bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Generalized habitual tendencies in alcohol dependent rats

Francesco Giannone[1], Arian Hach[1], Magda Chrószcz[2], Marion M. Friske[1], Marcus Meinhardt[1], Rainer Spanagel[1], Wolfgang H. Sommer[1*], Anita C. Hansson[1*]

[1]Institute of Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, 68159 Mannheim, Germany
[2]Maj Institute of Pharmacology Polish Academy of Sciences, Department of Molecular Neuropharmacology, Krakow, 31-343, Poland

*Corresponding authors:

Anita C. Hansson
Central Institute of Mental Health
Institute of Psychopharmacology
Square J5
68159 Mannheim, Germany
Phone: +49 0621/1703 6293
Fax: +49 621/1703 6255
E-mail: anita.hansson@zi-mannheim.de

Wolfgang H. Sommer
Central Institute of Mental Health
Institute of Psychopharmacology
Square J5
68159 Mannheim, Germany
Phone: +49 0621/1703 6293
Fax: +49 621/1703 6255
E-mail: wolfgang.sommer@zi-mannheim.de

Key words: alcohol addiction, post-dependent animals, goal-directed, habit learning, devaluation, DREADD, natural reward, dorsomedial striatum

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Habitual responses and ultimately compulsive behavior are thought to be at the core of addiction including alcohol use disorder (AUD). Little is known whether the habitization concerns exclusively the response towards alcohol or generalizes to other daily activities. Here, we address this question in a well-established animal model of AUD – the post-dependent rat model – by testing habitual responses towards a sweet palatable reward in two striatal learning paradigms: spatial navigation and reward conditioning. For the spatial navigation task, alcohol-dependent and control rats were tested on a sequential decision-making test after short and prolonged T-Maze training; for the reward conditioning task, rats were trained under a random interval schedule for a short and prolonged period and tested in a satiety devaluation test at each time point. Another cohort of alcohol-naive rats was trained and tested on both paradigms under DREADD (designer receptors exclusively activated by designer drugs)-mediated inactivation of the dorsomedial striatum (DMS) which controls goal-directed behavior. Our results show that alcohol-dependent rats displayed increased habitual behavior to obtain saccharin reward on both paradigms, with overall more habitual choices after prolonged training on the spatial navigation task, and increased habitual responses already after short training on the reward conditioning task. Finally, DREADD-mediated inactivation of the DMS increased habitual behavior in non-dependent rats on both paradigms. Our results provide evidence that a history of alcohol dependence produces a bias towards habitual responding that generalizes to a natural reward in rats. Similarly, a habitual bias was induced in non-dependent rats after inactivation of the DMS, thus confirming the critical role of this region in maintaining goal-directed behavior and suggesting its diminished control in AUD.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

A habit is a consolidated response to a stimulus, which once acquired is relatively resistant to interference and as such largely independent of action-outcome contingencies (Yin and Knowlton, 2006; Yin et al., 2009). A habit develops by extensive stimulus-response pairings and repetition of the same response sequence in the same context, and it is thought to depend on striatal learning processes. There is considerable evidence, mostly coming from rodent studies, suggesting the involvement of two different learning processes, the dorsomedial striatum (DMS) controlling the acquisition of goal-directed actions mainly via dopamine D1 receptors, and the dorsolateral striatum (DLS) acquisition of habits and skills via dopamine D2 receptors (Yin et al., 2009; Sommer et al., 2014). The two striatal regions receive excitatory inputs from different areas of the cortex. The DMS receives modulatory input from prefrontal cortices, whereas the DLS primarily receives major inputs from sensorimotor and premotor cortices in both humans and rats (Balleine et al., 2009; Balleine and O'Doherty, 2010; Corbit and Janak, 2016a). Dependent on DMS activity, goal-directed behavior is mainly involved during the early instrumental conditioning, and with overtraining habitual responding occur under control of the DLS (Yin et al., 2004, 2005a; Yin et al., 2005b; Yin and Knowlton, 2006; Yin et al., 2006; Balleine et al., 2009). Some studies suggest that the DMS and DLS operate in parallel during the transition from goal-directed to habitual responses (Yin et al., 2009; Vandaele et al., 2019), with the DLS already active at the early onset of training (Kupferschmidt et al., 2017; Bergstrom et al., 2018; Smith et al., 2021), and it appears that both subregions act via an opposing and competitive relationship (Smith and Graybiel, 2013; Turner et al., 2022).

The harmful use of alcohol is a global problem causing 3 million deaths per year (which accounts to 5.3% of deaths worldwide) and is attributed to the world's third largest risk factor for premature mortality, disability and loss of health (Rehm et al., 2009; WHO, 2018). When considering the total harm inflicted by alcohol including its societal costs, alcohol is by a wide margin the most harmful drug in the western world (Nutt et al., 2010). Alcohol use disorder (AUD) is characterized by the loss of flexible control over drug use despite negative consequences (Rehm et al., 2009; Connor et al., 2016; Koob, 2021). Current addiction theories postulate that the loss of control may be explained with increased habitual behavior in which drug use persists even when the drug is no longer rewarding, and ultimately leading to

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

compulsive alcohol drinking (Everitt and Robbins, 2005; Belin et al., 2009; Belin et al., 2013; Everitt and Robbins, 2016) and habitual drug-seeking (Everitt et al., 2008; Ostlund and Balleine, 2008; Belin et al., 2009; Root et al., 2009). Recent research has also implicated the dorsal striatum in the control of decision-making regulated by reward, particularly the role of the DMS in the integration of reward-related processes with action control (Balleine et al., 2007). The DLS has been linked with habitual responding for alcohol in animals (Corbit et al., 2012; Serlin and Torregrossa, 2015; Giuliano et al., 2019) and natural rewards (Dickinson et al., 2002; Corbit et al., 2012; Houck and Grahame, 2018; Turner et al., 2022) and humans (Vollstadt-Klein et al., 2010).

Most studies assessing habitual responses in animals use operant conditioning, and there are only few publications demonstrating the impact of alcohol exposure. Both acute (Houck and Grahame, 2018) and chronic ethanol exposure (Corbit et al., 2012; Renteria et al., 2018; Towner and Spear, 2021) shifted animals toward behaving more habitually. Most of these studies assess habitual responses based on their sensitivity to devalued reinforcers. The results are difficult to interpret, since they rather may relate from impaired instrumental learning abilities than habits (Balleine and Dezfouli, 2019).

In rodents, the spatial navigation task has been successfully used for assessment of goal-directed and habitual behavior (Packard, 1999; Yin and Knowlton, 2004; Lee et al., 2008; McCracken and Grace, 2013). Thus, spatial navigation could present an alternative to reward conditioning, although there are currently no data available from spatial navigation tasks assessing alcohol-related habitual responses in rodents.

This study aimed to investigate how habit formation processes are affected by a history of alcohol dependence using two striatal learning paradigms, the spatial navigation and reward conditioning task. We tested habitual responses towards a sweet palatable reward in both tasks in a well-established animal model of AUD – the post-dependent (PD) rat model. Alcohol dependence was introduced in adult rats by repeated cycles of intermittent exposure (CIE) to alcohol vapor intoxication and withdrawal (Rimondini et al., 2002). We further analyzed the effects of a designer receptors exclusively activated by the designer drugs (DREADD)-mediated inactivation of the DMS in habitual behavior.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Materials and Methods

Experimental animals and design

Male Wistar rats were obtained from Envigo (Germany), 6 weeks of age at arrival. Rats were group-housed (four animals per Makrolon type 4 cage, cage dimension: 85 x 85 x 67 cm) under a 12h light/dark cycle (lights off at 6 AM, lights on at 18:00 PM) with ad libitum access to water and food, and left to acclimatize to the environment for one week. They were handled daily for the following week before beginning the experimental procedures. All experimental procedures started in the dark phase (2hrs after lights off) in adult rats (≥ 8 weeks). The experimental procedures were approved by the Regierungspräsidium Karlsruhe, Germany (license: G-166/21).

Three separate cohorts of male rats were used for this research, each allocated to a different experiment. Study 1 was performed to assess the effects of alcohol dependence on habitual behavior in the context of spatial navigation (Experiment 1) and reward learning (Experiment 2) tasks (**Figure 1**). Study 2 was performed to evaluate the effects of DREADD-mediated posterior dorsomedial striatum (pDMS) inactivation on habitual behavior in the context of both spatial navigation and reward learning tasks (Experiment 3, **Figure 1**).

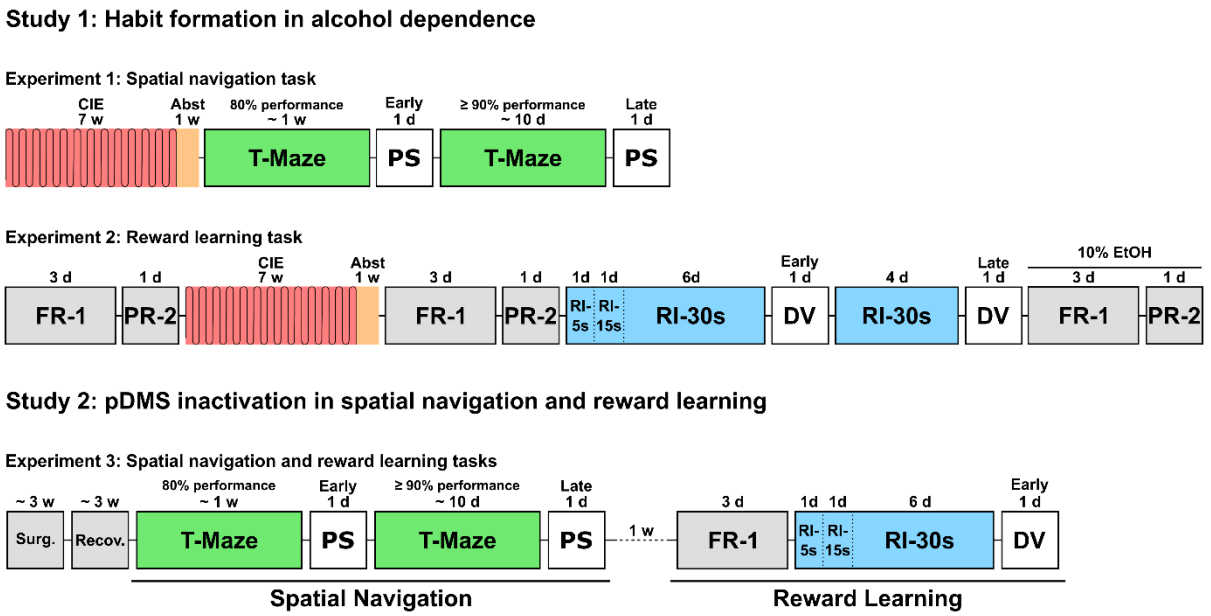


Figure 1. Experimental timelines. In Study 1 habitual responses (spatial navigation, reward learning) were analyzed in two cohorts of alcohol dependent and non-dependent rats. At the end of experiment 2 rats were trained to self-administer 10% ethanol to confirm the alcohol dependent phenotype. In Study 2 rats were tested for both behavior paradigms after inactivation of the posterior part of the dorsomedial striatum (**pDMS**). In all three experiments 0.2% saccharin was given as a natural reward. **CIE:** Chronic Intermittent Ethanol vapor exposure, **Abst:** abstinence, **FR-1:** Fixed Ratio 1 operant schedule, **PR-2:** Progressive Ratio-2 operant schedule, **RI:** Random Interval operant schedule (5s, 15s, and 30s indicate the interval length in seconds), **PS:** probe session composed of 5 consecutive trials, **DV:** Devaluation test, **Surg.**: surgery, **Recov.**: recovery from surgery.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Material details

Chemicals, drugs and kits

Alexa Fluor 488 donkey anti-mouse IgG(H+L) (Thermo Fisher, Ref: AB_141607), Alexa Fluor 594 donkey anti-rabbit IgG(H+L) (Thermo Fisher, Ref: AB_141637), anti-fade mounting medium (Shandon Immu-Mount^{TM}, Thermo Scientific), Carprofen (50 mg/ml, Zoetis Deutschland GmbH, Germany), Clozapine N-oxide (CNO) dihydrochloride (BIOZOL Diagnostica Vertrieb GmbH, Germany), DREADDs (ssAAV-8/2 virus expressing either hM4D(Gi)-mCherry (titer: 7.3 x 10^{12}) or eGFP (titer: 6.9 x 10^{12}), University of Zürich and ETH Zürich, Neuroscience Center Zürich ZNZ, Viral Vector Facility VVF, Switzerland), isoflurane (1 ml/ml, CP-Pharma Handelsgesellschaft GmbH, Germany), ethanol (96% vol, VWR BDH Chemicals, France), ketamine (bela-pharm GmbH & Co. KG, Germany), xylazine (Wirtschaftsgenossenschaft deutscher Tierärzte eG, Germany), kit for measurements of blood alcohol concentration (BAC, containing Alcohol buffer pH 7.4, alcohol oxidase (AOD) and AM1 analyzer, Analox Instruments Ltd, United Kingdom), lidocaine (2%, beta-pharm GmbH & Co KG, Germany), mouse anti-GFP (Invitrogen, Ref: A22230), paraformaldehyde (PFA, Carl Roth GmbH & Co. kg, Germany), rabbit anti-RFP (Rockland, Ref: 600-401-379), saccharin (saccharin sodium salt hydrate, Sigma-Aldrich Chemie GmbH, Germany), Sodium azide (Merk KGaA, Germany).

Method details

Induction of alcohol dependence

Alcohol dependence (Experiments 1 and 2, **Figure 1**) was induced via chronic intermittent ethanol (CIE) vapor exposure (Rimondini et al., 2002; Hirth et al., 2016; Meinhardt et al., 2013). The vapor exposure was conducted into custom-made vapor chambers that could hold four Type-4 cages each. After one week of habituation to the chambers, the vapor exposure procedure was initiated. The rats were exposed to alcohol vapors for 16 hrs per day every day for 7 weeks. The exposure was initiated 4 hrs before the beginning of the active cycle and ended at the end of the active cycle. The ethanol was delivered by dosing pumps into stainless-steel coils (Knauer, Berlin, Germany) that were heated to 70°C and released into the chambers at a rate of 16 L/min, mixed with air. Twice per week, right after the end of the vapor exposure session, 4 animals per chamber were randomly picked to measure BACs. The pump rate was adjusted in order to keep the BACs between 150 and 300 mg/dL throughout the procedure. At the end of the seventh week of exposure, the animals were maintained into abstinence (air

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

exposure only) for one week before initiating the behavioral trainings. Seven hours after the last ethanol vapor exposure cycle each rat was monitored for up to 5 min to assess tremor, tail rigidity, vocalization, abnormal gait, and ventral limb retraction. Severity of each withdrawal sign was ranked as 0 (non-existent), 1 (mild), and 2 (severe) added up and expressed as total withdrawal scores (Macey et al., 1996; Uhrig et al., 2017).

Spatial navigation (T-Maze) task

The spatial navigation task was performed in a custom-made T-Maze. The maze was elevated 50 cm above the floor and had four closed identical arms (50 cm (L) x 10 cm (W) x 50 cm (H). The access to either the north or the south arm was blocked by using a 9.8 cm cuboid. The training and testing protocols for habitual behavior were adapted from Packard & McGaugh (1996) and Yin & Knowlton (2004). On the first day, the animals underwent a habituation session. The access to the north arm of the maze was blocked; each rat was positioned at the end of the south arm and was free to explore the maze for 5 minutes. On the following day, the training started. The access to the north arm was still blocked, on the wall at the end of the west (left) arm a visual cue was positioned (baited arm) while no optic cue was presented at the east (right) arm (unbaited arm). At the end of both side arms a transparent plastic well was attached on the floor of the maze. The well at the west arm always contained 1 ml of 0.2% saccharin solution, while the well at the east arm was always kept empty.

For both Experiment 1 and 3, each training trial started by positioning the rat at the end of the south arm and ended after 5 minutes or after the animal found and consumed the reward, after which the animal was removed from the maze. Each training session consisted of 5 consecutive trials. For each trial, the performance of the animal was scored: A correct choice was considered when the animal walked from the south arm directly into the baited arm and consumed the reward, in all other cases (the animal walked into the baited arm but did not consume the reward or it walked into the unbaited arm) the choice was considered as incorrect. Each animal was trained daily (one session of 5 consecutive trials per day) until 80% performance (four correct choices out of 5 trials) for two consecutive sessions was reached, to ensure the animals properly learned the task. Then a probe session was performed on the following day (Early time point, Figure 1). After the early probe session, the training resumed and was continued daily until a criterion of at least 90% performance over the course of 10

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

consecutive days was reached, to ensure overtraining, after which a second probe session was performed (Late time point, Figure 1).

The probe sessions consisted of 5 consecutive probe trials, each trial was identical to the training trials except that animals were inserted into the north arm of the maze while the south arm was blocked; the animal was removed from the maze immediately after making a choice (baited or unbaited arm) and walked back into the center of the maze. When the animal walked into the baited arm, this was considered as a goal-directed choice. When the animal walked into the unbaited arm it was considered as a habitual choice.

Reward learning (operant conditioning) task

Before initiating the training sessions, all animals from experiments 2 and 3 underwent a 30 minutes habituation session in which no lever and no cue was present, the reward (0.2% saccharin, 13 μl drop) was released into the receptacle at a random interval 60s (RI-60s) schedule.

Following the habituation, the animals from Experiment 2 (**Figure 1**) were trained under fixed ratio 1 (FR-1) schedule until they reached a criterion of at least 50 rewards for 3 consecutive sessions, after which, a progressive ratio 2 (PR-2) test was performed. Based on the breakpoint value, the animals were then pseudo-randomly assigned to the treatment groups (CIE: chronic intermittent ethanol vapor exposed group, air-exposed control group) and the vapor exposure was initiated. One week after the last ethanol vapor exposure, the animals underwent three more sessions under FR1 schedule and were again tested in a PR-2 test. After the last PR test, all animals underwent one session under random interval 5s (RI-15s) schedule, one session under RI-15s schedule, and 6 sessions under random interval 30s (RI-30s) schedule. The training was performed from Mondays to Thursdays. After the 6th day of RI-30s training the animals underwent the devaluation (DV) test with a satiety devaluation procedure of 30 min followed by a 5 minutes extinction test (Early time point). The following week the RI-30 training was resumed for another 4 sessions and the DV test was repeated (Late time point).

For Experiment 3 (**Figure 1**) the animals were also pre-trained for three days under FR-1 schedule, after which they were directly trained under random interval schedule (RI-5s, RI-15s, RI-30s) and tested in a devaluation test at the Early time point only.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

FR-1 pre-training sessions in Experiment 2 ended after 30 minutes, FR-1 pre-training sessions in Experiment 3 ended after 30 minutes or after 50 rewards were earned.

All random interval sessions ended after 30 minutes or after 30 rewards were earned. Animals that did not earn 30 rewards in the last two sessions prior to the DV test were not included for statistical analysis.

During the DV test all animals from Experiment 2 and 3 were put into single cages and brought into a novel environment, where they were given one bottle containing 0.2% saccharin and one bottle containing tap water, and they were allowed to freely drink for 30 minutes. Immediately after the end of the drinking session, all animals were brought into the operant boxes where they underwent a 5 minutes extinction session in which the house light was turned on, the lever was present but no reward was released. At the end of the extinction session all animals were brought back into the single cages and were given another 30 minutes of free drinking to confirm the satiated state.

In Experiment 3, 30 minutes prior to each probe session (spatial navigation) and to the extinction session (reward learning), all animals received a 10 mg/ml CNO ip injection (in saline).

Progressive ratio test

Prior to the beginning and one week in abstinence all animals in experiment 2 were tested in a PR-2 test. This test provides an estimate about the motivation for an animal to acquire the reward. During the test only the active lever was available, this was the same lever used for the FR1 and the RI trainings. In a PR schedule the number of lever presses (n) required for obtaining the reward increased with each reinforcement (1, 2, 4, 6, 8, 10, … n+2) within a session. The session ended after 2 hours or after no lever press was performed for three consecutive minutes. During the test the house light signaled the beginning and the end of the session. The maximum number of lever presses for a single reinforcement within a session was considered as breakpoint.

Alcohol operant training

The animals that completed the reward learning task in experiment 2 were further trained to self-administer 10% ethanol in tap water. The training sessions were identical to the previous saccharin training sessions except that the lever assigned to the ethanol was the opposite

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

lever in the operant chamber. The reward was delivered on the receptacle next to the active lever, on the opposite side compared to the receptacle and the lever used for saccharin training. The animals were trained until they earned at least 50 ^ rewards for three consecutive sessions, after which they were tested in a PR2 test that ended after 2 hours or after 3 minutes of inactivity (see **Figure 1**).

Stereotaxic surgeries

Serotype 8/2 adeno associated viral vectors (ssAAV-8/2) expressing either hM4D(Gi)-mCherry (titer: 7.3 x 10^{12}) or eGFP (titer: 6.9 x 10^{12}) under the control of human synapsin promoter (hSyn1) were used. Surgeries were performed as described in Broccoli et al. (2018) and Meinhardt et al. (2022). Inactivating DREADD virus was injected bilaterally into the pDMS (coordinates: A/P = -0.4; M/L = ± 2; D/V = -4) using 10 µl NanoFil syringes equipped with 33 gauge beveled NanoFil needles (World Precision Instruments, Inc.) at a volume of 0.7 µl and a rate of 0.1 µl per minute. Before activating the micropump (MicroSyringe Pump Controller, Micro 4 ^{TM}, World Precision Instruments, Inc.) for the injection, the needle was first brought into position based on the aimed coordinates and left in place for 5 minutes. After injection, the needle was kept in place for another 10 minutes before slowly withdrawing it.

Four% isoflurane-oxygen mixture was used for induction of anesthesia, 2-2.5% isoflurane-oxygen mixture was used for maintenance (flow rate: 1 L/min). Carprofen (5 mg/kg, sc) and Lidocaine (1% sc, under the scalp) were injected at beginning of the surgery, Carprofen treatment was continued for three days post-surgery.

Perfusions

At the end of the experimental procedures, all rats that received bilateral ssAAV-8/2 injection were perfused. Briefly, rats were deeply anesthetized with a mixture of ketamine-xylazine (100 mg ketamine/kg, 3 mg xylazine /kg, ip), they were then perfused with 150 mL of 1xPBS followed by 100 ml of ice-cold 4% PFA (flow rate: 15 mL/min). The brains were then removed and post-fixed into 4% PFA at 4°C for 48 hours before being transferred into 1 x PBS + 0.1% sodium azide (4°C).

Immunofluorescence

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

For immunofluorescence, consecutive coronal sections (50 µl thickness) were obtained using a vibrating-blade microtome (Leica VT1000 S, Germany) covering the entire dorsal striatum. The slices were then permeabilized into 1% Triton X-100 in 1 x PBS (30 minutes, 4°C), washed three times into 1 x PBS (5 min each, RT), blocked with 10% donkey serum in 1 x PBS (1 hour, RT), and incubated o/n with either 1:1000 mouse anti-GFP (Invitrogen, Ref: A22230) or 1:1000 rabbit anti-RFP (Rockland, Ref: 600-401-379). The slices were then washed three times into 1 x PBS (5 min each, RT), incubated with either 1:500 Alexa Fluor 488 donkey anti-mouse or 1:500 Alexa Fluor 594 donkey anti-rabbit for 1 hour, RT. They were finally washed three times in 1 x PBS (10 min each, RT) before being mounted on glass slides with anti-fade mounting medium. The expression of both eGFP and mCherry was detected by an epifluorescent microscope (Zeiss, Axio Imager M2, Germany).

Data analyses and statistics

For the spatial navigation task, two different variables were used to assess habitual behavior: the number of habitual choices within the first trial of each probe session and the overall choice performance calculated as performance ratio: number of goal-directed choices per session divided by the total number of trials per session.

For measuring habitual behavior in the reward learning task, the cumulative devaluation index was used. This was calculated as follows:

$(baseline_{cumulative} - test_{cumulative}) / (baseline_{cumulative} + test_{cumulative})$

"$Baseline_{cumulative}$" represents the cumulative average number of lever presses performed for each 1-minute bin within the first 5 minutes of the last two sessions prior to the test; "$test_{cumulative}$" represents the number of lever presses performed for each 1-minute bin within the first 5 minutes of the test. Using the cumulative devaluation index it is possible to analyze how the behavioral approach of the animals (goal-directed or habitual) changes over the course of the test.

The difference between groups in the spatial navigation task using the number of habitual choices within the first trial of each session was analyzed via χ^2 test. Habitual behavior in the reward learning task was assessed via one-sample t-test of the devaluation index against 0.

The withdrawal scores were compared between groups via Mann-Whitney U test. The remaining data were analyzed via Repeated Measure Mixed Model (Gueorguieva and Krystal, 2004). The estimation method used was Maximum Likelihood (ML), the covariance matrix was

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

chosen based on the lowest BIC. When applicable, LSD post-hoc test was performed. For statistical analysis IBM SPSS Statistics 27 was used.

Sample size

Experiment 1: 8 Ctrl and 8 CIE animals were used for the spatial navigation task. One animal died during the vapor exposure, therefore the final sample size was 8 Ctrl and 7 CIE.

Experiment 2: For the reward learning task, 26 rats were used in total (12 Ctrl + 14 CIE). Of these, 2 rats died during the vapor exposure, 3 controls and 3 CIE were excluded from the analysis because they did not reach the 30 rewards per session criterion during the random interval training. The final sample size was 12 Ctrl and 9 CIE.

Experiment 3: A total of 28 rats were used, 14 GFP and 14 hM4Di injected rats. Of these, 2 hM4Di rats were excluded due to misplacement of injection site, 3 hM4Di rats did not acquire the lever pressing behavior, 2 Ctrl and 3 hM4Di were excluded from the analysis, because they did not reach the 30 rewards per session criterion during the random interval training. The final sample size was 14 GFP and 12 hM4Di for the spatial navigation task, and 12 Ctrl and 7 hM4Di for the reward learning task.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Results

Study 1: Habit formation in alcohol dependence

Habitual and goal-directed behavior was determined in alcohol dependent and non-dependent rats using spatial navigation (T-Maze, Study 1, Experiment 1) and reward conditioning (operant boxes, Study 1, Experiment 2) tasks with 0.2% saccharin as reward.

Dependence was introduced by chronic intermittent cycles of ethanol vapor exposure (CIE, (Rimondini et al., 2002; Hirth et al., 2016). Average BAC per daily cycle and experiment was 269.72 ± 16.04 mg/dL. Dependence was further confirmed by measurement of withdrawal scores 7 hrs in abstinence after the last vapor exposure cycle. The average total withdrawal score was 4.38 ± 0.36 for CIE animals compared to control rats (1.65 ± 0.22, U = 22.50, z = -4.60, p < .000).

Experiment 1: Spatial Navigation (T-Maze) task

The training on a T-Maze started in one week abstinent rats. Rats had to navigate to a well containing the reward (0.2% saccharin solution) signaled by a visual cue, always located on the left arm of the maze. Each rat was tested individually on different days depending on its performance. For the Early time point each rat was tested in a probe session the following day after making at least 4 correct choices out of five trials (80% performance) for two consecutive sessions. For the Late time point rats were tested after making 10 sessions with an overall performance of at least 90%.

As expected, we found no difference in terms of training performance between groups ($F_{1,25}$ = .02, p = .878) but the performance significantly differed between time points ($F_{2,48}$ = 117.5, p = .000; **Figure 2A**). The post-hoc test indicated that prior to the early probe session the animals properly learned the task as their performance increased compared to the beginning of the training (Start vs Early: p= .000), and prior to the Late probe session the animals improved their performance even more (Early vs Late: p=.001).

The results of the probe sessions, considering the first trial of each session (**Figure 2B**) and the overall performance for each session (**Figure 2C**) indicate that CIE animals displayed an increased habitual behavior response compared to controls. Based on the first trials, CIE animals made significantly more habitual turns compared to controls at the Late time point (χ^2 = 5.5, p = .019), 86% CIE vs 25% Controls. From the Early to the Late time point the number of CIE animals that made habitual turns increased (from 43% to 86%) but such increase was

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

not significant (χ^2 = 2.8, p = .094). However, control animals did not show any change in their choices with overtraining as at both time points only two animals made a habitual turn (**Figure 2B**). Based on the overall performance, measured as correct choice ratio per time point, the CIE animals showed a reduced performance, hence an overall increased number of habitual turns, specifically at the Late time point after overtraining (LSD: p = .013; Group*Time interaction: $F_{1,30}$ = 8.1, p = .008). Control animals showed an increased overall performance from the Early to the Late probe session (p=.000) which was not the case for the CIE group (p=1).

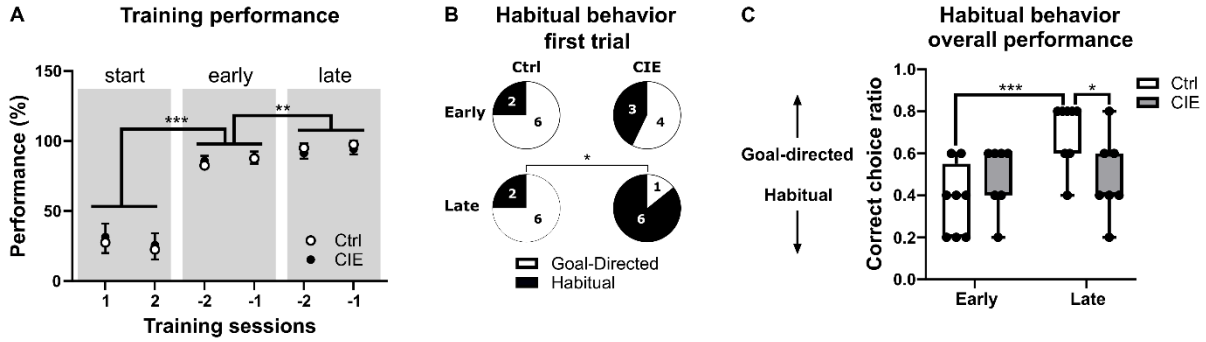


Figure 2. Pronounced habitual behavior in alcohol dependent rats (Study 1, Experiment 1). A. Training performance in percentage during the first two sessions, the last two sessions before the Early probe session, and the last two sessions before the Late probe session. Data are shown as mean ± SEM **B.** Pie chart showing number of habitual choices during the first trial of each probe session. **C.** Overall performance including all five trials of Early and Late probe session. Data for correct choice ratio are shown as box and whisker plot (median, first and third quartile range). Data in **B** were analyzed by χ^2 test. Data in **A** and **C** were analyzed by Repeated Measure Mixed Model, *p < 0.05, **p < 0.01, ***p < 0.001. CIE, cyclic intermittent alcohol vapor exposed rats; Ctrl, control rats. For Details, see Material & Methods.

Experiment 2: Reward learning (operant conditioning) task

In order to study habitual behavior in the context of reward learning, we first trained male Wistar rats to self-administer 0.2% saccharin under a FR-1 schedule until they obtained a minimum of 50 rewards for three consecutive sessions after which their motivation for the reward was tested in a progressive ratio PR-2 test. When analyzing the number of lever presses during the FR1 sessions prior to each PR-2 test, we found a significant Group x Time x Session interaction ($F_{2,168}$ = 5.1, p=.007). The post-hoc test showed no difference before initiating the vapor exposure (**Figure 3A**).

We used the breakpoint (BP) value obtained from this test to pseudo-randomly assign the animals into the two treatment groups (CIE and Ctrl) in order to control for the motivational

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

level toward the reward. The mean BP for the two groups was 25.1 ± 2.1 and 26.2 ± 2.0 (mean ± SEM) for the prospective Ctrl and the CIE groups, respectively (t_{19} = -.3, p = .731).

All animals were then tested again for 0.2% saccharin (on the same lever and receptacle) 1 week after the last cycle of ethanol vapor exposure. The CIE group showed significantly increased responding for saccharin in the last two sessions (2nd session: p=.002, 3rd session: p=.028). To confirm the alcohol dependent phenotype, at the end of the whole experiment animals were trained to self-administer 10% ethanol (on the opposite lever and receptacle). The BP obtained from the PR-2 tests for 0.2% saccharin and 10% alcohol, indicated that CIE animals displayed a higher motivation for both rewards compared to controls ($F_{1,40}$ =13.3, p = .001), but both groups strongly preferred saccharin over alcohol (effect of reward: $F_{1,40}$ = 16.0, p=.000) (**Figure 3B**).

After the second PR-2 test for saccharin, habitual behavior was induced by training the rats for one session under RI-5 schedule, one session under RI-15 schedule and for 10 sessions under RI-30 schedule. Despite the higher motivation for saccharin observed in the CIE animals, the two groups did not differ significantly in terms of lever presses during the Random Interval (RI) training ($F_{9,93}$ = 3.8, p = .081) (**Figure 3C**).

Importantly, also the baselines used for calculating the devaluation index did not differ between groups, as indicated by the lack of a Group effect ($F_{1,28}$ = .1, p = .751) or any interaction involving this factor.

Habitual and goal-directed behavior was tested after six RI-30s sessions (Early time point) and after four additional RI-30s sessions (Late time point). The test session (5 min) was performed under extinction condition and commenced immediately after 30 minutes of satiety devaluation. When analyzing the cumulative devaluation index of the two groups, we found a significant Group x Timepoint interaction ($F_{1,135}$ = 5.8, p = .018). The post-hoc test revealed that at the Early time point CIE animals had significantly lower devaluation index compared to controls (p = .011) while the groups did no longer show any difference at the Late time point (p = .432) (**Figure 3D**). Moreover, control animals significantly reduced their devaluation index from Early to Late time point (p = .000), while CIE animals did not change between the two time-points (p = .494) (**Figure 3D**). Finally, we found a significant main effect of Bin ($F_{4,52}$ = 10.1, p = .000) with no interaction involving this factor, indicating that both groups, at both

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

time points progressively became aware of the lack of reward during the test, correcting their behavior (reducing the response rate compared to baseline) as time progressed.

In order to evaluate at what point of the devaluation (DV) test the animals actually displayed habitual behavior, we compared the cumulative devaluation index of the groups for each minute of the test against the hypothetical devaluation index of 0, indicating full habitual behavior, via one-sample t-test. During the Early time point, the control group showed goal-directed behavior throughout the entire test as the devaluation index was significantly different from 0 at all minutes. During the first minute of the test the CIE group displayed habitual behavior but goal-directed behavior throughout the rest of the DV test. During the Late time point, both control and CIE groups displayed habitual behavior during the first two minutes of the test but shift towards goal-directed behavior afterwards (**Figure 3D**, **Table 1**).

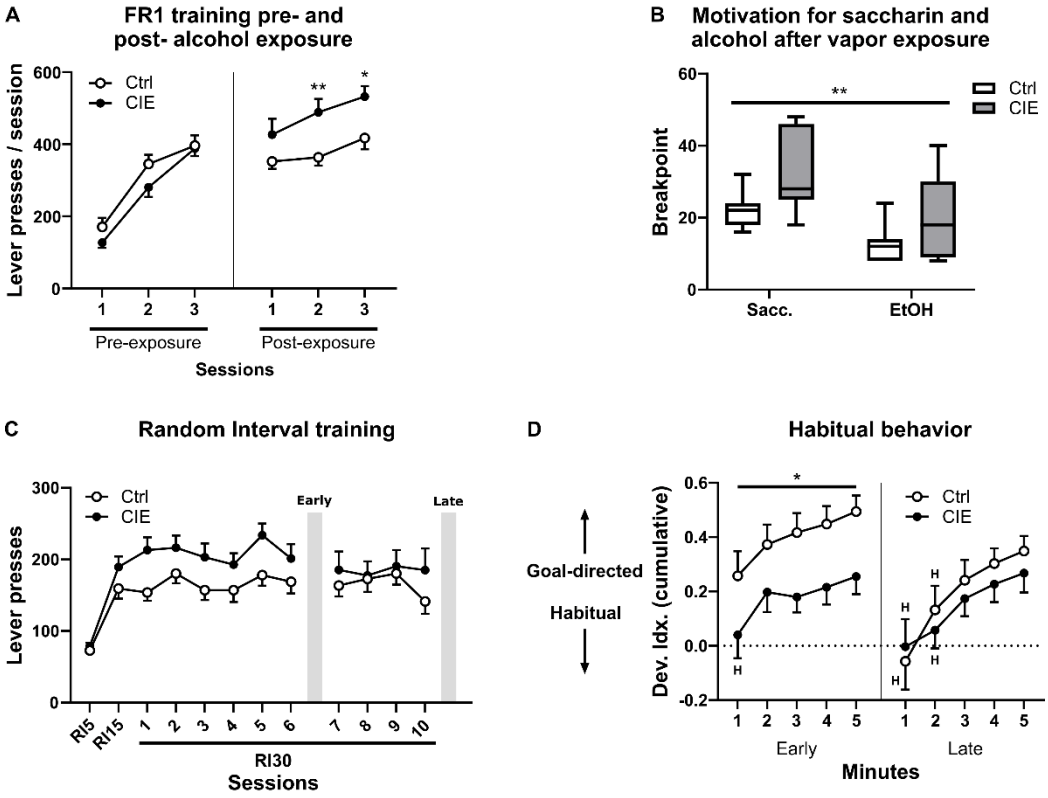


Figure 3. Alcohol dependent rats displayed habitual behavior for reward conditioning (Study 1, Experiment 2). A. Lever presses during FR1 training before and one week after the last cycle of ethanol vapor exposure. **B.** Breakpoint for 0.2% saccharin and 10% ethanol. **C.** Lever presses during random interval training. **D.** Devaluation index (Dev Idx) over the course of the extinction test for both early and late time-points. A dev Idx of 0 indicates full habitual behavior; a Dev Idx of 1 indicates full goal-directed behavior. Data points highlighted with "H" are indicating habitual behavior. Data in graphs **A**, **C** and **D** are shown as mean ± SEM. Data in graph **B** are shown as box and whisker plot (median, first and third quartile range). Data were analyzed via Repeated Measure Mixed Model. *p < 0.05, **p < 0.01, ***p < 0.001. CIE, cyclic intermittent alcohol vapor exposed rats; Ctrl, control rats. For Details, see Material & Methods.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Study 2: Role of pDMS in spatial navigation and reward learning

We next used a DREADD-based approach to inhibit the pDMS in naïve Wistar rats and evaluated the effects on goal-directed and habitual behavior in both the spatial navigation and reward learning tasks (Experiment 3). We first trained the animals on the spatial navigation task, starting three weeks after the end of the last surgery in order to allow full expression of the transgenes, we then trained the same animals on the reward learning task starting one week after the end of the spatial navigation task **(Figure 1)**. The quality of the injection of the virus was confirmed by immunohistochemistry for mCherry (expressing hM4Di, **Figure 4 A-B**) and GFP control virus within the pDMS region (**Figure 4 A**). Rats with misplaced virus injections were excluded from the entire analysis.

Spatial Navigation (T-Maze) task

As expected, we found no between groups difference ($F_{1,83}$ = .5, p = .495) in performance during the training, but both groups progressively improved ($F_{2,89}$ = 205, p = .000) going from the beginning of the training to prior the first probe session (p = .000) and from here to prior the second probe session (p = .000) (**Figure 4C**).

When analyzing the behavioral approach (habitual vs goal-directed) during the two probe sessions we did not find any difference between groups based on the first trial of each session (**Figure 4D**) as the hMD4i animals only made less than 20% more habitual turns during both time points, compared to GFP controls (early: χ^2 = .2, p = .671; late: χ^2 = .2, p = .619). However, when analyzing the overall performance considering all five trials of each probe session (**Figure 4E**) we found that hM4Di-treated rats displayed an overall increased number of habitual turns for both timepoints compared to GFP-expressing rats, as indicated by the presence of a statistically significant effect of Group ($F_{1,52}$ = 9.9, p = .003) but lack of a Timepoint effect ($F_{1,52}$ = .6, p = .424) and interaction ($F_{1,52}$ = 1.3, p = .260) effect between the two factors. However, the post-hoc test indicated a significant difference between the groups at the Early (p = .004) but not at the Late (p = .163) time point.

Reward learning (operant conditioning) task

After the end of the spatial navigation task, both groups were trained to self-administer 0.2% saccharin using the same FR1 pre-training used for Experiment 2 except that in this case the animals could not earn more than 50 rewards per session. This was done because we did not test the motivation of the animals in a progressive ratio test, because we had no reason to

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

assume a change in the rewarding properties by the experimental procedure, and in order to reduce the risk for overtraining of the lever-pressing behavior. Once the random interval training started, all conditions were identical to the ones used in Experiment 2 but the animals were only tested once, after 6 RI-30 sessions, when goal-directed behavior would be expected. During RI training (**Figure 4F**) we found no difference in lever presses between the hM4Di and GFP groups ($F_{1,26}$ = .7, p = .417).

For assessing habitual behavior, rats were treated with CNO (10 mg/Kg, ip.) 30 minutes before the test. When analyzing the behavior of the animals during the extinction test (**Figure 4G**), performed after reward satiety devaluation, we found a significant Group x Bin interaction ($F_{4,76}$ = 5.7, p = .000) with the post-hoc test indicating that hM4Di-expressing rats had a significantly lower devaluation index, compared to controls, specifically during the first minute of the test (p = .011). Also in this case both groups corrected their behavior as the test progressed, as indicated by the significantly higher devaluation index during the first minute of the test and at the end of the test (Ctrl: p = .020; hM4Di: p = .000). Based on the one-sample t-test analysis of the devaluation index against 0, the hM4Di-expressing animals displayed habitual behavior during the first 4 minutes of the test, while control animals consistently displayed goal-directed behavior (**Figure 4G**, **Table 2**).

We also analyzed the time required to make the first lever press response (first response latency, FRL) after the initiation of the extinction test and we compared it to the baseline FRL (**Figure 4H**), measured as the FRL averaged between the last two training sessions prior to the test. We found a significant Group x Time (baseline vs test) interaction ($F_{1,19}$ = 5.1, p = .036) with the post-hoc test revealing that the control group took a longer time to make the first response during the test as compared to their baseline (p = .015), while the hM4Di animals did not (p = .438). Moreover, hM4Di-treated rats had a significantly lower FRL during the test compared to the response latency of the control animals during the test (p = .016).

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

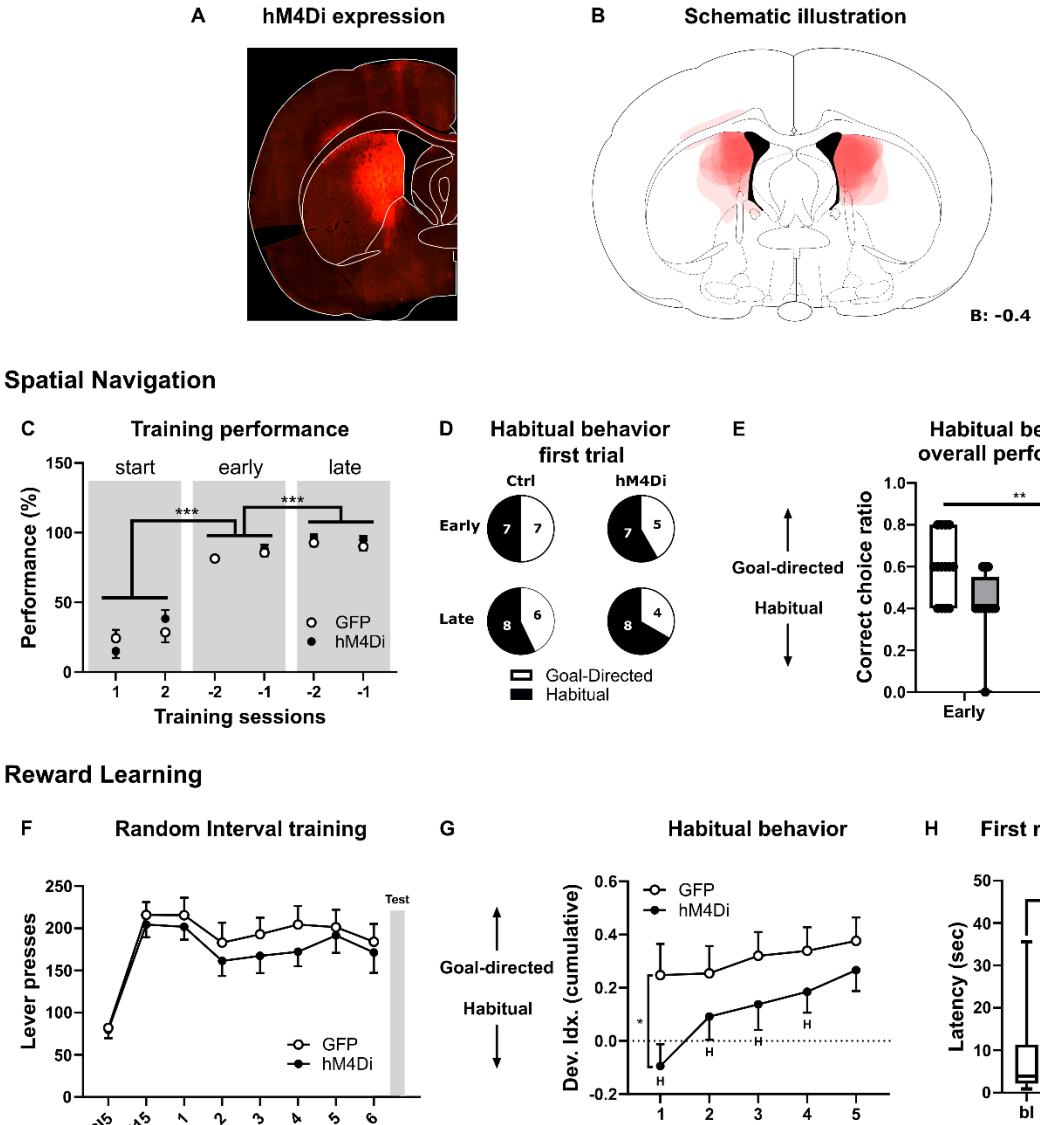


Figure 4. Inactivation of DMS led towards habitual behavior in naïve rats (Experiment 3). A. Coronal brain section showing the expression of hM4Di-mCherry within the pDMS. **B.** Schematic illustration showing quality of bilaterally injected hM4Di-mCherry virus at Bregma level: -0.4 mm. **C.** Training performance in percentage during the first two sessions, the last two sessions before the early probe session, and the last two session before the last probe session. **D.** Number of habitual choices made during the first trial for each probe session. **E.** Overall performance when considering all five trials for each probe session. **F.** Lever presses during random interval training. **G.** Devaluation index (Dev. Idx) over the course of the extinction test. A dev. Idx. of 0 indicates full habitual behavior, a Dev. Idx of 1 indicates full goal-directed behavior. Data points that were non-significantly different from 0 based on one-sample t-test are highlighted with "H". Data in graphs **C**, **F**, and **G** are shown as mean ± SEM. Data in graph **D** are shown as pie chart. Data in graphs **E** and **H** are shown as box and whisker plot (median, first and third quartile range). Data from graph D were analyzed via χ^2 test. Data from remaining graphs were analyzed via Repeated Measure Mixed Model. *p < 0.05, **p < 0.01, ***p < 0.001. DREADD-mediated inactivation of DMS was achieved by ip injection of CNO (Clozapine-N-oxide, 10mg/kg) 30 minutes prior each test. For Details, see Material & Methods.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

In this study we provide evidence for a generalized habitization in alcohol dependent rats. We investigated habitual responses towards a sweet palatable reward in both tasks (spatial navigation and reward conditioning) and estimated to which extent such habitual responses are changed in abstinent rats. We found that most dependent rats displayed increased habitual responses at an earlier time point of learning, while non-dependent rats still showed a more goal-directed response. Chemogenetic inactivation of the pDMS in alcohol-naïve animals increases habitual tendencies and suggests alcohol-induced damage to this region in post-dependent rats. In conclusion, we confirm the critical role of DMS in maintaining goal-directed behavior, which seems to be to some extend compromised in AUD.

Assessment of habitual behavior in the spatial navigation task

In study 1 we used post-dependent (CIE) and non-dependent rats for assessment of habitual behavior using a T-Maze navigation task. Both dependent and non-dependent controls displayed goal-directed responses at early time point of testing. Dependent rats switched to pronounced habitual responses after overtraining, while controls remained goal-directed (**Figure 2B-C**). The results imply that dependent rats developed habitual responses in a shorter time interval than controls, and overtraining was not effective in controls. A shift to habitual behavior in controls can be probably achieved after longer time of overtraining.

Navigation is an adaptive task in which stimulus-response associations or habits are incrementally acquired. Initially, a strategy based on a cognitive map (spatial map or `model-based`) are applied, while a more stereotyped response strategy (egocentric or 'model-free') can dominate after repeated training. Lesions of the dorsal striatum disrupt the latter strategy (Lee et al., 2008). AUD results in long-term damage of cortical regions, especially of the PFC but also hippocampus (Zahr et al., 2011; Meinhardt et al., 2013; Heilig et al., 2017; Meinhardt et al., 2021), suggesting that human subjects may preferentially adopt a striatal-based response strategy in e.g. 4-on-8 virtual maze or two step sequential Markov decision task.

The testing approach that we used in the spatial navigation task has one limitation: the five consecutive test trials during the early time point could influence the performance of the animals during the test performed at the late time point as the animals had the possibility to

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

experience the test condition (inverted maze) multiple times, which would explain the increased overall performance in the control group from the early to the late time point.

Assessment of habitual behavior in the operant conditioning task

Naïve rats were trained to self-administer 0.2% saccharin in operant chambers. Based on their motivation to self-administer saccharin, rats were assigned into alcohol dependent and control groups. After introduction of dependence by CIE, rats were trained for habitual behavior under a random interval (RI)-30 schedules that facilitated habitual behavior.

At the early and late time points rats were assessed for habitual behavior using a devaluation test procedure followed by an extinction test. At the early time point dependent rats displayed more habitual behavior (indicated by lower devaluation index) compared to controls. At the late time point both dependent and control rats showed a reduced devaluation index, indicating increased habitual behavior compared to the early time point (**Figure 3D**). Important to note, the dependent rats showed a higher motivation in the PR test for both saccharin and 10% alcohol during and at the end of the experiment, respectively, thus confirming the dependence-like phenotype (**Figure 3B**). An association between AUD and a sweet-liking phenotype has been in fact described in humans (Kampov-Polevoy et al., 2003; Pepino and Mennella, 2005, 2007; Bouhlal et al., 2018) and in high ethanol preferring rodents (Leventhal et al., 1995; Pecina and Berridge, 2005).

In the reward learning task, habitual behavior is typically tested in rodents after reward devaluation with an extinction test performed following a training under a random interval (RI) or random ratio (RR) schedule for a given number of sessions. Reward devaluation procedures, which can be accomplished via either specific satiety or conditioned-taste aversion, are typically used to dissociate habitual from goal-directed behavior (Dickinson, 1985; Barker and Taylor, 2014; McKim et al., 2016; Robbins and Costa, 2017). In this study the response to devaluation was assessed by pre-feeding the rats with saccharin solution prior to a 5-min extinction test. This test is usually short and it is performed in extinction conditions in order to avoid the animals experiencing the reward while in the devalued state and reduce extinction learning, this is particularly important if the animals are tested multiple times, as in our case. In this context, habitual behavior is defined as insensitivity to reward devaluation, meaning that animals that are displaying habitual behavior will press at a similar rate while in

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the devalued state (during the test) and when non-devalued (control condition/group) (Corbit et al., 2012). The length of the extinction test however is often different between studies that use this approach, lasting anywhere from 3 minutes to 30 minutes (Corbit et al., 2012; Mangieri et al., 2012; Clemens et al., 2014; Furlong et al., 2017; Iguchi et al., 2017). It is not clear if animals that show insensitivity to devaluation in a 3 minutes test would still show it in a 5 minutes test, or if animals found to be sensitive to the devaluation procedure in a 5 minutes test would also show sensitivity in a shorter test. Therefore, we have analyzed the data obtained from the devaluation-extinction test by calculating the devaluation index (Renteria et al., 2018; Towner and Spear, 2021) in cumulative form. The cumulative devaluation index is an estimate of how the behavior of the animals changes throughout the test.

In our experiment, following reward devaluation at the early time point, both dependent and control rats exhibited reductions in lever pressing, suggestive of a goal-directed response pattern. However, using the cumulative devaluation index we could show that dependent rats not only displayed overall increased habitual tendencies compared to controls (overall reduced devaluation index), but also that during the first minute of the test, dependent rats exhibited full habitual behavior, while control animals consistently exhibited goal-directed behavior throughout the test (**Figure 3D**). At the late time point, both dependent and control rats showed full habitual behavior during the first 2 minutes of the test while no longer differing in overall devaluation index, indicative of an effective overtraining.

A shift from goal-directed to habitual behavior is in line with the growing literature on dependence induced behavioral consequences, as there is evidence showing that chronic alcohol exposure can lead to increased habits also for taking behavior of a natural reward. Corbit and colleagues (2012) showed for example that non-contingent alcohol drinking performed during the course of operant training for a natural reward can induce insensitivity to outcome devaluation, while Towner and Spear (2021) showed that chronic intermittent ethanol intake, performed via oral gavage in late adolescent rats, also induces outcome insensitivity, indicating habitual behavior. Similarly, chronic intermittent alcohol vapor exposure has been shown to induce habitual behavior for a food reward in mice (Renteria et al., 2018). Interestingly, a recent study reported evidence of increased habitual behavior for

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

taking of a non-alcohol reward also after a single acute dose of ethanol, however in alcohol-preferring mice (Houck and Grahame, 2018).

This body of literature adds up to the evidence that responding for alcohol transitions towards habitual behavior faster compared to responding for a natural reward (Dickinson et al., 2002; Corbit et al., 2012; Mangieri et al., 2012). Moreover, previous research demonstrated that food habits form more quickly in females than in males, but the opposite is true for alcohol habit formation (Quinn et al., 2007; Barker et al., 2010).

Role of pDMS in spatial navigation and reward learning

Here we used a DREADD based approach to inhibit the posterior part of the dorsomedial striatum (pDMS) in naïve Wistar rats according to Yu et al., (2021). Inactivating DREADDs (hM4Di-mCherry) and GFP-expressing control viruses were bilaterally injected into the pDMS and effects on habitual behavior were assessed using both spatial navigation and operant tasks. We used a within subject design, with the same animals being tested for both behavior paradigms. In the T-Maze navigation task the DREADDs-mediated inactivation of the pDMS did induce an overall increased number of habitual choices at both time points, mostly pronounced during the early time point (**Figure 4E**). In the operant paradigm we could further demonstrate, that DREADD-treated rats displayed habitual behavior during the first 4 minutes of the devaluation test after short training, when goal-directed behavior would be expected, while the GFP-virus expressing control rats consistently displayed goal-directed behavior throughout the entire session. Thus, inactivation of pDMS by DREADDs led to habitual responses in both behavior paradigms, confirming the critical involvement of pDMS to maintain goal-directed behavior. Other studies found similar results after disengagement of the DMS and the formation of habitual behavior (Bassett et al., 2015; Kupferschmidt et al., 2017; Yu et al., 2021).

During the early goal-directed actions the DMS is mainly involved, while during the late habitual response, the DMS becomes weakened (Kupferschmidt et al., 2017) and the DLS instead becomes strengthened (Thorn et al., 2010). Thus, DMS and DLS seem to compete for control in the acquisition of habitual action sequences (Turner et al., 2022). Taken together, our findings support an opponent relationship of DMS and DLS in which loss of DMS function enhances the formation of habitual actions.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

In general, instrumental responding (using an operant box) and spatial navigation (using a T-Maze) share important features. Firstly, two stages can be identified in the formation of a well-established action sequence: an early acquisition phase, characterized by a steep learning curve, and a phase of slow gradual consolidation, in which the behavior is optimized and becomes less susceptible to external influences. Secondly, the transition from early acquisition to later consolidation involves a shift from ventral to dorsal striatal regions. Control of motivated behaviors is initiated in the ventral striatum. Clinical studies suggest that during the transition from reward or goal-directed to habit-driven behavior, control is shifted towards dorsal parts of the striatum (Vollstadt-Klein et al., 2010). The most direct evidence for this process comes from a series of studies, which directly showed ventral striatum activation to alcohol challenge in light social drinkers, attenuation of this signal in heavy social drinkers (some of whom met criteria for diagnosis), and lack of this response in hospitalized treatment-seeking patients (Gilman et al., 2008; Gilman et al., 2012; Spagnolo et al., 2014).

There are human fMRI studies showing a reduced sensitivity to devaluation after extensive instrumental training (Tricomi et al., 2009; Sjoerds et al., 2013), but other studies did not find evidence for habitual responses in human subjects after prolonged training (de Wit et al., 2009; Hogarth et al., 2019; Hogarth and Field, 2020; Luijten et al., 2020). Acute treatment of alcohol had been shown to impair goal-directed action (Hogarth et al., 2012). First evidence for a shift from goal-directed to habitual behavior in AUDs came from Sjoerds et al. (2013) using an instrumental learning task. Later on, another study compared AUD patients with healthy controls using the two-stage Markov decision task and found a reduction of goal-directed performance in AUDs, but no difference in habitual responding (Sebold et al., 2014). This finding may highlight rather the possibility of general cognitive impairments in AUDs that impact of habitual action control. Other studies using the same task compared a variety of compulsive disorders including abstinent AUD patients (Voon et al., 2015). Similarly to Sebold et al. (2014), they did not observe whole-sample differences between AUD patients and healthy controls. However, the study showed an association of early abstinence with greater habitual action control which shifts towards more goal-directed control with longer abstinence. Another study using a web-based version of the two-stage Markov decision task with 1413 participants showed a significant decrease in goal-directed control in participants with self-reported AUD (using AUDIT) which were attributed to impairment in goal-directed

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

control to a broader range of trans-diagnostic factors, which they grouped as compulsive behavior and intrusive thought (Gillan et al., 2016). A recent longitudinally designed study with a sample size of 188 young social drinkers found that the amount of regular alcohol consumption had no significant impact on action control and habits (Nebe et al., 2018). All these contradictory findings contrast the results on habitual overreliance in AUD patients (Tricomi et al., 2009; Sjoerds et al., 2013), which might be due to the different methodologies, sample features and relatively smaller sample sizes in the majority of studies.

In our animals, habitual behavior was not detected in rats with low withdrawal scores (data not shown) or least AUD-like behavior in terms of extensive homecage drinking (Smeets et al., 2022). Thus, a behavior heterogeneity of the AUD-like phenotype, such as e.g. amount of alcohol consumption/self-administration, number of lever presses/motivation, withdrawal severity, sweet-liking phenotype (Bouhlal et al., 2018), seem to be important output factors for assessment of habitual behavior in rodents. In addition, we observed variability and low effects sizes of the observed effects. A large individual variability in compulsivity has been observed in animals self-administrating drugs such as cocaine and alcohol (Giuliano et al., 2019; Siciliano et al., 2019).

Further potential limitations of the study were that we did not consider habitual behavior for alcohol reward and other aspects of the AUD-like phenotype such as e.g. loss of control for aversive-/footshock-resistant responding for alcohol- or saccharin reward. Finally, it is important to identify molecular mechanism of habitual behavior for natural rewards to better understand how this behavior developed in comparison with alcohol rewarding effects.

In summary, we conclude that dependence induced a generalized habitual response in rats for both spatial navigation and reward conditioning tasks. Inactivation of pDMS led to habitual responses, similar as seen in our dependent rats. Thus, pDMS is critical to maintain goal-directed behavior; modulation of its activity will lead to changes in habitual responses.
Based on our findings and a growing body of evidence we conclude a more minor role of habit formation in regard to AUD-like behavior. However, our data provide potential implications for understanding how dorsal striatal dysfunction contributes to AUD, and thus may provide new therapeutic strategies in support for regaining control over ‘bad habits’.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Acknowledgments

Financial support for this work was provided by the Bundesministerium für Bildung und Forschung (BMBF) funded SysMedSUDs consortium (FKZ: 01ZX1909A), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation—Project-ID: ME 5279/3-1 and Project-ID 402170461—TRR 265 (Heinz et al., 2020)) as well as the Hetzler Foundation for Addiction Research.

Author Contributions

ACH, RS, and WHS were responsible for study design and procured study funding. FG, AH, MC, MMF, and MM performed the behavioral experiments. FG analyzed data. FG, RS, WHS, and ACH wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Tables

Table 1. One-sample t-test against 0 of the devaluation index for each cumulative minute of the extinction tests at early and late time points (Study 1, Experiment 2). Non-significant results indicate habitual behavior (highlighted in bold).

Time point	Minutes (cumulative)	Controls			CIE		
		Mean ± SEM	t (df=11)	p-value	Mean ± SEM	t (df=8)	p-value
Early	1	0.26 ± 0.09	2.854	0.016	0.04 ± 0.09	0,467	**0.655**
	2	0.37 ± 0.07	5.075	0.000	0.20 ± 0.07	2,688	0.028
	3	0.42 ± 0.07	5.842	0.000	0.18 ± 0.06	3,186	0.013
	4	0.45 ± 0.07	6.761	0.000	0.22 ± 0.06	3,411	0.009
	5	0.49 ± 0.06	8.329	0.000	0.26 ± 0.06	3,954	0.004
Late	1	-0.06 ± 0.10	-0.546	0.596	-0.00 ± 0.10	-0.031	0.976
	2	0.13 ± 0.09	1.495	0.163	0.06 ± 0.07	0.869	0.410
	3	0.24 ± 0.07	3.229	0.008	0.17 ± 0.07	2.672	0.028
	4	0.30 ± 0.06	5.427	0.000	0.23 ± 0.07	3.445	0.009
	5	0.35 ± 0.06	6.337	0.000	0.27 ± 0.07	3.766	0.005

Table 2. One-sample t-test against 0 of the devaluation index for each cumulative minute of the DV test at the Early time point (Study 2, Experiment 3). Non-significant results indicate habitual behavior (highlighted in bold).

Minutes (cumulative)	Ctrl			hM4Di		
	Mean ± SEM	t (df=11)	p-value	Mean ± SEM	t (df=7)	p-value
1	0.26 ± 0.11	2.252	0.046	-0.09 ± 0.08	-1.149	**0.294**
2	0.25 ± 0.10	2.481	0.031	0.09 ± 0.09	1.047	**0.335**
3	0.32 ± 0.09	3.584	0.004	0.14 ± 0.10	1.438	**0.201**
4	0.34 ± 0.09	3.810	0.003	0.19 ± 0.08	2.340	0.058
5	0.38 ± 0.09	4.291	0.001	0.27 ± 0.08	3.379	0.015

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35:48-69.

Balleine BW, Dezfouli A (2019) Hierarchical Action Control: Adaptive Collaboration Between Actions and Habits. Front Psychol 10:2735.

Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. J Neurosci 27:8161-8165.

Balleine BW, Liljeholm M, Ostlund SB (2009) The integrative function of the basal ganglia in instrumental conditioning. Behav Brain Res 199:43-52.

Barker JM, Taylor JR (2014) Habitual alcohol seeking: modeling the transition from casual drinking to addiction. Neurosci Biobehav Rev 47:281-294.

Barker JM, Torregrossa MM, Arnold AP, Taylor JR (2010) Dissociation of genetic and hormonal influences on sex differences in alcoholism-related behaviors. J Neurosci 30:9140-9144.

Bassett DS, Yang M, Wymbs NF, Grafton ST (2015) Learning-induced autonomy of sensorimotor systems. Nat Neurosci 18:744-751.

Belin D, Belin-Rauscent A, Murray JE, Everitt BJ (2013) Addiction: failure of control over maladaptive incentive habits. Curr Opin Neurobiol 23:564-572.

Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ (2009) Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. Behav Brain Res 199:89-102.

Bergstrom HC, Lipkin AM, Lieberman AG, Pinard CR, Gunduz-Cinar O, Brockway ET, Taylor WW, Nonaka M, Bukalo O, Wills TA, Rubio FJ, Li X, Pickens CL, Winder DG, Holmes A (2018) Dorsolateral Striatum Engagement Interferes with Early Discrimination Learning. Cell Rep 23:2264-2272.

Bouhlal S, Farokhnia M, Lee MR, Akhlaghi F, Leggio L (2018) Identifying and Characterizing Subpopulations of Heavy Alcohol Drinkers Via a Sucrose Preference Test: A Sweet Road to a Better Phenotypic Characterization? Alcohol Alcohol 53:560-569.

Broccoli L, Uhrig S, von Jonquieres G, Schonig K, Bartsch D, Justice NJ, Spanagel R, Sommer WH, Klugmann M, Hansson AC (2018) Targeted overexpression of CRH receptor subtype 1 in central amygdala neurons: effect on alcohol-seeking behavior. Psychopharmacology (Berl) 235:1821-1833.

Clemens KJ, Castino MR, Cornish JL, Goodchild AK, Holmes NM (2014) Behavioral and neural substrates of habit formation in rats intravenously self-administering nicotine. Neuropsychopharmacology 39:2584-2593.

Connor JP, Haber PS, Hall WD (2016) Alcohol use disorders. Lancet 387:988-998.

Corbit LH, Janak PH (2016a) Habitual Alcohol Seeking: Neural Bases and Possible Relations to Alcohol Use Disorders. Alcohol Clin Exp Res 40:1380-1389.

Corbit LH, Nie H, Janak PH (2012) Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol Psychiatry 72:389-395.

de Wit S, Corlett PR, Aitken MR, Dickinson A, Fletcher PC (2009) Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. J Neurosci 29:11330-11338.

Dickinson A (1985) Actions and habits: The development of behavioural autonomy. Phil Trans R Soc Lond B 308:67-78.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Dickinson A, Wood N, Smith JW (2002) Alcohol seeking by rats: action or habit? Q J Exp Psychol B 55:331-348.

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

Everitt BJ, Robbins TW (2016) Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. Annu Rev Psychol 67:23-50.

Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW (2008) Review. Neural mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and addiction. Philos Trans R Soc Lond B Biol Sci 363:3125-3135.

Furlong TM, Supit AS, Corbit LH, Killcross S, Balleine BW (2017) Pulling habits out of rats: adenosine 2A receptor antagonism in dorsomedial striatum rescues meth-amphetamine-induced deficits in goal-directed action. Addict Biol 22:172-183.

Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND (2016) Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. Elife 5.

Gilman JM, Ramchandani VA, Crouss T, Hommer DW (2012) Subjective and neural responses to intravenous alcohol in young adults with light and heavy drinking patterns. Neuropsychopharmacology 37:467-477.

Gilman JM, Ramchandani VA, Davis MB, Bjork JM, Hommer DW (2008) Why we like to drink: a functional magnetic resonance imaging study of the rewarding and anxiolytic effects of alcohol. J Neurosci 28:4583-4591.

Giuliano C, Belin D, Everitt BJ (2019) Compulsive Alcohol Seeking Results from a Failure to Disengage Dorsolateral Striatal Control over Behavior. J Neurosci 39:1744-1754.

Gueorguieva R, Krystal JH (2004) Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. Arch Gen Psychiatry 61:310-317.

Heilig M, Barbier E, Johnstone AL, Tapocik J, Meinhardt MW, Pfarr S, Wahlestedt C, Sommer WH (2017) Reprogramming of mPFC transcriptome and function in alcohol dependence. Genes Brain Behav 16:86-100.

Heinz A et al. (2020) Addiction Research Consortium: Losing and regaining control over drug intake (ReCoDe)-From trajectories to mechanisms and interventions. Addict Biol 25:e12866.

Hirth N, Meinhardt MW, Noori HR, Salgado H, Torres-Ramirez O, Uhrig S, Broccoli L, Vengeliene V, Rossmanith M, Perreau-Lenz S, Kohr G, Sommer WH, Spanagel R, Hansson AC (2016) Convergent evidence from alcohol-dependent humans and rats for a hyperdopaminergic state in protracted abstinence. Proc Natl Acad Sci U S A 113:3024-3029.

Hogarth L, Field M (2020) Relative expected value of drugs versus competing rewards underpins vulnerability to and recovery from addiction. Behav Brain Res 394:112815.

Hogarth L, Attwood AS, Bate HA, Munafo MR (2012) Acute alcohol impairs human goal-directed action. Biol Psychol 90:154-160.

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. Eur J Neurosci 50:2513-2525.

Houck CA, Grahame NJ (2018) Acute drug effects on habitual and non-habitual responding in crossed high alcohol preferring mice. Psychopharmacology (Berl) 235:2167-2175.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Iguchi Y, Lin Z, Nishikawa H, Minabe Y, Toda S (2017) Identification of an unconventional process of instrumental learning characteristically initiated with outcome devaluation-insensitivity and generalized action selection. Sci Rep 7:43307.

Kampov-Polevoy AB, Garbutt JC, Khalitov E (2003) Family history of alcoholism and response to sweets. Alcohol Clin Exp Res 27:1743-1749.

Koob GF (2021) Drug Addiction: Hyperkatifeia/Negative Reinforcement as a Framework for Medications Development. Pharmacol Rev 73:163-201.

Kupferschmidt DA, Juczewski K, Cui G, Johnson KA, Lovinger DM (2017) Parallel, but Dissociable, Processing in Discrete Corticostriatal Inputs Encodes Skill Learning. Neuron 96:476-489 e475.

Lee AS, Duman RS, Pittenger C (2008) A double dissociation revealing bidirectional competition between striatum and hippocampus during learning. Proc Natl Acad Sci U S A 105:17163-17168.

Leventhal L, Kirkham TC, Cole JL, Bodnar RJ (1995) Selective actions of central mu and kappa opioid antagonists upon sucrose intake in sham-fed rats. Brain Res 685:205-210.

Luijten M, Gillan CM, de Wit S, Franken IHA, Robbins TW, Ersche KD (2020) Goal-Directed and Habitual Control in Smokers. Nicotine Tob Res 22:188-195.

Macey DJ, Schulteis G, Heinrichs SC, Koob GF (1996) Time-dependent quantifiable withdrawal from ethanol in the rat: effect of method of dependence induction. Alcohol 13:163-170.

Mangieri RA, Cofresi RU, Gonzales RA (2012) Ethanol seeking by Long Evans rats is not always a goal-directed behavior. PLoS One 7:e42886.

McCracken CB, Grace AA (2013) Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization. J Neurosci 33:17469-17482.

McKim TH, Shnitko TA, Robinson DL, Boettiger CA (2016) Translational Research on Habit and Alcohol. Curr Addict Rep 3:37-49.

Meinhardt MW, Giannone F, Hirth N, Bartsch D, Spampinato SM, Kelsch W, Spanagel R, Sommer WH, Hansson AC (2022) Disrupted circadian expression of beta-arrestin 2 affects reward-related micro-opioid receptor function in alcohol dependence. J Neurochem 160:454-468.

Meinhardt MW, Hansson AC, Perreau-Lenz S, Bauder-Wenz C, Stahlin O, Heilig M, Harper C, Drescher KU, Spanagel R, Sommer WH (2013) Rescue of infralimbic mGluR2 deficit restores control over drug-seeking behavior in alcohol dependence. J Neurosci 33:2794-2806.

Meinhardt MW et al. (2021) Psilocybin targets a common molecular mechanism for cognitive impairment and increased craving in alcoholism. Sci Adv 7:eabh2399.

Moykkynen T, Korpi ER, Lovinger DM (2003) Ethanol inhibits alpha-amino-3-hydyroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor function in central nervous system neurons by stabilizing desensitization. J Pharmacol Exp Ther 306:546-555.

Nebe S, Kroemer NB, Schad DJ, Bernhardt N, Sebold M, Muller DK, Scholl L, Kuitunen-Paul S, Heinz A, Rapp MA, Huys QJM, Smolka MN (2018) No association of goal-directed and habitual control with alcohol consumption in young adults. Addict Biol 23:379-393.

Nutt DJ, King LA, Phillips LD, Independent Scientific Committee on D (2010) Drug harms in the UK: a multicriteria decision analysis. Lancet 376:1558-1565.

Ostlund SB, Balleine BW (2008) On habits and addiction: An associative analysis of compulsive drug seeking. Drug Discov Today Dis Models 5:235-245.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Packard MG (1999) Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning. Proc Natl Acad Sci U S A 96:12881-12886.

Packard MG, McGaugh JL (1996) Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiol Learn Mem 65:65-72.

Pecina S, Berridge KC (2005) Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness? J Neurosci 25:11777-11786.

Pepino MY, Mennella JA (2005) Factors contributing to individual differences in sucrose preference. Chem Senses 30 Suppl 1:i319-320.

Pepino MY, Mennella JA (2007) Effects of cigarette smoking and family history of alcoholism on sweet taste perception and food cravings in women. Alcohol Clin Exp Res 31:1891-1899.

Quinn JJ, Hitchcott PK, Umeda EA, Arnold AP, Taylor JR (2007) Sex chromosome complement regulates habit formation. Nat Neurosci 10:1398-1400.

Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J (2009) Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 373:2223-2233.

Renteria R, Baltz ET, Gremel CM (2018) Chronic alcohol exposure disrupts top-down control over basal ganglia action selection to produce habits. Nat Commun 9:211.

Rimondini R, Arlinde C, Sommer W, Heilig M (2002) Long-lasting increase in voluntary ethanol consumption and transcriptional regulation in the rat brain after intermittent exposure to alcohol. FASEB J 16:27-35.

Robbins TW, Costa RM (2017) Habits. Curr Biol 27:R1200-R1206.

Root DH, Fabbricatore AT, Barker DJ, Ma S, Pawlak AP, West MO (2009) Evidence for habitual and goal-directed behavior following devaluation of cocaine: a multifaceted interpretation of relapse. PLoS One 4:e7170.

Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hagele C, Keller J, Junger E, Kathmann N, Smolka MN, Rapp MA, Schlagenhauf F, Heinz A, Huys QJ (2014) Model-based and model-free decisions in alcohol dependence. Neuropsychobiology 70:122-131.

Serlin H, Torregrossa MM (2015) Adolescent rats are resistant to forming ethanol seeking habits. Dev Cogn Neurosci 16:183-190.

Siciliano CA, Noamany H, Chang CJ, Brown AR, Chen X, Leible D, Lee JJ, Wang J, Vernon AN, Vander Weele CM, Kimchi EY, Heiman M, Tye KM (2019) A cortical-brainstem circuit predicts and governs compulsive alcohol drinking. Science 366:1008-1012.

Sjoerds Z, de Wit S, van den Brink W, Robbins TW, Beekman AT, Penninx BW, Veltman DJ (2013) Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. Transl Psychiatry 3:e337.

Smeets JAS, Minnaard AM, Ramakers GMJ, Adan RAH, Vanderschuren L, Lesscher HMB (2022) On the interrelation between alcohol addiction-like behaviors in rats. Psychopharmacology (Berl) 239:1115-1128.

Smith ACW, Jonkman S, Difeliceantonio AG, O'Connor RM, Ghoshal S, Romano MF, Everitt BJ, Kenny PJ (2021) Opposing roles for striatonigral and striatopallidal neurons in dorsolateral striatum in consolidating new instrumental actions. Nat Commun 12:5121.

Smith KS, Graybiel AM (2013) A dual operator view of habitual behavior reflecting cortical and striatal dynamics. Neuron 79:361-374.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Sommer WH, Costa RM, Hansson AC (2014) Dopamine systems adaptation during acquisition and consolidation of a skill. Front Integr Neurosci 8:87.

Spagnolo PA, Ramchandani VA, Schwandt ML, Zhang L, Blaine SK, Usala JM, Diamond KA, Phillips MJ, George DT, Momenan R, Heilig M (2014) Effects of naltrexone on neural and subjective response to alcohol in treatment-seeking alcohol-dependent patients. Alcohol Clin Exp Res 38:3024-3032.

Thorn CA, Atallah H, Howe M, Graybiel AM (2010) Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron 66:781-795.

Towner TT, Spear LP (2021) Rats exposed to intermittent ethanol during late adolescence exhibit enhanced habitual behavior following reward devaluation. Alcohol 91:11-20.

Tricomi E, Balleine BW, O'Doherty JP (2009) A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29:2225-2232.

Turner KM, Svegborn A, Langguth M, McKenzie C, Robbins TW (2022) Opposing Roles of the Dorsolateral and Dorsomedial Striatum in the Acquisition of Skilled Action Sequencing in Rats. J Neurosci 42:2039-2051.

Uhrig S, Vandael D, Marcantoni A, Dedic N, Bilbao A, Vogt MA, Hirth N, Broccoli L, Bernardi RE, Schonig K, Gass P, Bartsch D, Spanagel R, Deussing JM, Sommer WH, Carbone E, Hansson AC (2017) Differential Roles for L-Type Calcium Channel Subtypes in Alcohol Dependence. Neuropsychopharmacology 42:1058-1069.

Vandaele Y, Mahajan NR, Ottenheimer DJ, Richard JM, Mysore SP, Janak PH (2019) Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. Elife 8.

Vollstadt-Klein S, Wichert S, Rabinstein J, Buhler M, Klein O, Ende G, Hermann D, Mann K (2010) Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. Addiction 105:1741-1749.

Voon V, Derbyshire K, Ruck C, Irvine MA, Worbe Y, Enander J, Schreiber LR, Gillan C, Fineberg NA, Sahakian BJ, Robbins TW, Harrison NA, Wood J, Daw ND, Dayan P, Grant JE, Bullmore ET (2015) Disorders of compulsivity: a common bias towards learning habits. Mol Psychiatry 20:345-352.

WHO (2018) Global status report on alcohol and health 2018. In: World Health Organization, p 450. Geneva.

Yin HH, Knowlton BJ (2004) Contributions of striatal subregions to place and response learning. Learn Mem 11:459-463.

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464-476.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181-189.

Yin HH, Knowlton BJ, Balleine BW (2005a) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505-512.

Yin HH, Knowlton BJ, Balleine BW (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav Brain Res 166:189-196.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005b) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513-523.

bioRxiv preprint doi: https://doi.org/10.1101/2022.10.04.510642; this version posted October 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

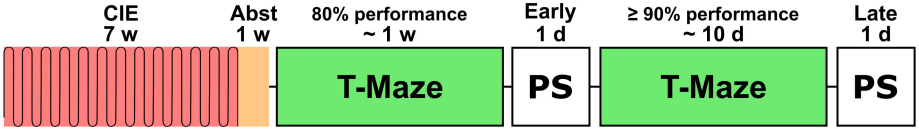
Yin HH, Mulcare SP, Hilario MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM (2009) Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat Neurosci 12:333-341.

Yu X, Chen S, Shan Q (2021) Depression in the Direct Pathway of the Dorsomedial Striatum Permits the Formation of Habitual Action. Cereb Cortex 31:3551-3564.

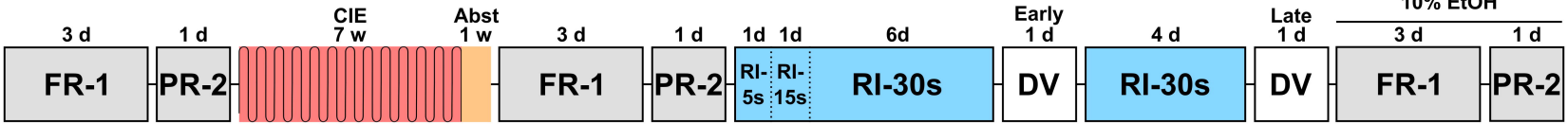
Zahr NM, Kaufman KL, Harper CG (2011) Clinical and pathological features of alcohol-related brain damage. Nat Rev Neurol 7:284-294.

Study 1: Habit formation in alcohol dependence

Experiment 1: Spatial navigation task

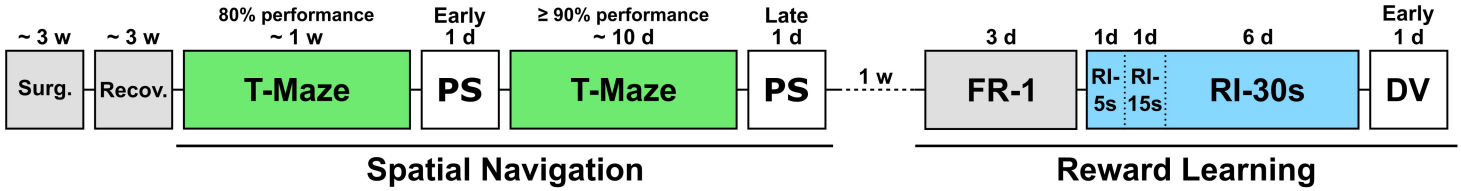


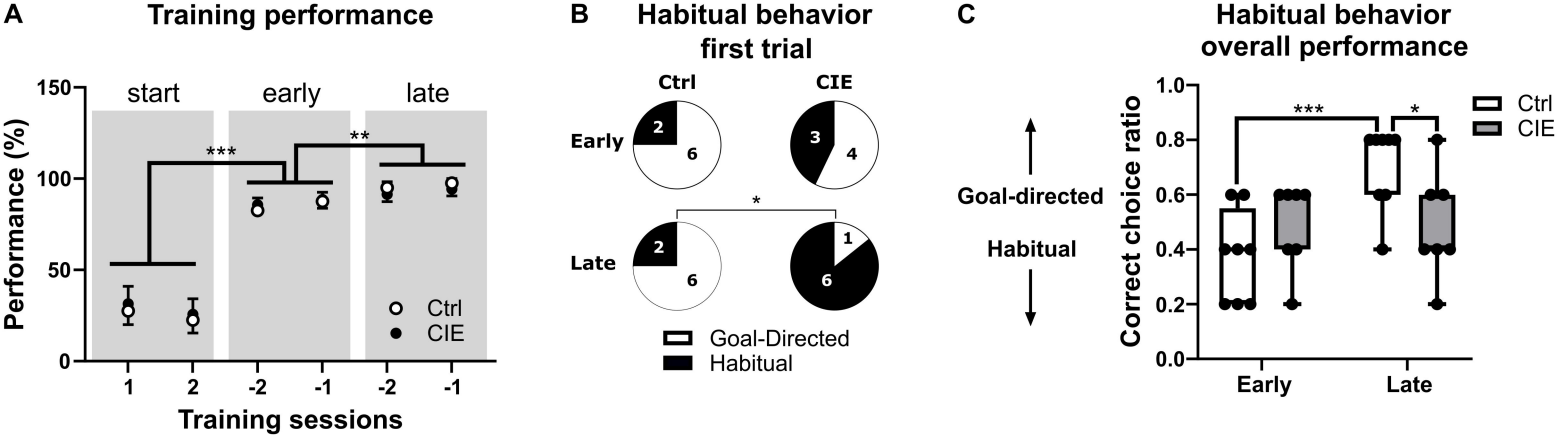
Experiment 2: Reward learning task

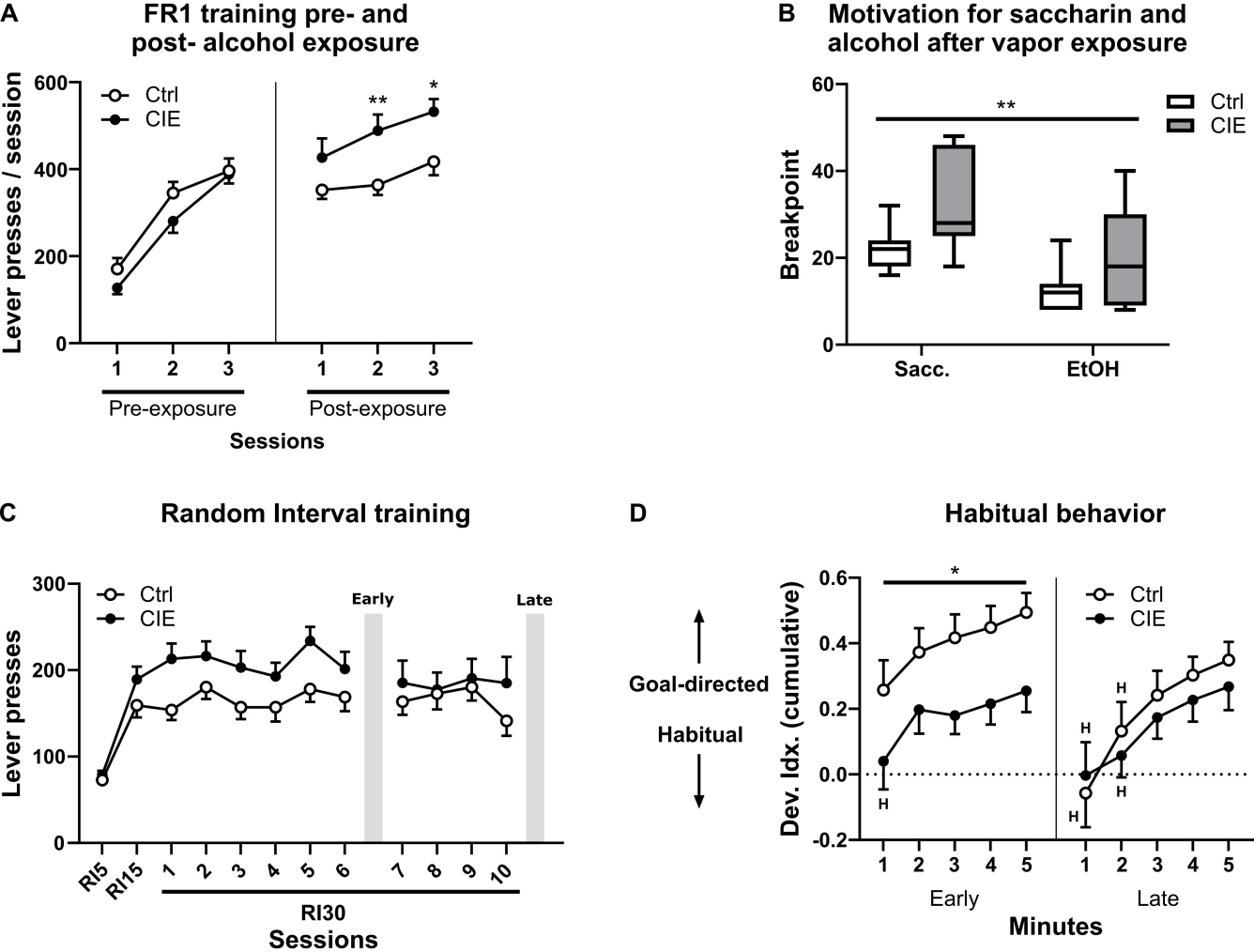


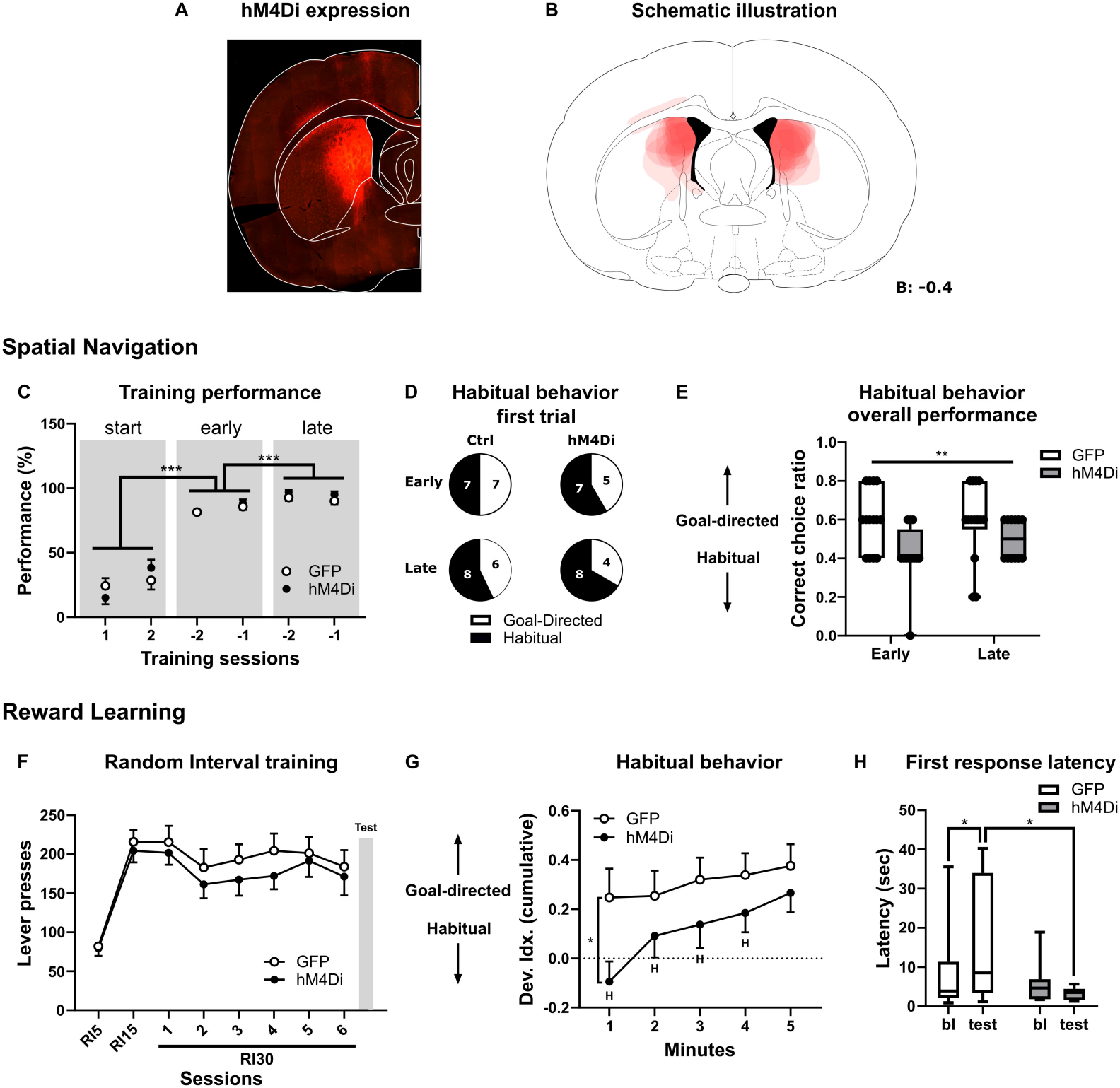
Study 2: pDMS inactivation in spatial navigation and reward learning

Experiment 3: Spatial navigation and reward learning tasks




A
Training performance
start
early
late
150
100
50
0
Performance (%)
***
**
Ctrl
CIE
1
2
-2
-1
-2
-1
Training sessions
B
Habitual behavior
first trial
Ctrl
CIE
Early
2
6
3
4
*
Late
2
6
1
6
Goal-Directed
Habitual
C
Habitual behavior
overall performance
Goal-directed
Habitual
Correct choice ratio
1.0
0.8
0.6
0.4
0.2
0.0
***
*
Ctrl
CIE
Early
Late


A
FR1 training pre- and post- alcohol exposure
Lever presses / session
600
400
200
0
Ctrl
CIE
**
*
1
2
3
1
2
3
Pre-exposure
Post-exposure
Sessions
B
Motivation for saccharin and alcohol after vapor exposure
Breakpoint
60
40
20
0
**
Ctrl
CIE
Sacc.
EtOH
C
Random Interval training
Lever presses
300
200
100
0
Ctrl
CIE
Early
Late
RI5
RI15
1
2
3
4
5
6
7
8
9
10
RI30
Sessions
D
Habitual behavior
Goal-directed
Habitual
Dev. Idx. (cumulative)
0.6
0.4
0.2
0.0
-0.2
*
Ctrl
CIE
H
H
H
H
H
1
2
3
4
5
1
2
3
4
5
Early
Late
Minutes


A
hM4Di expression
B
Schematic illustration
B: -0.4
Spatial Navigation
C
Training performance
start
early
late
***
***
Performance (%)
150
100
50
0
GFP
hM4Di
1
2
-2
-1
-2
-1
Training sessions
D
Habitual behavior
first trial
Ctrl
hM4Di
Early
7
7
7
5
Late
8
6
8
4
Goal-Directed
Habitual
E
Habitual behavior
overall performance
Goal-directed
Habitual
Correct choice ratio
1.0
0.8
0.6
0.4
0.2
0.0
**
GFP
hM4Di
Early
Late
Reward Learning
F
Random Interval training
Lever presses
250
200
150
100
50
0
Test
GFP
hM4Di
RI5
RI15
1
2
3
4
5
6
RI30
Sessions
G
Habitual behavior
Goal-directed
Habitual
Dev. Idx. (cumulative)
0.6
0.4
0.2
0.0
-0.2
GFP
hM4Di
*
H
H
H
H
1
2
3
4
5
Minutes
H
First response latency
GFP
hM4Di
Latency (sec)
50
40
30
20
10
0
*
*
bl
test
bl
test