity and relapse risk. *Eur Arch Psychiatry Clin Neurosci*. 269:295-308.

Accepted Manuscript

49. Martinovic J, Jones A, Christiansen P, Rose AK, Hogarth L, Field M (2014): Electrophysiological responses to alcohol cues are not associated with Pavlovian-to-instrumental transfer in social drinkers. *PLoS One*. 9:e94605.

50. Volkow ND, Wiers CE, Shokri-Kojori E, Tomasi D, Wang GJ, Baler R (2017): Neurochemical and metabolic effects of acute and chronic alcohol in the human brain: Studies with positron emission tomography. *Neuropharmacology*. 122:175-188.

51. Witkiewitz K, Marlatt GA (2004): Relapse prevention for alcohol and drug problems: that was Zen, this is Tao. *Am Psychol*. 59:224-235.

52. Sebold M, Garbusow M, Cerci D, Chen K, Sommer C, Huys QJ, et al. (2021): Association of the OPRM1 A118G polymorphism and Pavlovian-to-instrumental transfer: Clinical relevance for alcohol dependence. *J Psychopharmacol*.269881121991992.

53. Cartoni E, Moretta T, Puglisi-Allegra S, Cabib S, Baldassarre G (2015): The Relationship Between Specific Pavlovian Instrumental Transfer and Instrumental Reward Probability. *Front Psychol*. 6:1697.

54. Doñamayor N, Ebrahimi C, Garbusow M, Wedemeyer F, Schlagenhauf F, Heinz A (2021): Instrumental and pavlovian mechanisms in alcohol use disorder. *Current Addiction Reports*. 8:156-180.

55. Heinz A, Beck A, Halil MG, Pilhatsch M, Smolka MN, Liu S (2019): Addiction as Learned Behavior Patterns. *J Clin Med*. 8.

56. Lerma-Usabiaga G, Mukherjee P, Ren Z, Perry ML, Wandell BA (2019): Replication and generalization in applied neuroimaging. *Neuroimage*. 202:116048.

Accepted Manuscript

Figure/Table Legends

Figure 1. Behavioral PIT effect in relapsers (n = 47), abstainers (n = 37) and healthy controls (n = 74).

The behavioral PIT effect was strongest (steepest slope) in subsequently relapsed alcohol-dependent patients (AD-relapsers), followed by healthy controls (HCs), and smallest in abstinent patients (AD-abstainers). (A) shows the original number of button presses to each Pavlovian CS value; (B) shows the number of button presses relative to zero CS value. Group mean and standard error of the mean were shown with bars and error bars.

Figure 2: Neural PIT effect across all participants (n = 203).

Bilateral NAcc ROI (blue) and functional PIT activation (yellow; $p_{unc} < .005$ was used for illustration in the figure while $p_{FWE\text{-}SVC} < .05$ was used for data analyses).

Figure 3. Mean beta values in the left NAcc in subsequent relapsers (n = 47), abstainers (n = 37) and healthy controls (HCs; n = 74).

Group mean and standard error of the mean were shown with bars and error bars, and individual values were represented by colored dots. Relapsers showed a higher left NAcc PIT effect compared to abstainers. *: $p < .05$

Table 1. Sample characteristics and test statistics comparing alcohol-dependent (AD) patients (n = 129) to healthy controls (HCs) (n = 74), and comparing relapsers (n = 47) to abstainers (n =37).

Note. [a] Chi-square test; [b] Wilcoxon rank sum test; [c] t-test. ADS = Alcohol Dependence Scale, sum score with greater values indicating more severe alcohol dependence (36); FHAM: Family History Assessment Module (37); BIS-15 = Short German version of the Barrat Impulsiveness Scale-15, sum score with greater values indicating stronger trait impulsivity

Accepted Manuscript

(38); OCDS-G = German version of the Obsessive Compulsive Drinking Scale, sum score with greater values indicating stronger craving for alcohol within 7 days before assessment (39); HADS = Hospital Anxiety and Depression Scale, sum scores with greater values indicating stronger anxiety/depressivity within 7 days before assessment (40). *Three control subjects abstained from alcohol longer than two standard deviations above the mean. The median (1st quartile – 3rd quartile) of abstinence from alcohol in the control group is 4 (2 – 14) days.

Table 2. Behavioral PIT in AD patients (n = 129) and HCs (n = 74).

Note. AD: alcohol-dependent; CS: conditioned stimulus; HC: healthy control; PIT: Pavlovian-to-instrumental transfer.

Table 3. Behavioral PIT in relapsers (n = 47), abstainers (n = 37) and HCs (n = 74).

Note. CS: conditioned stimulus; HC: healthy control; PIT: Pavlovian-to-instrumental transfer.

Accepted Manuscript

Table 1.

	129 AD patients	*74 HCs*	*p*	*47 relapsers*	*37 abstainers*	*p*
Sex (female)	n = 21 (16%)	n = 12 (16%)	.99[a]	n = 7 (15%)	n = 6 (16%)	.87[a]
Sample characteristics	M (SD)	M (SD)		M (SD)	M (SD)	
Age	44.3 (9.9)	44.1 (10.8)	.92[b]	44.8 (9.9)	44.9 (10.5)	.97[c]
Education (years)	15.3 (4.1)	15.9 (3.4)	.09[b]	15.7 (4.3)	14.7 (3.8)	.27[b]
Smokers (%)	77%	72%	.41[a]	67%	82%	.12[a]
AD severity (ADS)	16.7 (7.4)	1.9 (2.9)	< .001[b]	16.2 (7.0)	16.9 (7.4)	.66[c]
With family history of alcohol dependence (FHAM)	39%	12%	< .001[a]	33%	47%	.18[a]
Time since last alcoholic drink (days)	22.1 (12.4)	75.0 (310.2)*	< .001[b]	22.1 (13.7)	18.9 (7.2)	.68[b]
Alcohol intake per day in past year (g of pure ethanol)	162 (134)	9.6 (10.6)	< .001[b]	148 (94.4)	149 (121)	.76[b]
Lifetime alcohol intake (kg of pure ethanol)	1728 (1291)	314 (936)	< .001[b]	1834 (1390)	1859 (1265)	.80[b]
Craving for alcohol (OCDS-G)	12.6 (7.9)	2.8 (2.6)	<.001[b]	12.0 (7.3)	12.9 (8.9)	.78[b]
Trait impulsivity (BIS-15)	31.5 (6.2)	29.8 (5.3)	.03[c]	31.6 (6.4)	32.2 (5.9)	.68[b]

Accepted Manuscript

Current anxiety (HADS)	4.3 (3.2)	1.9 (1.9)	< .001[b]	4.7 (2.9)	4.2 (3.4)	.29[b]
Current depressivity (HADS)	3.6 (3.7)	1.2 (1.8)	< .001[b]	3.8 (3.3)	4.3 (4.5)	.74[b]

Accepted Manuscript

Table 2.

Parameter	***Estimate (SE)***	***z***	***P***
Intercept	1.42 (0.04)	37.07	< .001
Pavlovian CS value	0.28 (0.003)	108.27	< .001
Instrumental condition (go vs. no-go)	0.59 (0.05)	10.91	< .001
Group (AD patient vs. HC)	0.10 (0.05)	1.87	.061
Pavlovian CS value × group	0.03 (0.005)	5.21	< .001
Instrumental condition × group	-0.23 (0.01)	-15.88	< .001

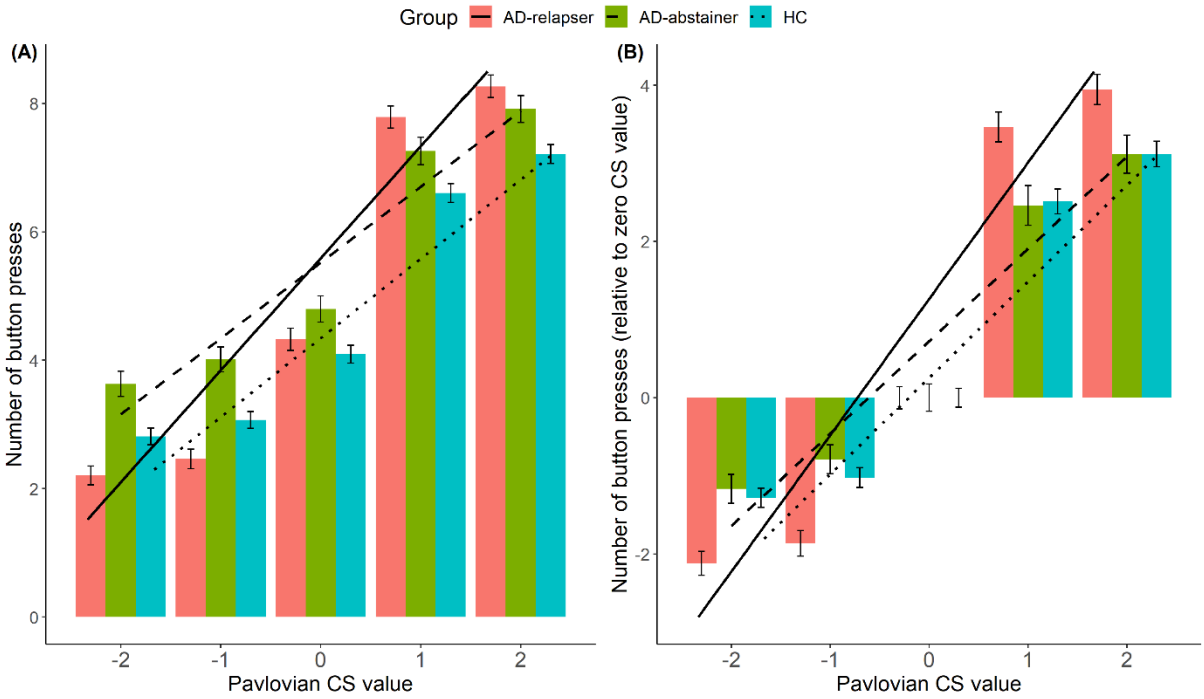
Accepted Manuscript

Table 3.

Parameter	***Estimate (SE)***	***z***	***P***
Intercept	1.59 (0.06)	24.58	< .001
Pavlovian CS value	0.22 (0.005)	40.29	< .001
Instrumental condition (go vs. no-go)	0.50 (0.06)	8.86	< .001
Group (relapser vs. abstainer)	-0.20 (0.08)	-2.56	.010
Group (HC vs. abstainer)	-0.22 (0.07)	-3.14	.002
Pavlovian CS value × group (relapser vs. abstainer)	0.15 (0.008)	20.24	< .001
Pavlovian CS value × group (HC vs. abstainer)	0.05 (.007)	7.24	< .001
Instrumental condition × group (relapser vs. abstainer)	-0.05 (0.02)	-2.52	.012
Instrumental condition × group (HC vs. abstainer)	0.21 (0.02)	10.78	< .001

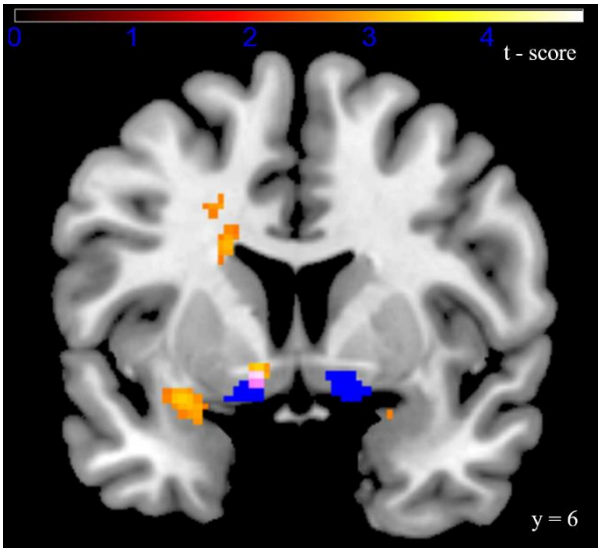
Accepted Manuscript

Figure 1.



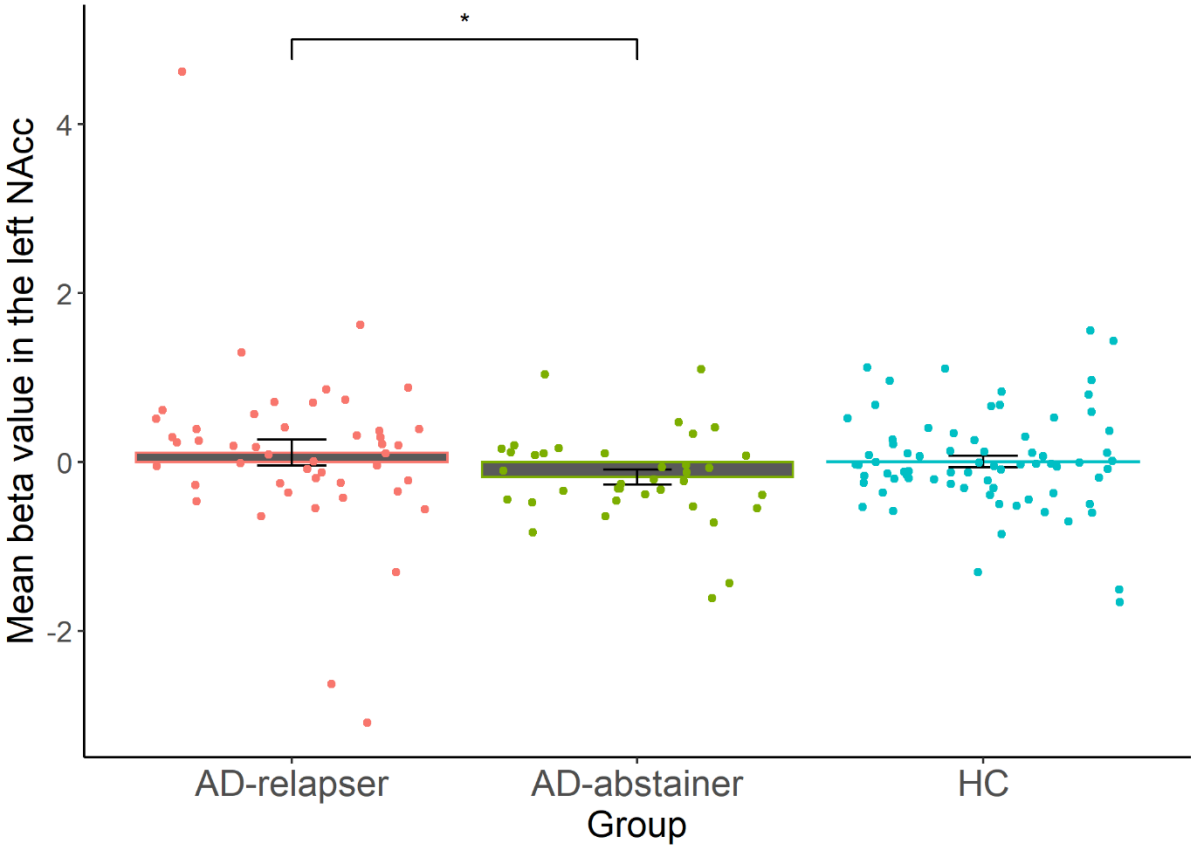
Accepted Manus

Figure 2.



Accepted Manuscript

Figure 3.



Accepted M