Counterconditioning Alcohol Cues: Neural and Behavioral Modulation of Automatic Tendencies and Pavlovian-to-Instrumental Transfer in Male Alcohol Users

Adarsh K. Verma[1*], Usha Chivukula[1], Maria Garbusow[2], Neeraj Kumar[3*]

[1]Department of Psychology, University of Hyderabad, India; [2]Department of Psychology, Medical School Berlin, Germany; [3]Department of Liberal Arts, Indian Institute of Technology Hyderabad, India

*Corresponding Author

Email: adars.du@gmail.com (AKV); neeraj.kumar@la.iith.ac.in (NK)

DECLARATIONS

Conflict of Interest: None

Data Availability: The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgement: This work was supported by the Ministry of Education, Government of India, through the Prime Minister's Research Fellowship (PMRF) awarded to AKV.

Abstract

Alcohol cues can influence ongoing instrumental behavior through Pavlovian-to-instrumental transfer (PIT), possibly promoting problematic alcohol consumption. This influence can be explained by automatic approach tendencies towards alcohol, which play a role in understanding alcohol abuse and relapse. However, individuals differ in these tendencies; some exhibit a strong approach, while others exhibit avoidance; yet, how such profiles shape PIT and intervention outcomes remains unclear. Existing interventions predominantly target approach tendencies, often overlooking avoidance profiles that may represent distinct motivational control. Building on evidence that cue-evoked responding reflects learned expectancies about response effectiveness, we investigated whether counterconditioning (CC), i.e., reversing alcohol cue valence by pairing it with an aversive outcome, could reduce approach tendencies and attenuate alcohol-related PIT effects across both approach and avoidance groups. To enhance interventional effect transfer, we integrated retrieval cues to facilitate memory reactivation in non-treatment contexts. Thirty-nine alcohol users completed the Alcohol Approach-Avoidance Task (A-AAT) and PIT task before and after CC, with participants categorized by their baseline Alcohol Approach Index. EEG recorded neural responses, focusing on N2 and P3 ERPs as indices of conflict monitoring and cognitive control. Individuals with approach tendencies showed stronger PIT interference during avoidance-related actions in the presence of alcohol cues. CC attenuated these tendencies and restored instrumental responding, with retrieval cues supporting transfer effects. Post-intervention, enhanced frontal N2 and centro-parietal P3 amplitudes reflected improved conflict monitoring and top-down control. These findings demonstrate that retrieval cue-integrated CC can recalibrate automatic approach tendencies and restore adaptive instrumental behavior, with corresponding neural enhancements, offering a practical intervention to reduce cue-driven drinking.

Keywords: Counterconditioning, alcohol approach tendencies, Pavlovian-to-Instrumental Transfer (PIT), event-related potentials

Introduction

Alcohol dependence is a complex process, often initiated by repeated pairings of alcohol cues with positive experiences as a short-term consequence of alcohol intake, such as euphoria (Drummond et al., 1990; Lutz & Childs, 2021). Over time, these associations become robust and manifest as automatic tendencies that influence behavior through fast, implicit processes operating below conscious awareness. When individuals encounter alcohol-related stimuli in everyday contexts, such as passing a wine shop while commuting to work, automatic responses can interfere with intended actions and undermine adaptive behavioral control, even in those motivated to abstain (Abiero et al., 2022; Campanella et al., 2019; Cofresí et al., 2022; Field et al., 2008; Garbusow et al., 2022; Wiers & van Dessel, 2023). One mechanism through which alcohol cues exert such influence is Pavlovian-to-instrumental transfer (PIT), wherein conditioned stimuli modulate ongoing instrumental behavior (Hogarth et al., 2014, 2019, Garbusow).

Debate continues as to whether PIT reflects habitual responding or goal-directed control (Hogarth et al., 2019; Hogarth & Chase, 2011; Seabrooke et al., 2019; Watson et al., 2014). Some evidence suggests that cue-elicited instrumental responding can operate independently of outcome value, consistent with habitual control (Hogarth & Chase, 2011; Watson et al., 2014), whereas other studies demonstrate that PIT effects are sensitive to outcome devaluation, supporting goal-directed mechanisms (Hogarth et al., 2019; Seabrooke et al., 2019). However, these investigations have primarily focused on group-level effects, treating the absence of goal-directed control as a habitual behavior rather than examining how individual differences in automatic response tendencies might differentially contribute to PIT effects. Indeed, considerable heterogeneity exists in how individuals respond to alcohol cues, with some showing strong approach tendencies while others exhibit avoidance patterns despite continued drinking (Kenney et al., 2018; McNaughton et al., 2016; Morris et al., 2020; Verma et al., 2025). Whether these opposing automatic tendencies produce distinct patterns of cue-driven interference with instrumental responding, and whether these tendencies modulate habitual versus goal-directed processes, remains unexplored. Hence, it is crucial to understand how automatic tendencies influence instrumental responding and how these cue-driven processes can be targeted through interventions.

Building on this perspective, interventions have been developed to either reduce automatic approach biases or to diminish the impact of alcohol cues on instrumental behavior. These include behavioral extinction (Bouton, 2002; Torregrossa & Taylor, 2012; Wiers et al., 2011), cognitive bias modification (CBM; Cristea et al., 2016; Rinck et al., 2018), avoidance training (Di Lemma et al., 2020; Rosenthal et al., 2023; Wiers et al., 2013), and neuromodulation (den Uyl et al., 2016). For example, Rosenthal et al. (2023) reported that alcohol cue-avoidance training did not reliably resolve alcohol cue-induced interference. These interventions show important limitations. Their efficacy remains mixed, with effects often being short-lived, and transfer to a non-interventional context has been limited (Cristea et al., 2016; Rinck et al., 2018; Van Gucht, Van den Bergh, et al., 2010; Wiers et al., 2013). A key reason behind the limited success is the context dependency of these interventions, where benefits observed in controlled settings do not generalize to real-world drinking situations. This underscores the need for interventions that address both the reduction of maladaptive automatic tendencies and the restoration of goal-directed control, while also accounting for the context-dependency of interventional effects.

Based on principles of associative learning, counterconditioning (CC) represents a promising approach for directly targeting the motivational significance of alcohol cues while targeting both cue-driven processes. Unlike interventions that focus on strengthening cognitive control, CC operates by creating competing associative memories that can diminish the automatic influence of alcohol cues on behavior (Baker & Cannon, 1979; Das et al., 2015; Elkins, 1980; Gale et al., 2021; Goltseker et al., 2016, 2021; Wiens et al., 1976). Evidence suggests that cue-evoked drug-seeking is sustained by hierarchical instrumental expectancies, the belief that a particular response will be more effective in producing the drug outcome when the cue is present (Hogarth et al., 2014). Hogarth et al. (2014) demonstrated that while simple Pavlovian extinction (presenting the cue without the outcome) failed to reduce PIT effects, discriminative extinction (degrading the stimulus and response-outcome expectancy) and explicit verbal instructions successfully abolished cue-driven responding. This pattern suggests that interventions targeting the hierarchical expectancy structure are more effective than those targeting simple stimulus-outcome associations. In this context, CC may operate through a related mechanism by reassigning negative motivational value to alcohol cues, thereby weakening their hierarchical signaling function and reducing their capacity to bias instrumental behavior.

Moreover, integrating retrieval cues within the CC protocol may enhance this effect by providing a contextual scaffold that supports the reactivation and generalization of the new aversive associations formed during training. Retrieval cues are neutral, salient stimuli presented during intervention and later reintroduced during testing to reactivate the associative memories formed during training (Collins & Brandon, 2002; Stasiewicz et al., 2007; Treanor et al., 2017). With analogy to Hogarth et al. (2014) findings that discriminative procedures effectively abolished PIT by altering cue-outcome expectancies, CC combined with retrieval cues may operate similarly, i.e., establishing competing associations wherein alcohol cues now signal aversive rather than rewarding outcomes. Retrieval cues can then facilitate access to this new learning across contexts, enabling generalization of CC effects beyond the training environment. By violating alcohol-based reward expectancies, CC could reduce approach biases and disrupt the learned associations that sustain cue-driven responding.

Building on this rationale, the present study investigated the effectiveness of CC combined with retrieval cues in reducing alcohol approach tendencies and cue-induced interference with instrumental behavior. To capture the interplay between automatic and controlled processes, we examined neural activity at frontal and centro-parietal regions. The dorsolateral prefrontal cortex (dLPFC) is critical for both cue-driven responses and cognitive control: the left dLPFC has been implicated in regulating cue-exposure processes such as craving (da Silva et al., 2013; A. M. McNeill et al., 2022), whereas the right dLPFC has been linked to controlled processes like conflict monitoring and response inhibition (A. McNeill et al., 2018; A. M. McNeill et al., 2022). At these frontal sites, the N2 event-related potentials reflect conflict resolution arising from cue-processing and inhibitory demands (Buss et al., 2011; Espinet et al., 2012), while centro-parietal activity (P3) indexes controlled attentional processes (Polich, 2007; Reed et al., 2022; Sawaki & Katayama, 2008).

We expected individuals with approach tendencies to exhibit stronger PIT effects, i.e., greater cue-driven interference with instrumental responding, compared to those with avoidance tendencies. Further, we hypothesized that CC would attenuate approach tendencies toward counterconditioned alcohol cues and reduce alcohol-related PIT effects by degrading the hierarchical instrumental (stimulus and response-outcome) expectancies that sustain cue-driven responding, with retrieval cues facilitating the generalization of these effects. These changes were expected to be reflected in both behavioral performance (reduced approach bias

and diminished PIT interference) and enhanced neural markers of cognitive control (increased N2 and P3 amplitudes).

Method

Sample

A priori power analysis was conducted using G*Power 3.1 (Faul et al., 2007) to estimate the required sample size for a repeated measures design with a within–between interaction. The analysis was based on an assumed moderate effect size (f = 0.25), an alpha level of 0.05, and a statistical power of 0.80. The analysis indicated that in total, a minimum sample size of 24 participants would be required to detect the expected effects.

To ensure sufficient representation of both approach and avoidance subgroups, 40 male alcohol users were recruited from the student community through advertisements. Males were included to minimize gender-related variability in alcohol use patterns and approach–avoidance tendencies. Since the current study aimed to investigate automatic tendency-based modulation of the Pavlovian-to-Instrumental Transfer (PIT) effect and the influence of counterconditioning, inclusion of male participants was deemed appropriate. Participants' approach and avoidance tendencies were identified post hoc based on their scores on the Alcohol Approach–Avoidance Task (A-AAT) administered during the experiment. Consequently, the number of participants in each subgroup could not be controlled or balanced prior to data collection.

During statistical analysis, one participant's data was dropped due to large EEG artifacts. The remaining 39 participants' (mean age = 21.7 years, SD = 2.32) data were included in the analysis. All participants reported being involved in alcohol consumption and were screened using the Alcohol Use Disorders Identification Test (AUDIT; Saunders et al., 1993), where 27 participants' scores ranged from 1 to 7 (indicating low-risk alcohol consumption) and 12 participants' scores ranged between 8 to 14 (indicating harmful alcohol consumption habits). All participants were right-handed or mixed right-handed, as assessed using the Edinburgh Handedness Inventory-SF (Veale, 2014), and had no history of neurological conditions. Post-A-AAT administration, 28 participants were classified into the alcohol approach group (having a positive approach-avoidance index [AAI] on a continuous scale) and 11 into the alcohol avoidance group (negative AAI) for statistical analysis (detailed classification reported in Data Analysis section).

At the time of participation, none of the individuals met the diagnostic criteria for Alcohol Use Disorder (AUD). All participants were informed in advance to abstain from alcohol consumption for a minimum of 24 hours prior to the experimental session. The study adhered to the ethical guidelines of the Declaration of Helsinki and was approved by the Institutional Ethics Committees of the University of Hyderabad (UH/IEC/2025/267) and the Indian Institute of Technology, Hyderabad (IITH/IEC/2025/01/13). Written informed consent was obtained from all participants prior to the commencement of the study.

Experimental Design and Tasks

The experiment was programmed in MATLAB using the Psychophysics Toolbox version 3.0.19 (Brainard, 1997) and presented on a 27-inch, 120 Hz monitor in portrait orientation. All tasks were completed in a dark, sound-attenuated room to minimize distractions and maintain consistency. The study followed an ABA design: participants first completed the Alcohol Approach-Avoidance Task (A-AAT) and Pavlovian-to-Instrumental Transfer (PIT) task in Context A (teal background), followed by a counterconditioning (CC) phase in Context B (gray background), and then repeated the A-AAT and PIT tasks in Context A (Figure 1A) (Stasiewicz et al., 2007; Torregrossa & Taylor, 2012). CC was implemented in this distinct context to test generalization beyond the training context, with retrieval cues integrated within the counterconditioning protocol to enhance the transfer of interventional effects across contexts. These context shifts simulated real-world treatment generalization, mimicking transitions between clinical and everyday settings. A five-minute break was provided between tasks to reduce carryover effects.

Alcohol Approach-Avoidance Task (A-AAT)

The adapted A-AAT was used to measure automatic approach or avoidance towards alcohol-related stimuli (Figure 1B) (Korucuoglu et al., 2016; Rinck & Becker, 2007). It included 180 trials divided into three blocks: two experimental blocks (80 trials each) and a neutral washout block (20 trials). In Block 1, participants were instructed to pull the joystick in response to alcohol images (approach) and push it for non-alcohol images (avoidance). Pulling or pushing resulted in a dynamic zoom-in or zoom-out effect, respectively, proportional to the joystick displacement along the y-axis. Alcohol and non-alcohol images were presented on 50% of trials each, with the sequence programmed to prevent more than three consecutive presentations of the same image type. In the washout block, eight neutral images (e.g., bells

and light bulbs) were used, where participants performed similar joystick movements (pulling bells and pushing light bulbs) to minimize any motor learning biases from Block 1. In Block 2, the instructions were switched, i.e., participants now pushed alcohol images and pulled non-alcohol images.

Each trial began with a 500-ms fixation cross, during which participants were required to maintain a neutral joystick position for at least 200 ms to prevent unintentional/premature responses. The stimulus then appeared for a maximum of 3000 ms, during which participants executed the required response. During practice trials, incorrect responses were flagged with a red cross, while a green tick indicated correctness. During the main trials, only feedback for incorrect responses was shown to help correct future trials.

Pavlovian-to-Instrumental Transfer (PIT) Task

The PIT paradigm consisted of three phases: instrumental training, Pavlovian conditioning, and transfer testing (adapted from Belanger et al., 2022; Garbusow et al., 2019). It was employed to assess whether Pavlovian cues modulate ongoing instrumental behavior.

Instrumental Learning: In an instrumental learning phase, participants completed 96 trials to learn, through trial-and-error, the reward or loss probabilities associated with six different shell stimuli (varied in color, shape, and texture). Each trial starts with a 500 ms fixation, followed by a 2000 ms stimulus screen displaying a colored shell at the center, a treasure chest at the bottom, and a trash can at the top (Figure 1C). Randomized across subjects, three shells were designated as "*good*" and assigned a positive value, while the remaining three were "*bad*" and assigned a negative value. Participants collected shells by pulling the joystick toward the treasure chest and rejected shells by pushing the joystick toward the trash can. Collecting a *good* shell produced a reinforcement of 50 tokens (gain), whereas rejecting it resulted in a punishment of 50 tokens (loss), with the reverse contingencies for *bad* shells. Each shell type appeared in 50% of trials, and approximately one-third of trials pseudo-randomly reversed the expected outcome to maintain uncertainty and prevent perfect memorization. Feedback indicating a gain or loss of 50 tokens was presented for 1500 ms after each response. Through this phase, participants acquired instrumental responses by learning the action–outcome contingencies associated with different shell types, enabling them to select the optimal action (collect vs. reject) to maximize token gains. For each trial, response accu-

racy (correct vs. incorrect), joystick position, and timestamps were recorded. Movement velocity was computed offline as the rate of change in joystick position over time.

Pavlovian Conditioning: For the Pavlovian conditioning phase, participants completed 100 trials to establish stimulus–outcome associations. Each trial began with a 3000-ms compound audiovisual cue, consisting of either an alcohol or a non-alcohol image presented at the center of the screen along with a specific tone delivered through headphones. This was followed by a 3000-ms outcome screen (Figure 1D). In 80% of trials, alcohol cues were paired with positive outcomes (e.g., +100 tokens) and non-alcohol cues with negative outcomes (e.g., –100 tokens), while in the remaining 20% of trials these contingencies were reversed. Following conditioning, an acquisition test assessed participants' explicit knowledge of the learned cue–outcome contingencies by requiring them to indicate whether each image–tone pair had been associated with reward or punishment. The accuracy of their responses was recorded.

PIT Phase (Pre-CC): This phase evaluated whether Pavlovian cues influenced instrumental decisions, with the task structure similar to the instrumental phase (Figure 1E). The task consists of 120 trials. Each trial began with a 500 ms fixation, followed by a 2000 ms stimulus screen. In 80% of the trials, the shell stimuli were presented against a tiled background containing the previously conditioned alcohol cues and associated tone, and the remaining 20% against non-alcohol cues and respective tone. In alcohol congruent trials (40%), positively conditioned alcohol cues appeared with *good* shells; in alcohol incongruent trials (40%), they appeared with *bad* shells. During the task, no feedback was provided to prevent the formation of new cue-outcome contingencies, but participants were told their decisions still affect their total token earnings.

Counterconditioning (CC)

The CC phase was conducted to reverse previously formed positive associations between alcohol cues and reward (Furlong et al., 2015; Stasiewicz et al., 2007). Across 120 trials, alcohol images were paired with monetary losses under varying stimulus conditions (Figure 1F). Specifically, in 80% of the trials, alcohol images with the respective tone were followed by loss outcomes, either a loss of 100 tokens or a larger loss of 500 tokens when accompanied by a retrieval cue (a Greek symbol, φ). To maintain learning variability and introduce decision uncertainty, 20% of the trials retained prior reward contingencies. These included audiovisual alcohol stimulus resulting in a gain of 100 tokens, or with another cue (□), lead-

ing to a larger gain of 500 tokens. The CC task was administered in Context B (gray background) to simulate a different setting in which interventions are typically delivered, thereby promoting contextual separation and supporting the generalization of new learning. Retrieval cues, visible only on screen and not explained to participants, were intended to later serve as symbolic reminders of the counterconditioning context during subsequent testing. Each trial began with a 500 ms fixation, followed by a 3000 ms stimulus display and a 3000 ms outcome screen. Similar to the conditioning phase, participants completed an acquisition test after CC to evaluate the newly formed cue–outcome associations.

Post-CC PIT and A-AAT

In the Post-CC PIT, participants completed a structurally identical task to the pre-CC PIT (Figure 1G). In this phase, the background grids contained counterconditioned alcohol images in 80% of trials. In congruent trials, negatively conditioned alcohol cues accompanied *bad* shells, while in incongruent trials, they were presented with *good* shells. Among the trials containing counterconditioned stimuli, half included retrieval cues presented in an alternate grid (Figure 1G), while the other half were presented without retrieval cues, allowing for the examination of the influence of symbolic reminders on behavior.

In the Post-CC A-AAT, the task structure remained identical to the original A-AAT, but counterconditioned alcohol stimuli were used (Figure 1H). Each image was randomly presented with or without a retrieval cue, in 50% of trials each. This manipulation allowed assessment of the impact of the retrieval cue on approach and avoidance responses.

Compensation and Debriefing

At the completion of the experiment, participants were monetarily compensated based on their performance, with a fixed minimum amount, and fully debriefed about the experiment.

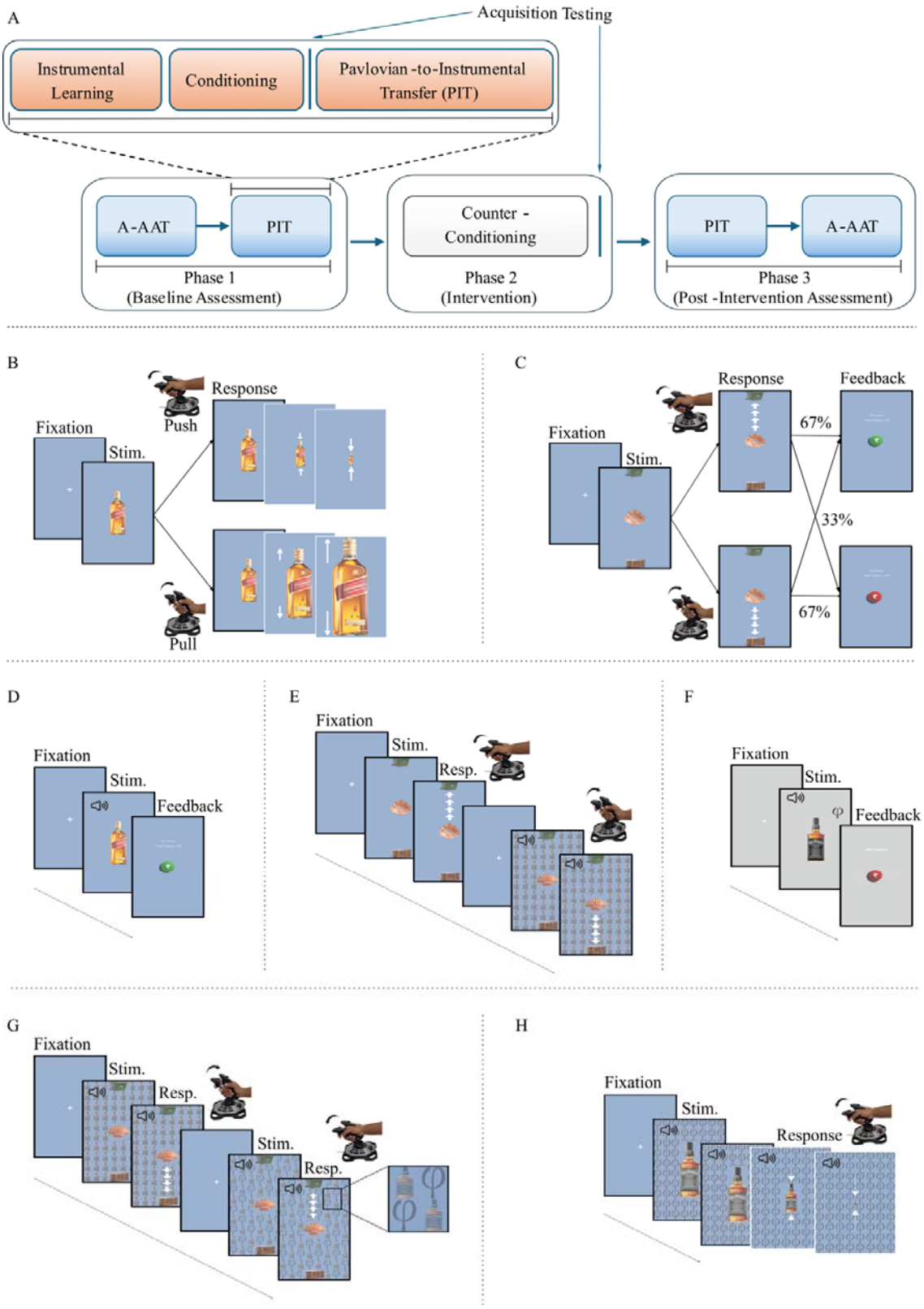

Figure 1. *Overview of the experimental paradigm. The study followed an ABA design, with tasks administered across three phases: baseline, pre-counterconditioning (pre-CC), and post-counterconditioning (post-CC)*

(Panel A). In the Approach–Avoidance Task (A-AAT, Panels B and H), participants used a joystick to pull (approach) or push (avoid) alcohol and non-alcohol stimuli. Movements triggered a simulated zoom-in or zoom-out effect along the y-axis, though these visual effects were not visible to participants. Panel C depicts the instrumental training phase, during which participants learned through trial and error to pull "good" shells toward a chest and push "bad" shells toward a bin. Panel D illustrates Pavlovian conditioning, in which alcohol-related images were repeatedly paired with monetary rewards. Panel E displays the pre-CC Pavlovian-to-Instrumental Transfer (PIT) task, in which background alcohol cues modulate instrumental responses. During counterconditioning (Panel F), alcohol cues were paired with negative outcomes, including monetary loss, with and without symbolic retrieval cues. Panel G presents the post-CC PIT task, where symbolic retrieval cues were reintroduced to assess their effect on transfer learning. Finally, Panel H shows the post-CC A-AAT, used to evaluate changes in automatic action tendencies. The white arrows shown in the figure (in Panels B, C, E, G, and H) were not visible to participants and are included solely to illustrate movement direction and cue dynamics. NR = no retrieval-cue trials post-CC, R = trials with a retrieval cue post-CC

EEG Data Acquisition and Processing

EEG data were recorded using the Brain Products ActiChamp Plus system with 64 Ag/AgCl electrodes at a sampling rate of 1 kHz. During preprocessing, a 50 Hz notch filter was applied to remove powerline interference. A bandpass filter (0.1–30 Hz) was then applied, followed by epoching (from -1000 to 2000 ms), rejection of artifactual trials, and independent component analysis (using the runica algorithm) to remove eye blink and muscle artifacts. All preprocessing was conducted using the EEGLAB toolbox in MATLAB (Delorme & Makeig, 2004).

The ERPs of interest, i.e., N2 at frontal regions and P3 at parietal regions, were extracted based on prior literature. To optimize signal clarity of concerned ERPs, a 0.2–10 Hz bandpass filter was applied (Zhang et al., 2024). Within the frontal regions, analyses targeted electrode clusters over the left and right dorsolateral prefrontal cortex (dLPFC) regions. The left frontal region of interest (LF), corresponding to the left dLPFC, was defined as the average activity from electrodes F1, F3, F5, F7, AF3, FC3, and FC5. The right frontal region of interest (RF), corresponding to the right dLPFC, was defined using electrodes F2, F4, F6, F8, AF4, FC4, and FC6. The centro-parietal region of interest (CP) was the average of CPz and Pz (Sochurková et al., 2006; Zhang et al., 2024). The mean amplitude of N2 was computed as the average amplitude between 250–650 ms post-stimulus at LF and RF, while P3 was extracted from CP in the 350–650 ms window based on a visual inspection of grand-averaged waveforms.

Data Analysis

Response time (RT) in the A-AAT was defined as the time elapsed between stimulus onset and the maximum displacement of the joystick on the y-axis (±1). For each participant, automatic approach scores for alcohol and non-alcohol images were calculated by subtracting median pull RT from push RT for each image type. The continuous Alcohol Approach Index (AAI) was then derived by subtracting the non-alcohol score from the alcohol score, with positive values (above zero) indicating an alcohol-approach bias and negative values (below zero) indicating avoidance. Accordingly, participants were classified into the approach and avoidance groups.

Demographic characteristics (age, education, and AUDIT scores) were compared using Wilcoxon rank-sum tests (for continuous variables) or Chi-square tests (for categorical variables) to ensure group comparability. Instrumental learning was operationalized as ≥80% accuracy in the last 16 trials (Belanger et al., 2022), with between-group learning accuracy compared using Chi-square tests.

Behavioral and neural data were analyzed using parametric ANOVAs; when the assumptions for these tests were violated, aligned rank transform ANOVAs (ART-ANOVA; Wobbrock et al., 2011) were applied instead. For A-AAT, interactions among automatic tendency, stimulus type, and action were examined. For PIT, error rates, interference scores (pull–push error differences), and response vigor (peak velocity) were assessed. For ERP analyses (N2 at left/right frontal region [LF/RF]; P3 at centroparietal region [CP]). Significant ANOVAs were followed by post-hoc Wilcoxon/t-tests. To assess the effect of CC, comparisons were made across baseline, post-CC without a retrieval cue (No R-Cue), and post-CC with a retrieval cue (R-Cue) conditions.

Spearman correlations examined associations between: (1) AAI and ERP alcohol–non-alcohol differences, and (2) action-related ERP differences (pull–push) and PIT interference scores. Simple linear regression tested whether baseline AAI predicted CC-induced changes (difference scores: No R-Cue - Baseline, R-Cue - Baseline). For the R-Cue condition, the Yeo-Johnson transformation (Yeo & Johnson, 2000) was applied to normalize difference scores before regression.

All analyses were conducted in R, with ART ANOVA using `ARTool` (Wobbrock et al., 2011) and Wilcoxon effect sizes using `wilcox_effsize` (Tomczak & Tomczak, 2014)

packages. Benjamini–Hochberg corrections were applied for multiple comparisons. For all null results, a complementary Bayesian analysis was conducted using the `BayesFactor` package, with Bayes factors (BF_{10}) computed using the Jeffreys–Zellner–Siow prior (r = 0.707) to assess the relative evidence for the null versus the alternative hypothesis.

Although ERP amplitudes at the left frontal region (LF) were analyzed, all comparisons and correlations at this site were non-significant across tasks; therefore, these results are not reported in the Results section. This lack of LF differences is consistent with its primary role in regulating craving-related responses, as shown in previous substance use studies (da Silva et al., 2013; A. M. McNeill et al., 2022). Because the present study employed tasks requiring rapid responses, LF activity may not have been modulated sufficiently, reflecting primarily unconscious preferential processing of alcohol cues rather than overt craving-related modulation.

Results

The current study examined how counterconditioning (CC) can influence the impact of environmental stimuli on instrumental behavior (PIT) and automatic alcohol approach tendencies. The study followed an ABA design with three phases: (1) Baseline assessment using the Alcohol Approach-Avoidance Task (A-AAT) and PIT task, (2) CC intervention in a distinct context with integrated retrieval cues, and (3) Post-intervention reassessment using the same tasks (Figure 1A). Participants were categorized into approach and avoidance groups based on their baseline Alcohol Approach Index (AAI). Our analyses address four key questions: (a) How do approach and avoidance groups differ in their baseline neural responses, and do these distinct automatic tendencies differentially influence instrumental behavior? (b) Does CC reduce approach tendencies and enhance associated neural markers? (c) Does CC reduce alcohol-related PIT effects on instrumental behavior? (d) How do behavioral and neural measures relate before and after CC?

Preliminary Demographics Exploration

Preliminary analyses explored demographic differences between the groups. The two groups, i.e., the Approach and Avoidance groups, did not differ significantly in age (W = 93, p = 0.056), overall AUDIT scores (W = 115.5, p = 0.233), or education level (Undergraduate / Postgraduate / Doctorate: $\chi^2_{(2,\ N = 39)}$ = 1.31, p = 0.520). Additionally, the distribution of par-

ticipants across AUDIT risk categories (low, moderate, and high) did not differ significantly between groups ($\chi^2_{(2, N = 39)} = 4.38$, $p = 0.112$). The absence of significant group differences in these variables suggests that they are unlikely to act as confounding covariates. Descriptive statistics for all variables are presented in Table 1.

Table 1. *Demographic and baseline characteristics of Approach and Avoidance groups. AUDIT scores are categorized according to drinking patterns: low-risk (1–7), moderate-risk (8–14), and high-risk (15+).*

		Approach	Avoidance
N [% of total]		28 [71.79]	11 [28.21]
Age (mean ± SD)		21.21 ± 2.01	22.82 ± 2.75
AUDIT (n [% within group])	low risk (1-7)	21 [75.00]	5 [45.45]
	moderate risk (8-14)	4 [14.29]	5 [45.45]
	high risk (15+)	3 [10.71]	1 [9.09]
AAI (Mean ± SE)		0.125 ± 0.018	-0.081 ± 0.025
Educational program (n [% within group])	Undergraduate	17 [60.71]	5 [45.45]
	Postgraduate	9 [32.14]	4 [36.36]
	Doctorate	2 [7.14]	2 [18.18]

Approach Tendencies Elicit Attenuated Frontal–Parietal Neural Activity and Reduced Alcohol-Related PIT During Approach (Pull) Actions

Reduced Neural Responses in Individuals with Automatic Approach Tendencies Compared to Those with Avoidance Tendencies

During baseline assessment, participants were divided into approach and avoidance groups based on their Alcohol Approach Index (AAI) (Figure 2A). This classification was manifested in the ERP data. The approach group showed weaker markers of conflict monitoring and cognitive control in response to alcohol stimuli compared to the avoidance group, as reflected by attenuated N2 and P3 amplitudes at the right-frontal ($F_{(1,111)} = 51.212$, $p < 0.001$, median difference = 2.39, $W = 244$, $p = 0.004$, $r = 0.461$; Figure 2B) and centroparietal regions ($F_{(1,111)} = 39.493$, p < 0.001, median difference = –3.850, $W = 70$, $p = 0.008$, $r = –0.425$; Figure 2C). Within the approach group, stimulus-specific attenuation was observed, with reduced RF N2 (median difference = 0.148, $W = 500$, p = 0.078, $r = 0.236$) and CP P3 (median difference = –2.120, $W = 264$, $p = 0.036$, $r = –0.280$) amplitudes for alcohol stimuli compared to non-alcohol stimuli. This pattern suggests that individuals with approach tendencies respond more automatically to alcohol cues. These ERP differences were not seen to

be affected by movement direction during the A-AAT (RF, $F_{(1,111)} = 0.983$, $p = 0.324$, $BF_{10} \approx 0.032$; CP, $F_{(1,111)} = 0.983$, $p = 0.324$, $BF_{10} \approx 0.017$).

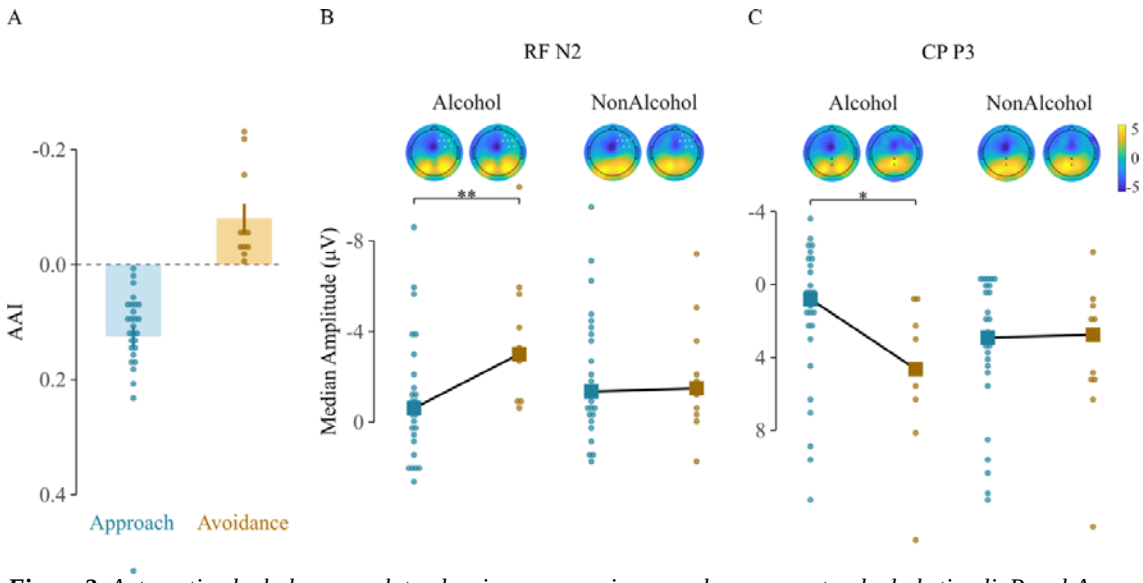

Figure 2. *Automatic alcohol approach tendencies compromise neural responses to alcohol stimuli. Panel A depicts individuals classified into approach and avoidance groups based on the Alcohol Approach Index (AAI). At both RF (Panel B) and CP (Panel C), individuals with approach tendencies show significantly attenuated neural responses to alcohol cues compared to those with avoidance tendencies. Blue and brown colors indicate approach and avoidance group data points, respectively. AAI = Alcohol Approach Index; RF = right frontal region; CP: centro-parietal region. * $p < 0.05$; ** $p < 0.01$*

Differential Influence of Automatic Tendencies on Pavlovian-to-Instrumental Transfer: Approach Tendencies Reduce Cue-Driven Interference During Pull Actions

To examine how automatic approach and avoidance tendencies influence instrumental behavior, participants completed the PIT task where alcohol cues were presented in the background. During the instrumental learning phase, both groups successfully learned the instrumental contingencies (approach participants: 81.2% accuracy; avoidance participants: 86.4% accuracy), as reflected in the last 16 trials (Belanger et al., 2022).

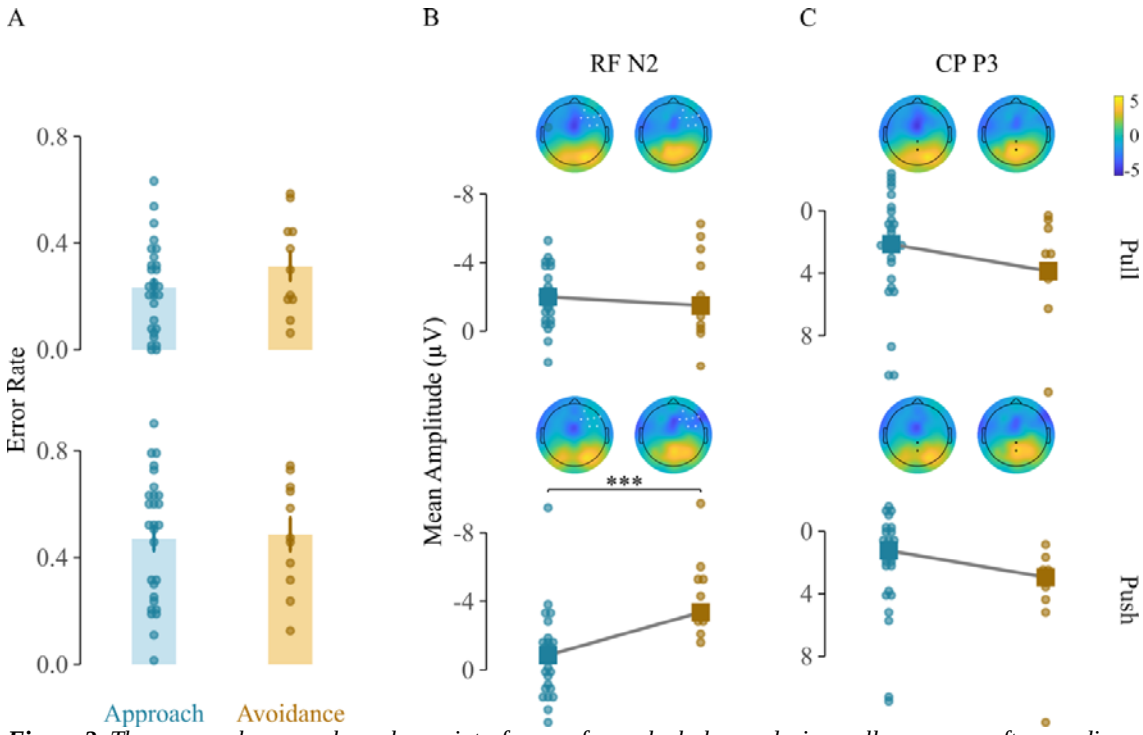

Figure 3. *The approach group shows lower interference from alcohol cues during pull responses after conditioning (during baseline PIT, Panel A), with a corresponding reduction in neural response at the RF during push responses (Panel B). However, no such change was evident at CP (Panel C). Blue and brown colors indicate approach and avoidance group datapoints, respectively. RF = right frontal region; CP = centro-parietal region. * p < 0.05; ** p < 0.01; *** p < 0.001*

During the transfer test, we compared error rates across three conditions: (1) between groups (approach vs. avoidance), (2) between actions (pull vs. push), and (3) their interaction. Consistent with predictions, alcohol cues differentially affected instrumental responding based on automatic tendencies. While overall interference from alcohol cues did not differ significantly between approach and avoidance groups ($F_{(1,74)}$ = 0.906, p = 0.344, BF_{10} ≈ 0.346; Figure 3A), or produce a group-action interaction ($F_{(1,74)}$ = 0.378, p = 0.541, BF_{10} ≈ 0.426), examination of action-specific patterns revealed stronger PIT effects in the approach group.

Specifically, within the approach group, we compared pull versus push error rates. The approach group committed significantly fewer errors during pull actions compared to the push actions ($t_{(27)}$ = -4.31, p < 0.001, d = -1.15). This pattern demonstrates that approach tendencies systematically modulate how alcohol cues influence instrumental responding, i.e., individuals with approach biases showed fewer errors during approach-consistent actions and greater interference with approach-inconsistent actions. The same comparison within the

avoidance group showed only a marginal significance for the same ($t_{(10)}$ = -2.10, p = 0.049, d = 0.90, BF_{10} ≈ 1.27), indicating weak evidence for a true difference between pull and push actions. Together, these results demonstrate that PIT effects, i.e., the influence of Pavlovian cues on instrumental responding, were strongest in individuals with approach tendencies, with alcohol cues systematically facilitating approach-consistent actions while interfering with approach-inconsistent actions.

Attenuated Frontal ERP Amplitudes During Pull Actions in the Approach Group

At the neural level, we examined whether the observed behavioral patterns were accompanied by differences in conflict monitoring (N2 at right frontal regions) and attentional allocation (P3 at centroparietal regions) during the PIT task. At the neural level, the differences in ERPs were observed at the RF ($F_{(1,37)}$ = 69.637, p < 0.001; Figure 3B) based on group, as well as action, i.e., whether the participants were required to make pull or push actions in the presence of alcohol cues in order to maximize the reward. Particularly, the ERP responses in the approach group were significantly reduced during push actions, compared to the avoidance group (median difference = 2.500, W = 274, p < 0.002, r = 0.600; Figure 3B). This attenuated N2 amplitude during push (approach-incongruent) actions in the approach group suggests reduced engagement of conflict monitoring mechanisms when Pavlovian alcohol cues interfere with avoidance responses. The combination of increased behavioral errors and reduced neural conflict monitoring during push actions indicates that approach-incongruent responses are particularly vulnerable to Pavlovian interference in individuals with strong approach tendencies. On the other hand, P3 activity at the centroparietal region ($F_{(1,37)}$ = 0.125, p = 0.726, BF_{10} ≈ 0.106; Figure 3C) remained largely unaffected, suggesting that alcohol cues primarily interfere with conflict monitoring rather than attentional allocation stages during instrumental responding.

We also examined whether automatic tendencies affected response vigor (peak velocity) during the PIT task. Automatic tendencies did not influence the peak velocity for pull or push actions ($F_{(1,70)}$ = 0.001, p < 0.974, BF_{10} ≈ 0.028), despite the presence of alcohol distractors. It suggests that Pavlovian alcohol cues primarily influenced neural and error-based measures of PIT, but not the vigor of instrumental movements.

Counterconditioning (CC) Reduces AAI and Enhances Frontal–Parietal Neural Responses, particularly in the Presence of a Retrieval Cue

Counterconditioning (CC) paired alcohol cues with monetary loss to reassign negative motivational value to alcohol cues and thereby weaken their capacity to bias instrumental behavior, along with reducing alcohol approach tendencies. By pairing alcohol cues with aversive outcomes, CC violates alcohol-based reward expectancies and establishes competing expectancies that alcohol cues now signal negative rather than positive outcomes. CC was administered in a distinct context (grey background), and retrieval cues (symbols) were introduced to support transfer of these effects to the testing context. Following CC, both the approach and avoidance groups associated alcohol cues with negative outcomes, with no significant group differences ($\chi^2_{(1, N=39)} = 0.01$, $p = 0.92$). This confirms that CC successfully altered participants' explicit associations with alcohol cues.

Selective Reduction of Behavioral AAI in the Approach Group Following CC

We compared AAI scores across three conditions: baseline, post-CC without retrieval cues, and post-CC with retrieval cues, separately for approach and avoidance groups. Following CC, approach tendencies significantly decreased compared to baseline ($F_{(2,74)} = 11.843$, $p < 0.001$; Figure 4A), accompanied by enhanced N2 and P3 amplitudes at RF ($F_{(2,407)} = 43.153$, $p < 0.001$; Figure 4C) and CP ($F_{(2,407)} = 14.840$, $p < 0.001$; Figure 4D), respectively. This reduction in approach tendencies following CC suggests that reassigning a negative value to alcohol cues successfully weakened their automatic motivational pull.

Within the approach group, pairwise comparisons revealed significant changes in the approach group. The reduction in AAI after CC was prominent in individuals who showed a baseline approach tendency irrespective of the absence (median difference = -0.085, $W = 324$, $p = 0.002$, $r = 0.404$) or the presence of retrieval cues (median difference = -0.098, $W = 281$, $p < 0.001$, $r = 0.584$). However, the presence of a retrieval cue showed an additional reduction in approach tendency (No R-Cue vs R-Cue: median difference = -0.013, $W = 281$, $p = 0.028$, $r = 0.212$; Figure 4A). The enhanced effectiveness with retrieval cues suggests that these cues facilitated access to the new aversive associations across contexts, consistent with their intended role in promoting generalization.

Within the avoidance group, the same comparisons showed no significant change after CC (No R-Cue: median difference = –0.048, $W = 17$, p = 0.360, $BF_{10} \approx 0.835$; R-Cue: median difference = 0.002, $W = 19$, $p = 0.189$, $BF_{10} \approx 0.611$). This specificity, i.e., CC effects observed only in the approach group, suggests that the intervention selectively targeted individuals for whom alcohol cues held positive motivational value.

Notably, despite the significant reduction in AAI for the approach group, their scores remained higher than the avoidance group's baseline levels, even with retrieval cues (median difference = –0.068, W = 238, p = 0.009, r = 0.420), indicating that counterconditioning reduced approach biases but did not fully normalize them.

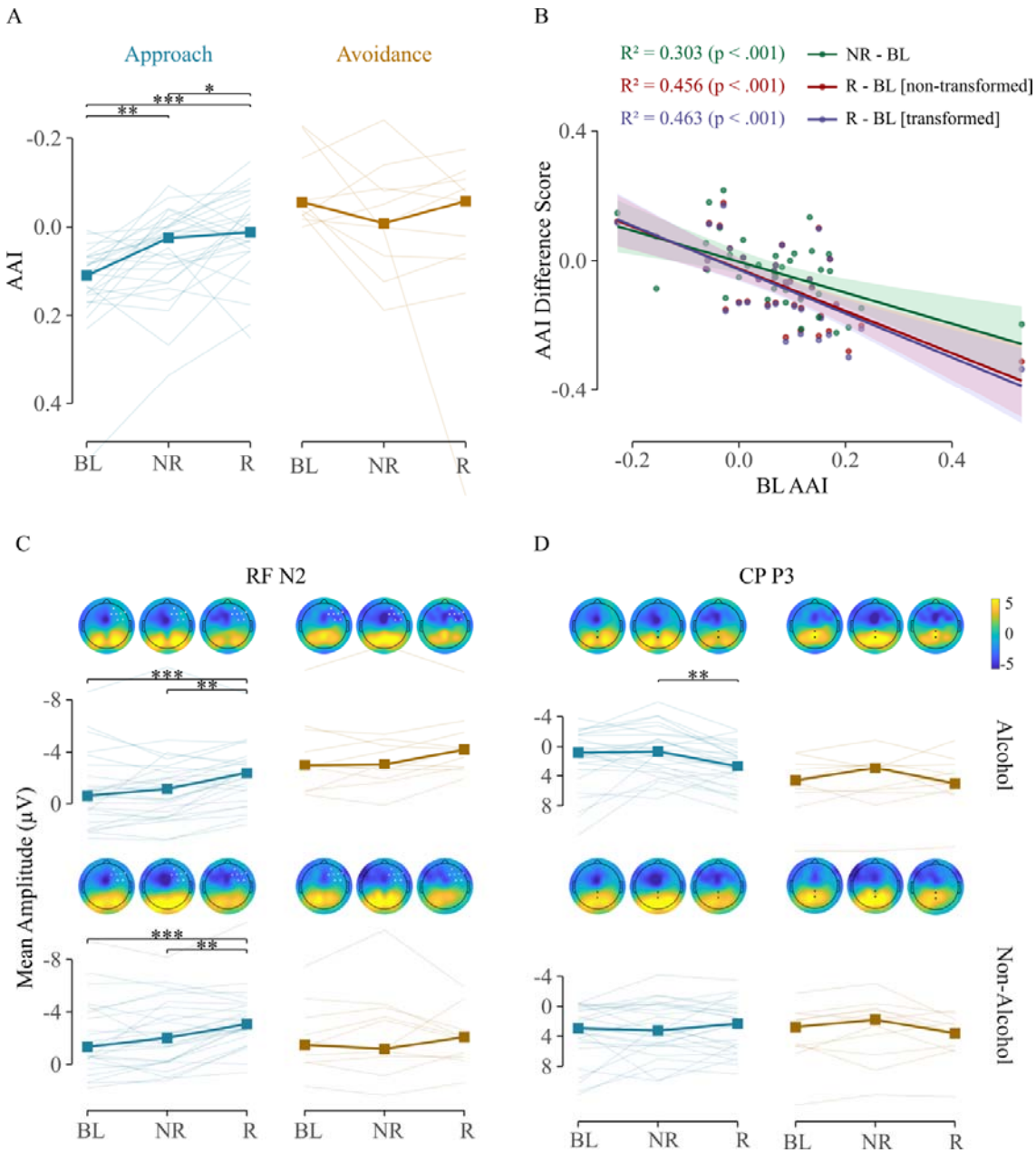

Figure 4. *After CC, individuals with approach tendencies showed significant reductions in automatic approach tendency towards alcohol cues (Panel A), along with neural enhancements at both RF (Panel C) and CP (Panel D), particularly in the presence of a retrieval cue. Panel B shows regression lines predicting the magnitude of change in AAI after CC based on initial automatic tendency scores. The R-Cue - Baseline data, with Yeo-Johnson transformation (due to violation of normality and homoscedasticity assumptions), along with non-transformed data, were plotted, showing the predictive value of baseline approach scores in estimating the*

*interventional effect of CC. Baseline: baseline PIT response (before CC); No R-Cue = no retrieval cue trails post-CC; R-Cue = retrieval cue trails post-CC; RF = right frontal region; CP = centro-parietal region. * p < 0.05; ** p < 0.01; *** p < 0.001*

Post-CC Enhancement of Right Frontal ERP Amplitudes in Individuals with Approach Tendencies

N2 amplitudes at the right frontal (RF) region were analyzed across the same three conditions. In the approach group, CC enhanced N2 amplitude at RF in response to alcohol stimuli when retrieval cues were present (median difference = 1.756, $W = 39$, $p < 0.001$, $r = 0.706$), but not without cues (median difference = 0.526, $W = 166$, $p = 0.412$, $BF_{10} \approx 0.343$; no retrieval cue vs retrieval cue trials: median difference = -1.230, $W = 61$, $p = 0.001$, $r = 0.611$; Figure 4C). The enhanced N2 amplitude suggests increased engagement of conflict monitoring processes, potentially reflecting the detection of conflict between the newly established aversive associations and the pre-existing approach tendencies.

Notably, the impact of the retrieval cue was not restricted to alcohol stimuli. It enhanced frontal ERP responses even on non-alcohol trials (Baseline vs No R-Cue: median difference = -0.693, $W = 141$, $p = 0.164$, $BF_{10} \approx 0.445$; Baseline vs R-Cue: median difference = -1.729, $W = 40$, $p < 0.001$, $r = 701$; No R-Cue vs R-Cue: median difference = -1.036, $W = 57$, $p = 0.001$, $r = 0.629$; Figure 4C), suggesting that by reminding participants of the counterconditioning relationship, the cue may have reactivated the aversive associations and engaged the conflict-monitoring mechanisms in anticipation of alcohol trials.

For the avoidance group, no significant N2 changes are observed after CC from the baseline response towards alcohol cues, irrespective of absence (median difference = 0.069, $W = 60$, $p = 0.999$, $BF_{10} \approx 0.341$) or the presence of a retrieval cue (median difference = 1.200, $W = 48$, $p = 0.657$, $BF_{10} \approx 0.491$).

CC Produces Alcohol- and Approach Tendency-Specific Increases in Centroparietal P3 Amplitudes

For the approach group, CP P3 amplitudes after CC remained largely unchanged compared to the baseline, indicating that attentional and controlled stimulus processing was unaffected (No R-Cue: median difference = 0.153, $W = 363$, $p = 0.643$, $BF_{10} \approx 0.805$; R-Cue: median difference = –1.833, $W = 497$, $p = 0.087$, $r = 0.229$, $BF_{10} \approx 0.561$). However, comparing the two post-CC conditions, the presence of a retrieval cue did show P3 enhancement compared

to its absence when responding to the alcohol stimuli (median difference = –1.986, W = 358, p = 0.001, r = 0.667; Figure 4D).

For the avoidance group, both non-retrieval (median difference = 1.733, W = 57, p = 0.847, $BF_{10} \approx 0.356$) and retrieval trials (median difference = –0.395, W = 63, p = 0.898, $BF_{10} \approx 0.311$) showed no significant change at the centroparietal (CP) region.

Baseline Approach Strength Predicts the Magnitude of CC-Induced Reduction in Approach Tendencies

We also examined whether the strength of baseline automatic tendencies predicted the magnitude of change in approach scores following CC. Higher baseline approach scores predicted greater reductions in AAI, regardless of the absence or presence of retrieval cues (No R-Cue: $b = -0.478$, $F_{(1,36)} = 15.66$, $p < 0.001$, $r^2 = 0.303$; R-Cue: $b = -0.650$, $F_{(1,36)}=30.14$, $p < 0.001$, $r^2 = 0.456$), with a stronger effect observed when a retrieval cue was present (Figure 4B). However, the difference scores in the retrieval cue condition violated the assumptions of normality and homoscedasticity. Hence, a Yeo-Johnson transformation was applied, which also yielded similar results ($b = -0.677$, $F_{(1,36)} = 30.14$, $p < 0.001$, $r^2 = 0.463$), confirming the robustness of this relationship. This dose-response relationship, i.e., a stronger baseline approach predicting greater reduction, suggests that CC specifically targets the motivational value assigned to alcohol cues, with greater initial positive value allowing for larger shifts toward negative value.

CC with Retrieval Cues Reduced PIT Interference for Push, but Increased for Pull Actions

Behavioral (error rates) and neural (RF N2 amplitude) indices during the PIT task were examined across baseline and post-CC conditions. If CC successfully degraded expectancies, we expected to see reduced interference with instrumental responding, particularly for actions that were incongruent with approach tendencies (push/avoidance actions). Results indicate CC significantly modulated both behavioral performance ($F_{(3, 259)} = 19.481$, $p < 0.001$; Figure 5A) and neural responses at the RF ($F_{(3, 259)} = 10.247$, $p < 0.001$; Figure 5B) compared to baseline PIT.

CC Diminishes Alcohol Cue-Induced Bias During Push (Avoidance) Actions in the Approach Group

We first examined whether CC reduced interference during push actions by comparing baseline push errors to post-CC push errors with and without retrieval cues. In the approach group, alcohol cue–induced interference during push (avoidance) trials was markedly reduced after CC, compared to the baseline, when retrieval cues were present (median difference = 0.438, $W = 336$, $p = 0.004$, $r = 0.570$), whereas no significant change was observed without the retrieval cues (median difference = -0.035, $W = 165$, $p = 0.393$, $BF_{10} \approx 0.276$; Figure 5A). This reduction in interference with avoidance actions is consistent with CC degrading the expectancy that alcohol cues signal at the baseline PIT. By establishing competing expectancies of aversive outcomes, CC appears to have weakened the capacity of these cues to bias instrumental behavior toward approach and away from avoidance.

This behavioral improvement (reduced avoidance interference) after CC was accompanied by stronger conflict-related neural activity (N2) at RF during push responses. Non-retrieval trials showed a moderate increase (median difference = 0.640, $W = 317$, $p = 0.016$, $r = 0.191$), while retrieval trials exhibited a larger effect (median difference = 2.950, $W = 403$, $p < 0.001$, $r = 0.712$). Direct comparison between non-retrieval cue and retrieval cue trials after CC confirmed a significant enhancement with retrieval cues (median difference = 2.310, $W = 9$, $p < 0.001$, $r = 0.646$; Figure 5B). The enhanced N2 during push actions after CC suggests heightened conflict monitoring, potentially reflecting the competition between the original hierarchical expectancy (alcohol cues → approach effective) and the newly established aversive association (alcohol cues → negative outcome). The retrieval cue's role in amplifying this neural response further supports its function in reactivating and maintaining access to the new aversive associations.

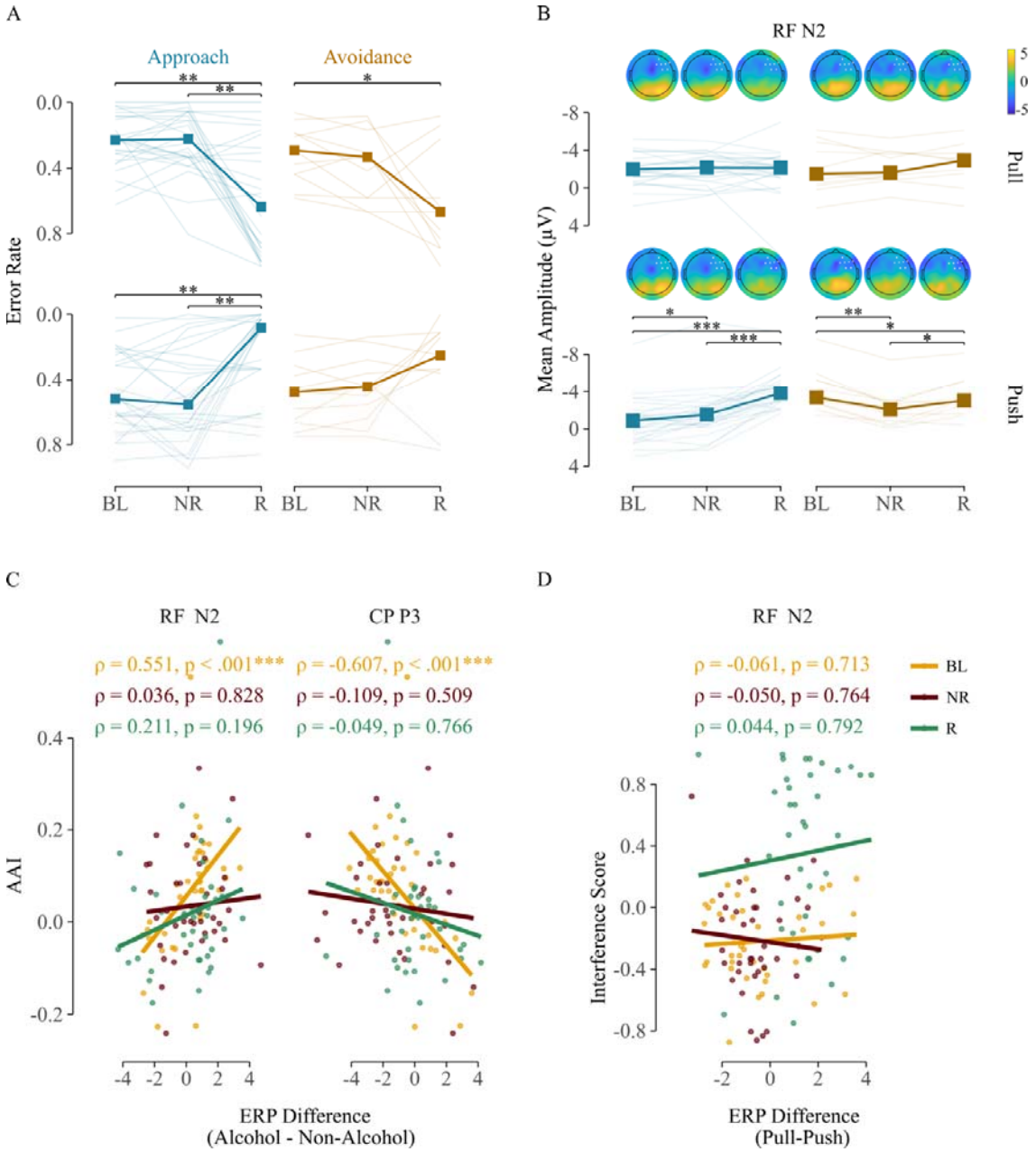

Figure 5. *The approach group shows prominent changes after CC, with enhanced push actions (lower error rates) in the presence of a retrieval cue, whereas both groups exhibit increased interference during pull actions (Panel A). Push actions were further complemented by enhanced neural response at RF in the approach group (Panel B). During the baseline, a significant relationship was observed between AAI and ERP difference scores, which later disappeared after CC (Panel C). However, no significant association emerged between behavioral and neural responses during PIT (Panel D). Baseline: baseline PIT response (before CC); No R-Cue = no retrieval cue trails post-CC; R-Cue = retrieval cue trials post-CC; AAI = approach avoidance index; RF = right frontal region; CP = centro-parietal region. * p < 0.05; ** p < 0.01; *** p < 0.001*

CC Increases Alcohol Cue-Induced Interference During Pull (Approach) Actions in the Approach Group

After CC, retrieval cue presence increased the interference for pull (approach actions) in the approach group (median difference = -0.408, $W = 62$, $p = 0.004$, $r = -0.574$; Figure 5A), accompanied by greater overall alcohol cue interference (pull-push error rate; median difference = -0.731, $W = 60$, $p = 0.002$, $r = -0.613$), though no neural changes were observed for pull trials (all $p > 0.05$). This shift in interference following CC was primarily driven by the retrieval cues, as evidenced by a significant increase in pull interference (median difference = -0.408, $W = 62$, $p = 0.001$, $r = -0.574$) and a decrease in push interference (median difference = 0.473, $W = 338$, $p = 0.001$, $r = 0.674$), relative to trials without retrieval cues.

In the avoidance group, behavioral performance showed no significant changes after CC (all p > 0.05). However,conflict-related neural activity (N2) at RF was enhanced during push responses, suggesting increased engagement of conflict monitoring mechanisms despite the absence of behavioral change (all behavioral $p > 0.05$; Figure 5B). Non-retrieval trials showed a modest increase (median difference = -0.342, $W = 7$, $p = 0.038$, $r = 0.203$), whereas retrieval trials exhibited a larger enhancement (median difference = -1.299, $W = 0$, $p = 0.006$, $r = 0.483$). Comparison between no retrieval cue and retrieval cue conditions confirmed a significant effect of retrieval cues (median difference = 0.957, $W = 5$, $p = 0.030$, $r = 0.357$).

CC Abolished A-AAT Neural–Behavioral Correlation; No PIT–ERP Relationship

To assess the relationship between behavioral outcomes and neural responses during the A-AAT and PIT tasks, correlations were performed. Before counterconditioning, alcohol approach tendencies (AAI) were significantly associated with ERP amplitude differences between alcohol and non-alcohol cues (N2: $\rho = 0.551$, $p < 0.001$; P3: $\rho = -0.607$, $p < 0.001$; Figure 5C). Specifically, individuals with stronger approach tendencies showed more attenuated N2 and P3 responses to alcohol cues, consistent with weaker conflict monitoring (N2) and reduced allocation of controlled attentional resources (P3) when processing alcohol stimuli.

After counterconditioning, these associations disappeared, both in the absence (P3: $\rho = -0.109$, $p = 0.509$; RF N2: $\rho = 0.036$, $p = 0.828$) and presence of retrieval cues (P3: $\rho = -0.049$, $p = 0.766$; RF N2: $\rho = 0.211$, $p = 0.196$; Figure 5C). This loss of correlation suggests that counterconditioning disrupted the functional coupling between automatic approach tenden-

cies and attenuated neural responses to alcohol cues, indicating that CC might have reassigned motivational value and engaged cognitive control processes.

For PIT, no correlations were observed between N2 difference (pull minus push) and behavioral interference scores (pull error rate minus push error rate) during the baseline (ρ = -0.061, p = 0.713) and after CC (No R-Cue: ρ = -0.050, p = 0.764; R-Cue: ρ = 0.044, p = 0.792; Figure 5D). This absence of association suggests that the neural and behavioral measures during PIT may capture different aspects of cue-driven interference, with neural responses reflecting conflict monitoring processes that operate somewhat independently of overt behavioral errors.

Discussion

The present study examined whether counterconditioning (CC) can reduce automatic alcohol approach tendencies and their influence on instrumental behavior via Pavlovian-to-Instrumental Transfer (PIT). Building on evidence that cue-evoked responding is sustained by hierarchical instrumental expectancies (Hogarth et al., 2014), we tested whether CC could degrade these expectancies by reassigning negative motivational value to alcohol cues. Our findings indicate that approach tendencies exert systematic interference during the PIT task, with corresponding reductions in ERP amplitudes over the right frontal region. Importantly, CC adaptively regulated the influence of alcohol cues on PIT responses by reducing interference with push (avoidance-related) actions and increasing interference with pull (approach-related) actions, particularly in the presence of retrieval cues. These behavioral changes were accompanied by the restoration of previously attenuated neural responses and were most pronounced in individuals with baseline approach tendencies. These findings provide novel evidence that individual differences in automatic approach tendencies systematically modulate cue-driven effects on instrumental behavior, and that CC, along with retrieval cues, can recalibrate these effects, consistent with the degradation of hierarchical expectancies.

Individual Differences in Automatic Tendencies Shape Cue-Driven Interference with Instrumental Behavior

The present findings highlight that individual differences in automatic response tendencies systematically shape how alcohol cues interfere with instrumental behavior through PIT. While debate continues regarding whether PIT reflects habitual or goal-directed processes

(Hogarth et al., 2019; Hogarth & Chase, 2011; Seabrooke et al., 2019; Watson et al., 2014), previous investigations have primarily examined group-level effects without considering how opposing automatic tendencies might produce distinct patterns of interference. Our findings demonstrate that alcohol approach tendencies facilitate pull (approach-related) actions at baseline but increase errors during push (avoidance-related) actions, suggesting that alcohol cues bias action selection toward approach-congruent responses. From a cognitive control perspective, this pattern reflects increased conflict when actions are incongruent with prepotent stimulus-driven tendencies (push) and relatively reduced conflict when actions are congruent with approach tendencies (pull) (Botvinick et al., 2001; Sharbanee et al., 2014). In individuals with approach tendencies, such facilitation likely arises from strong learned associations between alcohol cues and rewarding outcomes, which bias responses toward approach-congruent actions (Stacy & Wiers, 2010; Verma et al., 2025; Wiers et al., 2008).

The presence of similar interference across both actions in the avoidance group highlights a boundary condition of dual-process models that capture the facilitatory influence of approach biases but do not readily account for the behavior of the avoidance group. This is likely because avoidance is not an equally strongly rooted motivational bias but rather reflects the relative absence of approach drive (Carver, 2006). This asymmetry suggests that PIT effects depend on the strength of underlying motivational associations, with stronger approach biases producing more systematic interference patterns. In approach-oriented individuals, attentional and executive resources are disproportionately drawn toward cue-congruent responses (Bollen et al., 2021; Le Pelley et al., 2024), facilitating appetitive performance while increasing conflict for incongruent actions. Thus, our findings demonstrate that individual automatic tendencies determine both the magnitude and pattern of cue-driven interference, providing evidence for automatic tendency-based complexity in goal-directed accounts and habitual processes.

Counterconditioning Alters Cue Signaling and Reduces Interference by Attenuating Alcohol Approach Tendencies

The counterconditioning intervention, particularly when combined with retrieval cues, significantly reduced alcohol cue interference by weakening the automatic approach tendencies. By pairing alcohol cues with negative outcomes, CC established competing associations that altered what these cues signal, thereby weakening their capacity to bias instrumental behavior (Hogarth et al., 2014). This reduces their incentive salience, decreasing

the automatic pull toward alcohol-related responses (Houben et al., 2012; Wiers et al., 2011). The avoidance-oriented individuals showed little behavioral change after CC, consistent with the specificity of the intervention. CC targets existing positive motivational associations, which were minimal in the avoidance group. The extent of change depended on the strength of baseline tendencies, with stronger approach biases allowing greater reduction through CC, whereas weaker or avoidance-oriented biases showed minimal shifts. This dose-response relationship supports the interpretation that CC operates by modifying existing learned associations rather than establishing entirely new ones.

The inclusion of retrieval cues played a crucial role in CC. During counterconditioning, initially appetitive alcohol cues were paired with negative outcomes, and a retrieval cue was presented concurrently to enable subsequent reactivation of the newly established aversive alcohol memories (Goltseker et al., 2021). Just as Hogarth et al.'s (2014) discriminative extinction was more effective than simple Pavlovian extinction, retrieval cues may provide contextual scaffolds that support generalization of CC effects across contexts. By providing reactivation cues, this approach engages reconsolidation processes that support the strengthening of these aversive associations (Soeter & Kindt, 2015), enhancing both their recall and the generalization of counterconditioning effects beyond the immediate training context (Collins & Brandon, 2002; Stasiewicz et al., 2007; Treanor et al., 2017). Functionally, this restored balance in the PIT task by reducing automatic interference and enabling instrumental control to dominate.

Neural Mechanisms of Approach Tendency Modulation and Recalibration

The electrophysiological results provide insight into the cognitive mechanisms underlying the observed behavioral effects. In the A-AAT, approach-oriented individuals showed reduced N2 and P3 amplitudes over right frontal (RF) and centroparietal (CP) regions, reflecting impaired conflict monitoring (Buss et al., 2011; Espinet et al., 2012) and reduced attentional allocation (Polich, 2007; Reed et al., 2022; Sawaki & Katayama, 2008). These attenuated responses suggest that strong approach associations reduce engagement of cognitive control processes, consistent with more automatic processing. In the PIT, only RF N2 was reduced, indicating disruption of early conflict monitoring by approach biases, while P3 remained unaffected, likely because late-stage attentional and control processes are less sensitive to interference in instrumental responding in the presence of Pavlovian cues (Nadler et al., 2011).

After CC, earlier attenuated N2 and P3 amplitudes were enhanced in the approach group across tasks, indicating restored conflict monitoring and attentional control. The enhanced N2 amplitudes, particularly during push actions, suggest increased detection of conflict between the newly established aversive associations and pre-existing response tendencies. This ability of CC to modify early ERP components is likely explained by its direct weakening of automatic cue–outcome associations, which would otherwise evoke fast, stimulus-driven responses and diminish the recruitment of conflict monitoring mechanisms (Das et al., 2015; Van Gucht, Baeyens, et al., 2010). Critically, the disruption of the correlation between approach tendencies and ERP amplitudes after CC indicates fundamental reorganization of how alcohol cues are processed. However, P3 modulation was less consistent, particularly in the PIT, likely because CC predominantly targets automatic associative processes rather than deliberate attentional control (Das et al., 2015; Kerkhof et al., 2011; Van Gucht, Baeyens, et al., 2010).

Mechanistic Insights and Implications

Counterconditioning was found to be effective because it directly targeted the associative basis of approach biases rather than relying on effortful control strategies (Baker & Cannon, 1979; Goltseker et al., 2021; Kerkhof et al., 2011). Unlike interventions such as cognitive bias modification or inhibition training, which rely primarily on top-down regulatory mechanisms, CC operates by altering the underlying automatic associative links between alcohol cues and appetitive outcomes (Eberl et al., 2013). This mechanistic specificity explains why only the approach tendency participants improved, i.e., their appetitive associations provided a substrate for CC to modify, whereas avoidance-oriented participants lacked such ingrained biases.

By diminishing the approach drive towards alcohol cues, CC reduced conflict at both the behavioral and neural levels, allowing goal-directed systems to regain control (Keller et al., 2020). CC specifically targeted participants' automatic approach tendencies, which are thought to interfere with response conflict resolution and cognitive control, thereby mitigating their disruptive influence on task performance and recalibrating cognitive control mechanisms in the presence of alcohol cues (Yee et al., 2022). Retrieval cues introduced during counterconditioning become directly associated with the newly updated memory trace, functioning as "carriers" of the counterconditioned associations by making them accessible in

contexts beyond the training environment (Collins & Brandon, 2002; Goltseker et al., 2016; Soeter & Kindt, 2015; Stasiewicz et al., 2007).

From a translational perspective, retrieval cues provide a portable and easily implementable tool for supporting the expression of counterconditioned associations in real-world situations. Simple cues such as visual symbols or brief sensory reminders can help reinstate the updated memory network, thereby extending the impact of counterconditioning beyond the therapeutic setting (Collins & Brandon, 2002; Stasiewicz et al., 2007; Treanor et al., 2017). Consistent with this interpretation, Lörsch et al. (2024) have shown that interventions incorporating retrieval cues produce stronger and more persistent effects, highlighting their value as a mechanistically grounded strategy to enhance the effectiveness and generalizability of counterconditioning.

Limitations and Future Directions

While these findings provide strong evidence for the relationship between approach tendencies and PIT interference, as well as the effectiveness of counterconditioning in reversing this relationship, several limitations should be acknowledged. First, the laboratory-based PIT paradigm may not fully capture the complexity of real-world situations where approach tendencies interfere with adaptive behavior. Second, the specific CC procedure used in this study was designed to reverse the appetitive associations formed during the conditioning phase of the PIT task. As such, this paradigm may not be optimal for studies or clinical settings that do not employ PIT-based conditioning. In future applications outside of PIT frameworks, alternative aversive stimuli such as disgust-inducing images or other salient negative outcomes may be better suited for pairing with alcohol-related cues. Third, the unequal sample sizes between the approach and avoidance groups arose from post hoc classification based on A-AAT performance. Although this imbalance was unavoidable, as approach–avoidance tendencies could not be predetermined, future studies could incorporate preliminary screening to ensure more balanced group representation. Additionally, the use of a non-clinical sample in the current study limits the generalizability of findings to individuals diagnosed with alcohol-use disorders. Replication in clinical populations is crucial for determining the therapeutic utility of counterconditioning.

Conclusion

This study establishes that alcohol approach tendencies systematically modulate PIT responding, creating interference with instrumental avoidance behavior that may contribute to addiction maintenance. Importantly, counterconditioning with retrieval cues can reverse this interference by altering what alcohol cues signal, i.e., degrading hierarchical expectancies, thereby restoring adaptive behavioral control. The neural evidence demonstrates that this restoration occurs through normalization of conflict monitoring and attentional control processes, providing a mechanistic foundation for understanding therapeutic change. These findings advance both theoretical knowledge and clinical practice in several ways. First, by demonstrating that targeting the learned associations sustaining automatic approach tendencies can improve instrumental behavioral control, they provide a mechanistically grounded intervention strategy. Second, they contribute to debates about habitual and goal-directed control by demonstrating that cue-driven interference in PIT is modulated by automatic approach tendencies, which drive habitual responses. Critically, this interference remains sensitive to expectancy-based updating through counterconditioning, revealing an interaction between automatic and goal-directed systems shaped by individual differences in approach tendencies. This underscores the importance of personalized treatment approaches that account for individual automatic tendency profiles in addiction interventions.

References

Abiero, A. R., Ali, Z., Vissel, B., & Bradfield, L. A. (2022). Outcome-selective reinstatement is predominantly context-independent, and associated with c-Fos activation in the posterior dorsomedial striatum. *Neurobiology of Learning and Memory*, *187*, 107556. https://doi.org/10.1016/J.NLM.2021.107556

Baker, T. B., & Cannon, D. S. (1979). Taste aversion therapy with alcoholics: Techniques and evidence of a conditioned response. *Behaviour Research and Therapy*, *17*(3), 229–242. https://doi.org/10.1016/0005-7967(79)90038-X

Belanger, M. J., Chen, H., Hentschel, A., Garbusow, M., Ebrahimi, C., Knorr, F. G., Zech, H. G., Pilhatsch, M., Heinz, A., & Smolka, M. N. (2022). Development of Novel Tasks to Assess Outcome-Specific and General Pavlovian-to-Instrumental Transfer in Humans. *Neuropsychobiology*, *81*(5), 370–386. https://doi.org/10.1159/000526774

Bollen, Z., D'Hondt, F., Dormal, V., Lannoy, S., Masson, N., & Maurage, P. (2021). Understanding Attentional Biases in Severe Alcohol Use Disorder: A Combined Behavioral and Eye-Tracking Perspective. *Alcohol and Alcoholism*, *56*(1), 1–7. https://doi.org/10.1093/ALCALC/AGAA062

Botvinick, M. M., Carter, C. S., Braver, T. S., Barch, D. M., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Buss, K. A., Dennis, T. A., Brooker, R. J., & Sippel, L. M. (2011). An ERP study of conflict monitoring in 4–8-year old children: Associations with temperament. *Developmental Cognitive Neuroscience*, *1*(2), 131–140. https://doi.org/10.1016/J.DCN.2010.12.003

Campanella, S., Schroder, E., Kajosch, H., Noel, X., & Kornreich, C. (2019). Why cognitive event-related potentials (ERPs) should have a role in the management of alcohol disorders. *Neuroscience & Biobehavioral Reviews*, *106*, 234–244. https://doi.org/10.1016/j.neubiorev.2018.06.016

Carver, C. S. (2006). Approach, avoidance, and the self-regulation of affect and action. *Motivation and Emotion*, *30*(2), 105–110. https://doi.org/10.1007/S11031-006-9044-7/FIGURES/1

Cofresí, R. U., Kohen, C. B., Motschman, C. A., Wiers, R. W., Piasecki, T. M., & Bartholow, B. D. (2022). Behavioral response bias and event-related brain potentials implicate elevated incentive salience attribution to alcohol cues in emerging adults with lower sensitivity to alcohol. *Addiction*, *117*(4), 892–904. https://doi.org/10.1111/ADD.15728

Collins, B. N., & Brandon, T. H. (2002). Effects of extinction context and retrieval cues on alcohol cue reactivity among nonalcoholic drinkers. *Journal of Consulting and Clinical Psychology*, *70*(2), 390–397. https://doi.org/10.1037/0022-006X.70.2.390

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2016). The effectiveness of cognitive bias modification interventions for substance addictions: A meta-analysis. *PLoS ONE*, *11*(9), e0162226. https://doi.org/10.6084/M9.FIGSHARE.2059137

da Silva, M. C., Conti, C. L., Klauss, J., Alves, L. G., do Nascimento Cavalcante, H. M., Fregni, F., Nitsche, M. A., & Nakamura-Palacios, E. M. (2013). Behavioral effects of transcranial Direct Current Stimulation (tDCS) induced dorsolateral prefrontal cortex plasticity in alcohol dependence. *Journal of Physiology-Paris*, *107*(6), 493–502. https://doi.org/10.1016/J.JPHYSPARIS.2013.07.003

Das, R. K., Lawn, W., & Kamboj, S. K. (2015). Rewriting the valuation and salience of alcohol-related stimuli via memory reconsolidation. *Translational Psychiatry 2015 5:9*, *5*(9), e645–e645. https://doi.org/10.1038/tp.2015.132

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

den Uyl, T. E., Gladwin, T. E., & Wiers, R. W. (2016). Electrophysiological and Behavioral Effects of Combined Transcranial Direct Current Stimulation and Alcohol Approach Bias Retraining in Hazardous Drinkers. *Alcoholism: Clinical and Experimental Research*, *40*(10), 2124–2133. https://doi.org/10.1111/ACER.13171

Di Lemma, L. C. G., Stancak, A., Soto, V., Fallon, N., & Field, M. (2020). Event-related and readiness potentials when preparing to approach and avoid alcohol cues following cue avoidance training in heavy drinkers. *Psychopharmacology*, *237*(5), 1343–1358. https://doi.org/10.1007/S00213-020-05462-7

Drummond, D. C., Cooper, T., & Glautier, S. P. (1990). Conditioned learning in alcohol dependence: implications for cue exposure treatment. *British Journal of Addiction*, *85*(6), 725–743. https://doi.org/10.1111/j.1360-0443.1990.tb01685.x

Eberl, C., Wiers, R. W., Pawelczack, S., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2013). Approach bias modification in alcohol dependence: Do clinical effects replicate and for

whom does it work best? *Developmental Cognitive Neuroscience*, *4*, 38–51. https://doi.org/10.1016/J.DCN.2012.11.002

Elkins, R. L. (1980). Covert sensitization treatment of alcoholism: Contributions of successful conditioning to subsequent abstinence maintenance. *Addictive Behaviors*, *5*(1), 67–89. https://doi.org/10.1016/0306-4603(80)90023-4

Espinet, S. D., Anderson, J. E., & Zelazo, P. D. (2012). N2 amplitude as a neural marker of executive function in young children: An ERP study of children who switch versus perseverate on the Dimensional Change Card Sort. *Developmental Cognitive Neuroscience*, *2*(SUPPL. 1), S49–S58. https://doi.org/10.1016/J.DCN.2011.12.002

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Field, M., Kiernan, A., Eastwood, B., & Child, R. (2008). Rapid approach responses to alcohol cues in heavy drinkers. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(3), 209–218. https://doi.org/10.1016/j.jbtep.2007.06.001

Furlong, T. M., Pan, M. J., & Corbit, L. H. (2015). The effects of compound stimulus extinction and inhibition of noradrenaline reuptake on the renewal of alcohol seeking. *Translational Psychiatry 2015 5:9*, *5*(9), e630–e630. https://doi.org/10.1038/tp.2015.130

Gale, G., Walsh, K., Hennessy, V. E., Stemerding, L. E., Ni, K. S., Thomas, E., Kamboj, S. K., & Das, R. K. (2021). Long-term behavioural rewriting of maladaptive drinking memories via reconsolidation-update mechanisms. *Psychological Medicine*, *51*(16), 2875–2885. https://doi.org/10.1017/S0033291720001531

Garbusow, M., Ebrahimi, C., Riemerschmid, C., Daldrup, L., Rothkirch, M., Chen, K., Chen, H., Belanger, M. J., Hentschel, A., Smolka, M. N., Heinz, A., Pilhatsch, M., & Rapp, M. A. (2022). Pavlovian-to-Instrumental Transfer across Mental Disorders: A Review. *Neuropsychobiology*, *81*(5), 418–437. https://doi.org/10.1159/000525579

Garbusow, M., Nebe, S., Sommer, C., Kuitunen-Paul, S., Sebold, M., Schad, D. J., Friedel, E., Veer, I. M., Wittchen, H. U., Rapp, M. A., Ripke, S., Walter, H., Huys, Q. J. M., Schlagenhauf, F., Smolka, M. N., & Heinz, A. (2019). Pavlovian-To-Instrumental Transfer and Alcohol Consumption in Young Male Social Drinkers: Behavioral, Neural and Polygenic Correlates. *Journal of Clinical Medicine 2019, Vol. 8, Page 1188*, *8*(8), 1188. https://doi.org/10.3390/JCM8081188

Goltseker, K., Bolotin, L., & Barak, S. (2016). Counterconditioning During Reconsolidation Prevents Relapse of Cocaine Memories. *Neuropsychopharmacology 2017 42:3*, *42*(3), 716–726. https://doi.org/10.1038/npp.2016.140

Goltseker, K., Handrus, H., & Barak, S. (2021). Disruption of relapse to alcohol seeking by aversive counterconditioning following memory retrieval. *Addiction Biology*, *26*(3), e12935. https://doi.org/10.1111/adb.12935

Hogarth, L., & Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: Implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*(3), 261–276. https://doi.org/10.1037/A0022913

Hogarth, L., Lam-Cassettari, C., Pacitti, H., Currah, T., Mahlberg, J., Hartley, L., & Moustafa, A. (2019). Intact goal-directed control in treatment-seeking drug users indexed by outcome-devaluation and Pavlovian to instrumental transfer: critique of habit theory. *European Journal of Neuroscience*, *50*(3), 2513–2525. https://doi.org/10.1111/EJN.13961

Hogarth, L., Retzler, C., Munafò, M. R., Tran, D. M. D., Troisi, J. R., Rose, A. K., Jones, A., & Field, M. (2014). Extinction of cue-evoked drug-seeking relies on degrading hierarchical instrumental expectancies. *Behaviour Research and Therapy*, *59*, 61–70. https://doi.org/10.1016/J.BRAT.2014.06.001

Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer à no-go: learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction*, *107*(7), 1280–1287. https://doi.org/10.1111/j.1360-0443.2012.03827.x

Keller, N. E., Hennings, A. C., & Dunsmoor, J. E. (2020). Behavioral and neural processes in counterconditioning: Past and future directions. *Behaviour Research and Therapy*, *125*, 103532. https://doi.org/10.1016/J.BRAT.2019.103532

Kenney, S. R., DiGuiseppi, G. T., Meisel, M. K., Balestrieri, S. G., & Barnett, N. P. (2018). Poor mental health, peer drinking norms, and alcohol risk in a social network of first-year college students. *Addictive Behaviors*, *84*, 151–159. https://doi.org/10.1016/J.ADDBEH.2018.04.012

Kerkhof, I., Vansteenwegen, D., Baeyens, F., & Hermans, D. (2011). Counterconditioning: An effective technique for changing conditioned preferences. *Experimental Psychology*, *58*(1), 31–38. https://doi.org/10.1027/1618-3169/A000063

Korucuoglu, O., Gladwin, T. E., & Wiers, R. W. (2016). The effect of acute alcohol on motor-related EEG asymmetries during preparation of approach or avoid alcohol responses. *Biological Psychology*, *114*, 81–92. https://doi.org/10.1016/J.BIOPSYCHO.2015.12.012

Le Pelley, M. E., Watson, P., & Wiers, R. W. (2024). Biased Choice and Incentive Salience: Implications for Addiction. *Behavioral Neuroscience*, *138*(4), 235–243. https://doi.org/10.1037/BNE0000583

Lörsch, F., Kollei, I., & Steins-Loeber, S. (2024). The effects of a retrieval cue on renewal of conditioned responses in human appetitive conditioning. *Behaviour Research and Therapy*, *176*, 104501. https://doi.org/10.1016/J.BRAT.2024.104501

Lutz, J. A., & Childs, E. (2021). Alcohol conditioned contexts enhance positive subjective alcohol effects and consumption. *Behavioural Processes*, *187*, 104340. https://doi.org/10.1016/J.BEPROC.2021.104340

McNaughton, N., DeYoung, C. G., & Corr, P. J. (2016). Approach/Avoidance. In *Neuroimaging Personality, Social Cognition, and Character* (pp. 25–49). Elsevier. https://doi.org/10.1016/B978-0-12-800935-2.00002-6

McNeill, A. M., Monk, R. L., Qureshi, A. W., Makris, S., Cazzato, V., & Heim, D. (2022). Elevated ad libitum alcohol consumption following continuous theta burst stimulation to the left-dorsolateral prefrontal cortex is partially mediated by changes in craving. *Cognitive, Affective and Behavioral Neuroscience*, *22*(1), 160–170. https://doi.org/10.3758/S13415-021-00940-7

McNeill, A., Monk, R. L., Qureshi, A. W., Makris, S., & Heim, D. (2018). Continuous Theta Burst Transcranial Magnetic Stimulation of the Right Dorsolateral Prefrontal Cortex Impairs Inhibitory Control and Increases Alcohol Consumption. *Cognitive, Affective and Behavioral Neuroscience*, *18*(6), 1198–1206. https://doi.org/10.3758/S13415-018-0631-3/FIGURES/4

Morris, H., Larsen, J., Catterall, E., Moss, A. C., & Dombrowski, S. U. (2020). Peer pressure and alcohol consumption in adults living in the UK: A systematic qualitative review. *BMC Public Health*, *20*(1), 1–13. https://doi.org/10.1186/S12889-020-09060-2

Nadler, N., Delgado, M. R., & Delamater, A. R. (2011). Pavlovian to instrumental transfer of control in a human learning task. *Emotion*, *11*(5), 1112–1123. https://doi.org/10.1037/A0022760

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Reed, C. L., Siqi-Liu, A., Lydic, K., Lodge, M., Chitre, A., Denaro, C., Petropoulos, A., Joshi, J., Bukach, C. M., & Couperus, J. W. (2022). Selective contributions of executive function ability to the P3. *International Journal of Psychophysiology*, *176*, 54–61. https://doi.org/10.1016/J.IJPSYCHO.2022.03.004

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(2), 105–120. https://doi.org/10.1016/J.JBTEP.2006.10.001

Rinck, M., Wiers, R. W., Becker, E. S., & Lindenmeyer, J. (2018). Relapse Prevention in Abstinent Alcoholics by Cognitive Bias Modification: Clinical Effects of Combining Approach Bias Modification and Attention Bias Modification. *Journal of Consulting and Clinical Psychology*, *86*(12), 1005–1016. https://doi.org/10.1037/CCP0000321

Rosenthal, A., Chen, K., Beck, A., & Romanczuk-Seiferth, N. (2023). Modifying Pavlovian-to-instrumental transfer by approach avoidance training in healthy subjects: a proof of concept study. *Scientific Reports 2023 13:1*, *13*(1), 1–12. https://doi.org/10.1038/s41598-023-37083-3

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

Sawaki, R., & Katayama, J. (2008). Distractor P3 is associated with attentional capture by stimulus deviance. *Clinical Neurophysiology*, *119*(6), 1300–1309. https://doi.org/10.1016/J.CLINPH.2008.01.107

Seabrooke, T., Hogarth, L., Edmunds, C. E. R., & Mitchell, C. J. (2019). Goal-directed control in Pavlovian-instrumental transfer. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(1), 95–101. https://doi.org/10.1037/XAN0000191

Sharbanee, J. M., Stritzke, W. G. K., Jamalludin, M. E., & Wiers, R. W. (2014). Approach-alcohol action tendencies can be inhibited by cognitive load. *Psychopharmacology*, *231*(5), 967–975. https://doi.org/10.1007/S00213-013-3318-Z/FIGURES/1

Sochurková, D., Brázdil, M., Jurák, P., & Rektor, I. (2006). P3 and ERD/ERS in a visual oddball paradigm: A depth EEG study from the mesial temporal structures. *Journal of Psychophysiology*, *20*(1), 32–39. https://doi.org/10.1027/0269-8803.20.1.32

Soeter, M., & Kindt, M. (2015). Retrieval cues that trigger reconsolidation of associative fear memory are not necessarily an exact replica of the original learning experience. *Fron-*

tiers in Behavioral Neuroscience, *9*(MAY), 137411. https://doi.org/10.3389/FNBEH.2015.00122/BIBTEX

Stacy, A. W., & Wiers, R. W. (2010). Implicit cognition and addiction: A tool for explaining paradoxical behavior. *Annual Review of Clinical Psychology*, *6*(Volume 6, 2010), 551–575. https://doi.org/10.1146/ANNUREV.CLINPSY.121208.131444/CITE/REFWORKS

Stasiewicz, P. R., Brandon, T. H., & Bradizza, C. M. (2007). Effects of Extinction Context and Retrieval Cues on Renewal of Alcohol-Cue Reactivity Among Alcohol-Dependent Outpatients. *Psychology of Addictive Behaviors*, *21*(2), 244–248. https://doi.org/10.1037/0893-164X.21.2.244

Tomczak, E., & Tomczak, M. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *TRENDS in Sport Sciences*, *21*(1), 19–25. https://tss.awf.poznan.pl/The-need-to-report-effect-size-estimates-revisited-An-overview-of-some-recommended,188960,0,2.html

Torregrossa, M. M., & Taylor, J. R. (2012). Learning to forget: manipulating extinction and reconsolidation processes to treat addiction. *Psychopharmacology*, *226*(4), 659. https://doi.org/10.1007/S00213-012-2750-9

Treanor, M., Brown, L. A., Rissman, J., & Craske, M. G. (2017). Can Memories of Traumatic Experiences or Addiction Be Erased or Modified? A Critical Review of Research on the Disruption of Memory Reconsolidation and Its Applications. *Perspectives on Psychological Science*, *12*(2), 290–305. https://doi.org/10.1177/1745691616664725

Van Gucht, D., Baeyens, F., Vansteenwegen, D., Hermans, D., & Beckers, T. (2010). Counterconditioning Reduces Cue-Induced Craving and Actual Cue-Elicited Consumption. *Emotion*, *10*(5), 688–695. https://doi.org/10.1037/A0019463

Van Gucht, D., Van den Bergh, O., Beckers, T., & Vansteenwegen, D. (2010). Smoking behavior in context: Where and when do people smoke? *Journal of Behavior Therapy and Experimental Psychiatry*, *41*(2), 172–177. https://doi.org/10.1016/J.JBTEP.2009.12.004

Veale, J. F. (2014). Edinburgh Handedness Inventory – Short Form: A revised version based on confirmatory factor analysis. *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(2), 164–177. https://doi.org/10.1080/1357650X.2013.783045

Verma, A. K., Kumar, A. D., Chivukula, U., & Kumar, N. (2025). Neural mechanisms underlying approach and avoidance tendencies in alcohol use among males: An electrophysiological investigation. *Biological Psychology*, *202*, 109144. https://doi.org/10.1016/J.BIOPSYCHO.2025.109144

Watson, P., Wiers, R. W., Hommel, B., & De Wit, S. (2014). Working for food you don't desire. Cues interfere with goal-directed food-seeking. *Appetite*, *79*, 139–148. https://doi.org/10.1016/J.APPET.2014.04.005

Wiens, A. N., Montague, J. R., Manaugh, T. S., & English, C. J. (1976). Pharmacological aversive counterconditioning to alcohol in a private hospital; one-year follow-up. *Journal of Studies on Alcohol*, *37*(9), 1320–1324. https://doi.org/10.15288/jsa.1976.37.1320

Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining Automatic Action Tendencies Changes Alcoholic Patients' Approach Bias for Alcohol and Improves Treatment Outcome. *Psychological Science*, *22*(4), 490–497. https://doi.org/10.1177/0956797611400615

Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., & Ridderinkhof, K. R. (2013). Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology. *Clinical Psychological Science*, *1*(2), 192–212. https://doi.org/10.1177/2167702612466547

Wiers, R. W., Havermans, R., Deutsch, R., & Stacy, A. W. (2008). A mismatch with dual process models of addiction rooted in psychology. *Behavioral and Brain Sciences*, *31*(4), 460. https://doi.org/10.1017/S0140525X08004962

Wiers, R. W., & van Dessel, P. (2023). Mechanisms Underlying Alcohol Approach Bias and Its Modification. *Biological Psychiatry Global Open Science*, *3*(3), 317–318. https://doi.org/10.1016/J.BPSGOS.2023.02.010

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. *Conference on Human Factors in Computing Systems - Proceedings*, 143–146. https://doi.org/10.1145/1978942.1978963

Yee, D. M., Leng, X., Shenhav, A., & Braver, T. S. (2022). Aversive motivation and cognitive control. *Neuroscience & Biobehavioral Reviews*, *133*, 104493. https://doi.org/10.1016/J.NEUBIOREV.2021.12.016

Zhang, G., Garrett, D. R., & Luck, S. J. (2024). Optimal filters for ERP research II: Recommended settings for seven common ERP components. *Psychophysiology*, *61*(6), e14530. https://doi.org/10.1111/PSYP.14530