Too little and too much: medial prefrontal functional inhibition impairs early acquisition of operant reversal learning, whereas medial prefrontal disinhibition impairs established serial-reversal performance in rats

Abbreviated title

Medial prefrontal GABA function and reversal learning

Author names and affiliations (incl. postal codes)

Renström[1], Jacco G.; Taylor[1], Charlotte J.L; Grasmeder Allen[1], Rachel; Loayza[1], Joanna; O'Hara[1], Luke; von Heimendahl[2], Moritz; Deiana[2], Serena; Moran[1], Paula M.; Stevenson[3], Carl W.; Maggi[1], Silvia; Bast[1], Tobias

1. School of Psychology, University of Nottingham, NG7 2RD, United Kingdom.
2. Boehringer Ingelheim Pharma GmbH & Co. KG, 88400 Biberach, Germany.
3. School of Biosciences, University of Nottingham, LE12 5RD, United Kingdom.

Author contributions:

JR, MvH, SD, PMM, CWS, SM, and TB designed research.

JR, CT, LOH, RGA, JL performed research.

JR and SM analyzed data.

JR and TB wrote the original draft of the paper, with all other authors subsequently contributing to revising the draft; all authors have approved the final content.

Funding:

The study and JR were supported by a BBSRC CASE Award in partnership with Boehringer Ingelheim GmbH & Co. KG (grant number BB/M008770/1, project 2275703). CT was supported by a BBSRC CASE award in partnership with b-neuro (grant number BB/M008770/1, project 2270910), RGA was supported by a Nottingham BBSRC Doctoral Training Partnership (DTP) studentship (grant number BB/T008369/1, project 2595069) and JL was supported by the MRC IMPACT DTP (grant number MR/N013913/1, project 2271003). SM was supported by an Anne McLaren Fellowship from the University of Nottingham.

Conflicting interests statement:

TB has obtained research funding from Boehringer Ingelheim, b-neuro and Neuro-Bio.

MvH and SD are currently employed by Boehringer Ingelheim GmbH & Co. KG. The other authors declare no competing financial interests.

Corresponding author email

Jacco Renstrom – jaccorenstrom@gmail.com

Tobias Bast – tobias.bast@nottingham.ac.uk

Acknowledgements

We thank the staff of the Biological Service Unit (BSU) for their contributions to the husbandry and welfare of the rats used in the present study.

ABSTRACT

Schizophrenia has been linked to hypofrontality (reduced prefrontal activation) and neural disinhibition (reduced GABAergic inhibition) within the dorsolateral prefrontal cortex (dlPFC), as well as reversal learning deficits. Interestingly, whilst reversal learning has been strongly linked to the orbitofrontal cortex, research suggests that the primate dlPFC – or its rodent analogue, the medial PFC (mPFC) – is less important for this process. Nevertheless, we hypothesized that the mPFC may be required for reversal learning if the reversal is particularly demanding. Furthermore, even if the mPFC is not required, mPFC disinhibition may impair reversals, because it may disrupt processing in mPFC projection sites. To test this, we induced mPFC functional inhibition and disinhibition in rats, using microinfusions of the GABA-A receptor agonist muscimol or antagonist picrotoxin, respectively, and examined the impact on early and serial reversals on a food-reinforced two-lever discrimination task. Using classical performance measures and Bayesian trial-by-trial strategy analysis, we found that mPFC functional inhibition impaired early, but not serial, reversals by increasing perseveration and impairing exploratory (lose-shift) behavior. In contrast, disinhibition impaired serial reversals, which was associated with reduction in both exploratory (lose-shift) and exploitative (win-stay) behavior. These findings suggest that mPFC hypoactivation and disinhibition disrupt distinct aspects of reversal learning by different mechanisms.

INTRODUCTION

Schizophrenia has been associated with two key neural abnormalities in the dorsolateral prefrontal cortex (dlPFC): reduced activation (hypofrontality) (Carter et al., 1998; Minzenberg et al., 2009) and reduced GABAergic inhibition (neural disinhibition) (Schoonover et al., 2020). These two aspects of prefrontal dysfunction were suggested to be causally linked, whereby hyperexcitability due to neural disinhibition at early stages of schizophrenia contributes to hypofrontality at later stages (Krystal & Anticevic, 2015). Moreover, research in animal models suggests that such imbalanced prefrontal activity, both too little and too much, can cause cognitive impairments relevant to schizophrenia and other neuropsychiatric disorders (Rao et al., 2000; Enomoto et al., 2011; Pezze et al., 2014; Tse et al., 2015; Bast et al., 2017).

Schizophrenia has also been associated with reversal learning deficits (Murray et al., 2008; Leeson et al., 2009; Izquierdo et al., 2017). Reversal learning is an important aspect of cognitive flexibility and involves the reversal of response-reward associations, whereby a previously rewarded response is no longer rewarded, whereas a previously unrewarded response becomes rewarded. Reversal learning has mainly been associated with the orbitofrontal cortex (OFC), whereas many studies suggest that the dlPFC in primates and the medial PFC (mPFC) in rodents, which shares functional-anatomical properties with the dlPFC (Brown & Bowman, 2002; Uylings, 2003; Dalley et al., 2004; Seamans et al., 2008; Laubach et al., 2018), are less important for reversal learning (Leeson et al., 2009; Izquierdo et al., 2017). Key evidence for this comes from lesion and pharmacological inactivation studies showing that OFC, but not dlPFC or mPFC, was required for reversal learning (Dias et al., 1996; Chudasama & Robbins, 2003; McAlonan & Brown, 2003; Boulougouris et al., 2007; Murray et al., 2007; Floresco et al., 2008; Hervig et al., 2020). Notwithstanding this evidence, imbalanced mPFC activity may still impair reversal learning. First, mPFC lesions impaired reversal learning when task difficulty was particularly high (such as when difficult discriminanda were used or more than two stimuli had to be monitored simultaneously), which was suggested to reflect a higher demand on mPFC-mediated attentional processes (Bussey et al., 1997; Chudasama & Robbins, 2003; Brigman & Rothblat, 2008; Kosaki & Watanabe, 2012). Second, the mPFC has been implicated in overcoming prepotent behavioral responses (Miller & Cohen, 2001; Haddon & Killcross, 2006; Marquis et al., 2007), which may be particularly relevant at early stages of reversal learning. Third, even under conditions when the mPFC is not required, mPFC disinhibition may disrupt reversal learning by causing aberrant drive of mPFC projections and, thereby, disrupting processing in mPFC projections sites (Bast et al., 2017), which include the OFC (Sesack et al., 1989).

Here, we examined in rats the effect of mPFC functional inhibition and disinhibition on early (Experiment 1) and serial (Experiment 2) reversal performance, using a food-reinforced two-lever reversal task (Brady & Floresco, 2015; Gonçalves et al., 2023). Functional inhibition and disinhibition were induced by microinfusion of the GABA-A receptor agonist muscimol or antagonist picrotoxin, respectively, which we confirmed previously, using in vivo electrophysiology, to induce neural changes consistent with inhibition/disinhibition, namely reduced/enhanced neuron firing, especially burst firing (Pezze et al., 2014). Alongside classical whole-session measures of reversal performance, including responses to criterion (RTCs) and errors, we used a Bayesian analysis to examine, with single-trial resolution, changes in behavioral strategies (including win-stay and lose-shift) that may underlie any changes in reversal performance (Maggi et al., 2024). Early and serial reversal performance may differently depend on the mPFC, because early performance may require more mPFC-dependent attention and cognitive control to overcome pre-existing response biases. Therefore, we predicted that mPFC functional inhibition may impair early acquisition of reversal learning but have less impact on serial reversal learning once reversal performance has been established. Additionally, we predicted that mPFC disinhibition may affect reversal performance, regardless of whether the mPFC is required, because disinhibition may disrupt processing in mPFC projection sites.

MATERIALS AND METHODS

Rats

In total, 64 young adult male Lister hooded rats (Envigo, UK) were used, weighing 290-340 g (8-9 weeks old) at surgery. Experiment 1 used 48 rats in two batches of 24 rats. Experiment 2 used a single batch of 16 rats. For sample size justification, exclusions and final sample sizes, see *Experimental design* section. Rats were housed in groups of four, in individually ventilated 'Double Decker' cages (462 mm X 403 mm X 404 mm; Tecniplast, UK) with temperature and humidity control (21±1.5 °C, 50±8%) and 12-h light-dark cycle (lights on at 0700). All experimental procedures were carried out during the light phase. Rats had unlimited access to water throughout the study. Access to food (Teklad Global 18% Protein Rodent Diet 2018C; Envigo, UK) was *ad libitum* until the start of food restriction one day prior to the start of operant pretraining. During food restriction, rats received daily food rations of 15-21 g per rat and were weighed daily to maintain body weights at 85-90% of free-feeding weights, as projected by a pre-established weight growth curve. On pretraining and test days, rats were weighed before the day's operant task session and received their daily food ration after

completing the session (in addition to the food reward received during the session). All procedures were conducted in accordance with the United Kingdom (UK) Animals (Scientific Procedures) Act 1986, approved by the University of Nottingham's Animal Welfare and Ethical Review Board and run under the authority of Home Office project license PP1257468. Experimental methods and results are reported in line with the ARRIVE guidelines (Lilley et al., 2020).

Guide cannula implantation into mPFC

Guide cannulas were implanted into the mPFC using the same coordinates and similar methods as in our previous studies (Pezze et al., 2014). Rats were anesthetized with isoflurane (induction: 3%; maintenance: 1-3%) delivered in medical oxygen (1 L/min). Following induction of anesthesia, rats' scalps were shaved, and all rats received subcutaneous (s.c.) injections of analgesic (Rimadyl 50 mg/ml; Zoetis, UK, diluted 1:9 with sterile saline, 0.9%, and injected at 0.1 ml/100 g) and antibiotic (0.02 ml/100 g Synulox containing 14% Amoxicillin; Zoetis, UK). Rats were transferred to the stereotaxic frame where they were secured in the horizontal-skull position with ear bars coated with local anesthetic cream (EMLA 5%, containing 2.5% lidocaine and 2.5% prilocaine; AstraZeneca, UK). Eye gel (Lubrithal, Dechra, UK) was applied to the eyes to prevent drying out during surgery, and the shaved scalp was disinfected with alcoholic skin wipes (2% chlorhexidine, 70% alcohol; Clinell, UK). Throughout the surgery, body temperature was maintained at 37 °C via a homeothermic heating pad controlled by an external temperature probe placed under the rat. A small anterior-posterior incision was made into the scalp to expose the skull, and bregma was located. Two small holes were drilled through the skull at the following coordinates: +3 mm anterior and ±0.6 mm lateral from bregma. Bilateral infusion guide cannulas ("mouse" model C235GS-5-1.2; Plastics One, Bilaney Consultants, UK) were used, consisting of a 5-mm plastic pedestal that held two 26-gauge metal tubes, 1.2 mm apart and projecting 4.5 mm from the pedestal. The guide cannulas were then lowered to 3.5 mm ventral from the skull surface and secured to the skull with dental acrylic (Kemdent, UK) and stainless-steel screws (1.2 mm X 3 mm; MDK Fasteners, UK). Non-protruding double stylets (33 gauge; Plastics One, Bilaney Consultants, UK) were inserted into the guide cannulas and a dust cap was secured on top. At the end of surgery, rats were injected with 1 ml of saline (s.c.) to minimize the risk of dehydration. Following surgery, rats continued to receive daily prophylactic antibiotic injections for the duration of the study. Rats were allowed to recover for at least 7 days before the start of food restriction and pretraining on the operant task.

mPFC drug microinfusions

Drug doses and infusion parameters were based on our previous studies, where both mPFC picrotoxin and muscimol infusions caused marked attentional deficits on the 5-choice serial reaction time task (Pezze et al., 2014). Rats were gently restrained to insert 33-gauge injectors (Plastics One, Bilaney Consultants, UK) into the previously implanted guide cannulas. The injectors protruded 0.5 mm below the guides into the mPFC, resulting in the following target coordinates for the infusions: 3 mm anterior and 0.6 mm lateral from bregma and 4 mm ventral from skull. The injector ends were connected via polyethylene tubing to two 5-µl syringes (SGE, World Precision Instruments, UK) secured on a micro-infusion pump (sp200IZ, World Precision Instruments, UK). Prior to infusions, the tubing and syringe were backfilled with distilled water and an air bubble was included before any drug solution was pulled up. A volume of 0.5 µl/side of either 0.9% sterile saline (vehicle), GABA-A-receptor antagonist picrotoxin (300 ng/0.5 µl/side, C30H34O13, Sigma-Aldrich, UK) in sterile saline, or agonist muscimol (62.5 ng/0.5 µl/side, C4H6N2O2, Sigma-Aldrich, UK) in sterile saline was administered over the span of 1 min. Movement of the air bubble within the tubing was monitored to ensure solutions had been successfully injected into the brain. Injectors remained in place for an additional 1 min to allow for absorption of the infusion bolus by the surrounding brain tissue. The injectors were then removed, and the stylets replaced. Testing began 10 min after the infusion was complete. In both Experiment 1 and 2, each rat received six microinfusions into the mPFC. The number of microinfusions per rat was limited to six, to reduce the risk that reactive gliosis caused by repeated microinfusions (Cunningham et al., 2008) may reduce effectiveness of the infused drug. After each infusion session, tubing and microinjectors were flushed with distilled water, and cleaned with >98% ethanol.

Operant reversal task

The pretraining and testing protocols were adapted from our previous studies (Gonçalves et al., 2023) and based on original protocols by Brady and Floresco (2015).

Apparatus

All testing was conducted in eight individual operant boxes (30.5 cm X 24.1 cm X 21.0 cm; Med-Associates, St. Albans, VT, USA), equipped with a house light (40 lux), an extractor fan, two retractable response levers either side of a dish, into which sucrose reward pellets (5TUL-45mg, Testdiet, OCB Systems Ltd, UK) could be dispensed, as well as two LED cue lights (40 lux each), one above each lever. The LED cue lights above the levers were illuminated pseudo-randomly throughout all sessions. In the present study, the cue lights were not

relevant to the correct lever choice (which was always determined by lever position, left or right), and they were essentially acting as distractors. Each rat was assigned to an operant chamber, where it underwent all operant pretraining and test sessions. Chambers were cleaned with 20% ethanol between different rats. The stimuli presented, lever operation and data collection were controlled via an interface with the computer and using custom software (MED-PC software, Med-Associates, St. Albans, VT, USA).

Pretraining

One day before pretraining, rats were habituated to the apparatus by placing them in the operant box for 15 min, with exterior doors open and no levers extended. Pretraining started with several days of single lever-press training sessions (one session per day), during which one of the two levers was extended for a fixed 30-min period, with each lever press rewarded by one sucrose pellet (5TUL-45mg, Testdiet, OCB Systems Ltd, UK). The initial location (left or right) of the lever was counterbalanced across rats. Lever location was switched once at least 50 responses were made in one session. At the beginning of the first lever-press session, two pellets were placed in the magazine and crushed pellets on top of the extended lever to promote rats' engagement with the apparatus. This was repeated on day two if rats did not readily press the lever on the previous day (<10 presses in the first 10 min). During these sessions, rats could press levers without restriction, allowing for unlimited rewards. This stage was completed after the criterion of 50 responses was met for both levers.

Lever-press training was followed by a minimum of 5 days of 90-trial retractable-lever training sessions (one session per day), where rats had to press the extended lever within the 10-s response window, after which the lever would be retracted. Both levers were presented in a pseudo-random order, but the same lever would not be extended more than two times in a row. Each lever press was rewarded by one sucrose pellet. At the end of this stage, rats were expected to make 5 or fewer omissions during a 90-trial session. 51/64 rats achieved this criterion by the last day. All rats had achieved several consecutive days of <5 omissions prior to the last day, with all rats making an average of 3.67±0.58 (mean±*SEM*) omissions by the final day, and thus all rats progressed to the next stage.

On the final day of pretraining, immediately after the last 90-trial session, the side preference of each rat was determined via 7 trials consisting of several sub-trials. At the start of each main trial, both levers were extended into the chamber and the initial press was rewarded, regardless of location. Subsequent sub-trials only rewarded the first response on the opposite lever. Once the rat had pressed the opposite lever, the next main trial began. Therefore, each

of the 7 trials included one rewarded press on each lever. The rat's side preference was defined as the side of the lever, which the rat chose to start on at least 4 of the 7 main trials.

Spatial discrimination and reversal testing

Testing began with a spatial discrimination (SD) task, where rats were rewarded to press the lever opposite to their side preference established on the previous day. Trials began every 20 s, with a 10-s response window. On each trial, both levers were presented, but only the one opposite to the rats' side bias was rewarded. Each correct response was rewarded with one sucrose pellet. Sessions were terminated once rats had reached a criterion of 10 consecutive correct responses. If the criterion was not reached within 200 trials, rats were retested with the same lever the next day. In the present study, all rats reached criterion within a maximum of two SD sessions.

The day after completion of the SD stage, reversal testing began. Each reversal stage started with 20 reminder trials, where the same lever response was rewarded as on the previous day. The purpose of reminder trials was to assess if retrieval/expression of the rule learnt on the previous day was affected by the prefrontal manipulations and, thereby, to aid interpretation of any effects of these manipulations on reversal learning (see Brady and Floresco, 2015). The reversal occurred after trial 20, when the correct and incorrect levers were reversed, such that the lever opposite to the one rewarded on the previous day and during the reminder trials was now the correct lever and rewarded (**Fig. 1A**). A reversal session ended when rats reached a criterion of 10 consecutive correct responses or after a maximum of 200 reversal trials. What happened when a rat failed to reach criterion in 200 trials and between reversal sessions differed between Experiment 1 and 2 and is described below (*Experimental design*).

Experimental design: Impact of mPFC infusions on early reversal learning (Experiment 1) and on well-established serial reversal learning (Experiment 2)

Timelines of Experiment 1 and 2 are shown in **Fig. 1B**, including when mPFC infusions were administered. In Experiment 1, we used a between-subjects design to examine the impact of saline, picrotoxin and muscimol infusion on rats' performance during the SD stage and during the first five reversals (R1 to R5). Rats were allocated to one infusion group and received the same infusion 10 min before the SD and each reversal session, based on onset of behavioral and electrophysiological effects of mPFC muscimol and picrotoxin infusion in our previous studies (Pezze et al., 2014). If rats did not reach the criterion of 10 consecutive correct trials within a maximum of 200 trials, they continued to be tested at the same stage (i.e., with the

same lever rewarded), on the following day, 10 min after receiving their prefrontal infusion. Because we limited the number of infusions per rat to six, rats that required more than one daily 200-trial session to complete a reversal stage, would complete less than five reversal stages. Therefore, statistical analysis was restricted to SD and R1-3, as these were the stages completed by all rats, with only 8/15rats in the picrotoxin group successfully completing all 5 reversals. Our target sample size was n=16 rats per group to give a power of 80% for pairwise comparisons (two-tailed independent t-tests, $p<0.05$), to detect differences between infusion groups that correspond to an effect size of Cohen's d of around 1 (G*Power, Faul et al., 2007). Experiment 1 was run using two cohorts of 24 rats (n=8 per group in each cohort). Rats were allocated to one of the three infusion groups - saline, picrotoxin or muscimol - via a randomized block design, with at least one rat in each cage of four being allocated to each infusion group. One rat died following surgery, one rat was culled due to complications with the cannula implant prior to testing and another two rats had to be excluded from the analysis due to incomplete testing. This resulted in the following final sample sizes: saline, n=15; picrotoxin, n=14; muscimol, n=15.

In Experiment 2, we used a within-subjects design to examine the impact of mPFC infusions on serial reversals, once reversal learning was well established. For this purpose, rats first underwent the SD stage and four reversals (R1 to R4), to establish relatively stable reversal performance, based on previous pilot work (chapter 2 in Renstrom, 2024), before the impact of mPFC infusions on serial-reversal testing was tested within-subjects across six additional reversals (R5 to R10). Rats were only allowed one reversal session, i.e. maximally 200 trials, to reach criterion. To reduce the risk of carry-over effects, whereby any impairment caused by an mPFC infusion during one reversal may affect performance on the next reversal and, thereby, confound the within-subjects design, we interspersed 'retention days' between reversals. On retention days, rats were tested without any infusion and with the same lever rewarded as on the preceding infusion day (i.e., during the last reversal) until they reached the criterion of 10 consecutive correct responses or for 200 trials (all rats reached criterion within less than 200 trials on retention days). Retention days reinforced the lever press rule from the last reversal before testing rats on the next reversal, to ensure that all rats started each new reversal with the old rule from the previous stage well established. Each rat received two series of the three different mPFC infusions (saline, picrotoxin, and muscimol), with infusion series 1 consisting of infusions applied 10 min before R5-7 and infusion series 2 consisting of infusions applied 10 min before R8-10. Testing order of the three infusion conditions was counterbalanced within each series using a Latin square design. Individual rats were randomly allocated to one of the three testing orders determined according to the Latin square design, ensuring that at least one rat in each cage of four rats was allocated to each

of the three testing orders. Our target sample size was n=12-16 rats to give a power of >80% for pairwise comparisons (two-tailed paired t-tests, $p<0.05$) to detect differences between infusion conditions that correspond to an effect size of Cohen's d of around 1 (G*Power, Faul et al., 2007). Three rats were culled due to complications with the cannula implant, resulting in a final sample size of n=13.

In both experiments, the main experimenter (JR) was blinded with respect to the infusion groups and conditions. Blinding was only broken after data analysis (although the identity of the picrotoxin group/condition was evident from the increased omissions shown following mPFC picrotoxin infusions).

Verification of cannula placements

After the completion of the experiments, rats were overdosed with sodium pentobarbitone (1-2 ml Euthatal; sodium pentobarbitone, 200 mg/ml; Genus Express, UK) and transcardially perfused with 0.9% saline followed by 4% paraformaldehyde solution in saline. Following extraction, brains were post-fixed in 4% paraformaldehyde and cut into 70-µm coronal sections using a vibratome (Leica, UK). Sections were then mounted on slides and analyzed under a light microscope, where cannula placements were verified and mapped onto coronal sections of a rat brain atlas (Paxinos and Watson, 1998).

Data analysis

Performance measures

The main measure of operant task performance was RTC, i.e., the number of total trials a rat required to achieve the success criterion of 10 consecutive correct responses, excluding omitted trials when the rat did not respond within the 10-s response window. For the reminder trials, which continued even if rats made 10 consecutive correct responses, 'percentage correct' was used, which refers to the percentage of correct reminder-trial responses divided by the total number of reminder-trial responses, excluding omissions. For the analysis of errors during reversal sessions, previous studies used different approaches to differentiate between 'perseverative errors', thought to reflect difficulties with giving up responding according to the old rule, and 'regressive errors', thought to reflect chance or regression in learning the new rule (e.g., Deuel et al., 1971; Floresco et al., 2006; Brady & Floresco, 2015; Dalton et al., 2016; Dhawan et al., 2019). We compared several different approaches, and all yielded similar conclusions. Here, we only report the outcome of the approach outlined in Brady and Floresco (2015). We counted errors as perseverative errors until a threshold of less than 10 total errors in a block of 16 responses was reached. After this, all errors were counted as regressive.

Additionally, we analyzed omissions (i.e., trials when rats did not press a lever within the 10-s response window), as well as response latencies for correct and incorrect responses. Performance measures during SD and reversal trials were analyzed separately from the performance measures during the reminder trials at the beginning of each reversal stage.

Trial-by-trial Bayesian strategy analysis

Complementing the classical performance measures outlined above, we also used a trial-by-trial Bayesian strategy analysis (Maggi et al., 2024). This analysis estimates the probability of a rat applying a particular response strategy on any given trial based on evidence collected up to this trial. Each strategy (e.g., lose-shift, win-stay) was estimated independently from each other and required a different set of input parameters. These parameters included the choice (i.e. right or left lever), and the outcomes of the current or previous trials. For example, a 'lose-shift' strategy is defined as an incorrect/unrewarded response ('lose'), followed by a response 'shift' to the opposite lever on the following trial, whereas a 'win-stay' strategy is defined as a correct/rewarded response ('win'), followed by re-selection of the same lever ('stay') on the following trial. Estimating the lose-shift/win-stay strategy probabilities, therefore, requires information about the choice and outcome on trial t and the choice on trial t+1. Briefly, the Bayesian strategy analysis uses the Bayes rule to estimate the probability of a given strategy *i* being applied by the subject based on the history of choices up to the current trial *t,* $P(strategy_i(t)|choices(1:t))$. This *posterior* probability is updated based on the *likelihood* (how well the choices match the strategy) and the *prior* (initial probability estimate of using that strategy). The initial estimation of the prior at trial 0 was based on the uniform distribution. Maggi et al. (2024) showed that the choice of prior (whether based on the uniform or Jeffrey's distribution) has minimal impact on the estimated strategy probabilities after just a few trials because the actual data quickly dominates over the initial prior assumptions. In Experiment 1, we used the SD stage to accumulate evidence for each strategy before comparing the impact of mPFC manipulations on probabilities from R1 onwards. In Experiment 2, evidence was accumulated across the SD stage and R1-R4 before comparing the impact of mPFC manipulations on probabilities at the first infusion session. Following Maggi et al. (2024), we also used a forgetting rate that progressively reduces the influence of past evidence while giving more weight to recent choices in the probability calculations. If no evidence is available for or against a given strategy on a particular trial, the strategy's probability remained the same as on the previous trial.

We examined the following strategies:

1) Strategies that reflect if rats follow the task rule - ‘Go-previous’ strategy around reversal, and its complementary strategy ‘go-new’ around learning:

The go-previous strategy involves the rat following the spatial rule that was correct prior to rule change (i.e., go-right or go-left), whereas the go-new strategy involves the rat applying the correct choice strategy after the reversal. Go-previous and go-new strategies are mutually exclusive, i.e. their probabilities add up to P=1, whereby as one increases the other one decreases, and vice versa. We focused on go-previous around reversal and go-new around learning, because these provide information about how flexibly rats moved away from the previously learned rule and how quickly they learned the new rule, respectively.

2) Other explorative and exploitative strategies:

Win-stay and lose-shift strategies are both task-appropriate strategies that would support correct responding, because the rat continues to press the correct lever or ‘shifts’ to the other lever after an incorrect choice, respectively. In addition, we also examined strategies that were task-inappropriate, such as cue-light-based strategies, and situational strategies that are neither directly appropriate nor inappropriate, such as a ‘sticky’ strategy, which involved ‘sticking’ with the same response regardless of outcome. We only report these probabilities if they varied substantially from chance (P=0.5), which was only the case for sticky around new learning in Experiment 2. Finally, for all the strategies above, there is a strategy that is mutually exclusive (go-previous and go-new; win-stay and win-shift; lose-shift and lose-stay; sticky and alternating). Because the probability of one of these two mutually exclusive strategies is implied by the probability of the other, we only analyzed the probabilities of the dominant strategy (i.e., P>0.5) out of these pairs of mutually exclusive strategies.

We investigated strategies around reversal (i.e., responses immediately preceding and following the rule reversal) and preceding learning (i.e., responses preceding the 10 consecutive correct responses that we defined as criterion of successful learning and that ended the test sessions). These time points were of interest, as response strategies around rule reversal reveal how the rats respond to the rule change, whereas strategies preceding learning reveal how rats approach the acquisition of a new rule. We restricted our Bayesian strategy analysis to the number of responses that was completed by all rats across all sessions. Therefore, for both experiments, strategy analysis around reversal was restricted to the 6 reminder-trial responses preceding reversal and the 16 responses following reversal (i.e., 22 total responses around reversal). For the analysis of strategies preceding learning, we also excluded the last response that rats made before starting the 10 consecutive correct

responses that ended the session, as this was inevitably incorrect for each rat, and thus would violate the assumption of independence of observations for subsequent ANOVA. In Experiment 1, analysis of strategies preceding learning was restricted to 6 responses, whereas, in Experiment 2, it was restricted to 5 responses preceding the last incorrect response.

Statistical analysis

All statistical analyses were carried out in JASP (version 0.18.3), and all graphs were made in GraphPad prism (version 10.2.1). Data were analyzed by ANOVA. In Experiment 1, infusion group (saline, picrotoxin, muscimol) was used as a between-subjects factor and task stage (SD, R1-3) as a within-subjects factor. In Experiment 2, infusion condition (saline, picrotoxin, muscimol) and infusion series (1 or 2) were used as within-subject factors. ANOVAs of trial-by-trial strategy probabilities additionally included response number as a within-subjects factor (we refer to 'responses' due to the removal of omission trials in these analyses). Where the assumption of sphericity for within-subjects ANOVA was violated, Greenhouse-Geisser correction was applied to the degrees of freedom. Interactions were further examined by simple-main effects analysis, and simple main effects were further analyzed by pairwise comparisons, using Fisher's LSD test. The accepted level of significance was $p<0.05$ for all analyses.

RESULTS

Cannula placements within mPFC

In both experiments, all infusion cannula tips were placed within the mPFC, in an area that corresponded approximately to +2.7 to +4.2 mm anterior to bregma in the atlas by Paxinos and Watson (1998) (**Fig. 2**).

Experiment 1: mPFC functional inhibition impaired early reversal learning performance, whereas mPFC disinhibition facilitated early reversal learning

mPFC functional inhibition increased, whereas disinhibition reduced RTC during R2

First, we examined session-wide reversal performance indexed by RTC across SD and the following first three reversal stages, R1 to R3 (**Fig. 3A**). The saline group displayed an initial reversal cost, with higher RTC at R1 compared to SD (137.47±11.97 and 93.6±6.54,

mean±*SEM*, respectively), which decreased across reversals as learning progressed. A significant drug X stage interaction was observed ($F_{(6,123)}$=3.340, p=0.004). At R2, mPFC functional inhibition by muscimol increased RTC, whereas mPFC disinhibition by picrotoxin reduced the reversal cost ($F_{(2, 41)}$=10.892, p<0.001). RTC in the muscimol group were higher compared to picrotoxin (p<0.001) and tended to be higher compared to saline (p=0.051). Additionally, RTC in the picrotoxin were lower than in the saline group (p=0.010). Interestingly, muscimol improved SD learning ($F_{(2,41)}$=3.655, p=0.035), reducing RTC compared to saline (p=0.010; all other pairwise comparisons, smallest p=0.170). No significant differences were observed at R1 ($F_{(2,41)}$=2.540, p=0.091) or R3 ($F_{(2,41)}$=2.060, p=0.140).

mPFC functional inhibition increased, whereas disinhibition reduced perseverative errors during R2

Next, to examine the group differences at R1 and R3 further, we analysed perseverative and regressive errors (**Fig. 3B** and **C**). For perseverative errors, we observed a significant drug X stage interaction ($F_{(4,82)}$=4.852, p=0.001). There was a main effect of drug at R2 ($F_{(2,41)}$=11.625, p<0.001), where mPFC muscimol increased perseverative errors compared to both saline (p=0.009) and picrotoxin (p<0.001), and mPFC picrotoxin reduced perseverative errors compared to saline (p=0.040). In addition, a trend towards a main effect of drug was observed at R3 ($F_{(2,41)}$=3.134, p=0.054), but not R1 (F<1). Regressive errors were similar across drug groups ($F_{(2,82)}$=2.043, p=0.143) and stages (main effect of stage and drug X stage interaction, F<1), suggesting that mPFC manipulations specifically affected perseveration without causing non-specific learning impairments during early reversals.

mPFC disinhibition increased omissions, limiting reminder-trial responses

mPFC picrotoxin markedly increased the number of omissions during reminder-trial responses ($F_{(2, 41)}$=173.014, p<0.001), across all task stages (there was only a trend towards a drug X stage interaction, $F_{(3.432, 70.365)}$=2.416, p=0.066), compared to saline (p<0.001) and muscimol (p<0.001), whereas saline and muscimol did not differ (p=0.333) (**Table 1**). Omissions across the initial SD learning trials and the reversal trials at R1-3 were also markedly increased by picrotoxin. There was a drug X stage interaction ($F_{(3.895, 79.856)}$=4.291, p=0.004), with a main effect of drug infusion at every stage (smallest $F_{(2, 41)}$=19.436, p<0.001). *Post-hoc* pairwise comparisons highlighted that, at all stages, mPFC picrotoxin increased omissions compared to both saline and muscimol (both p<0.001), which did not differ from each other at any stage (smallest p=0.333). This increase in omissions by mPFC neural disinhibition is consistent with previous findings (Pezze et al., 2014; Auger et al., 2017).

mPFC functional inhibition did not affect correct reminder-trial responses

The marked increase in reminder trial omissions in the picrotoxin group limited the number of reminder-trial responses (average reminder-trial responses across R1-R3 out of a maximum of 20, mean±*SEM*: 3.86±0.86) compared to the saline (19.78±0.09) and muscimol (19.66±0.11) group. This low number of reminder trials rendered the analysis of percentage correct measures for reminder trials unreliable in the picrotoxin group. Therefore, the picrotoxin group was excluded from the analysis of the percentage correct values during reminder trials. Analysis of percentage correct measures limited to the muscimol and saline groups, revealed that prefrontal muscimol did not affect correct responses during reminder trials compared to saline (no main effect of drug or stage, and no drug X stage interaction, largest $F_{(2, 52)}$=2.149, p=0.127) (**Fig. 3D**).

mPFC disinhibition increased response latencies

Next, we examined latencies of correct and incorrect responses across reminder trials and reversal trials (**Table 1**). Due to the high number of omissions, only one rat receiving mPFC picrotoxin made at least one correct and one incorrect response during the 20 reminder trials at all three reversal stages (compared to all 15 rats in both the saline and muscimol groups). Therefore, we restricted the analysis of reminder-trial latencies across the three reversal stages to the saline and muscimol groups, which revealed no main effect or interaction involving drug (largest $F_{(1, 28)}$=1.773, smallest p=0.194). All three groups could be included in the analysis of latencies during the SD trials and the reversal trials of the three reversal stages (R1-3). Consistent with previous studies showing increased response latencies following mPFC disinhibition (Enomoto et al., 2011; Pezze et al., 2014; Auger & Floresco, 2015; Auger et al., 2017), this analysis revealed increased response latencies following mPFC picrotoxin regardless of response type and at all task stages (significant drug X stage X response type interaction, $F_{(6, 117)}$=2.367, p=0.034). We further analyzed this triple interaction, using a 2-way ANOVA at each stage (SD, R1-3), with drug and response type as factors. This showed a main effect of drug at each stage (smallest $F_{(2, 39)}$=3.954, p=0.027), as well as a main effect of response type at SD, R1 and R3 (smallest $F_{(1, 39)}$=7.652, p=0.009). Furthermore, there was a significant drug X response type interaction at R1 ($F_{(2, 39)}$=6.364, p=0.004), but not the other stages (largest $F_{(2, 39)}$=1.191, p=0.315). *Post-hoc* pairwise comparisons at SD, R2 and R3 revealed that, at all these stages, picrotoxin increased response latencies compared to saline and muscimol infusions (largest p=0.008), which did not differ (smallest p=0.157). Simple main effect analysis of the interaction at R1 revealed significant simple main effects of drug on both

correct and incorrect latencies (both $F_{(2, 39)}$>12.8, p<0.001). *Post-hoc* pairwise comparisons revealed that, also at R1, picrotoxin increased both correct and incorrect response latencies compared to saline and muscimol (largest p=0.004), which did not differ from each other (smallest p=0.333).

Bayesian analysis of strategies around reversal: mPFC functional inhibition enhanced use of the old strategy (go-previous) and disrupted exploration (lose-shift) at R2

To complement classical whole-session measures of reversal performance, and to gain an insight into the underlying behavioral changes across the trials within a session, we first examined rats' response strategies during responses around rule reversal (6 reminder-trial responses and 16 reversal-trial responses), using the Bayesian trial-by-trial strategy analysis (Maggi et al., 2024). The picrotoxin group was excluded from this analysis, because of the low number of reminder trials completed by this group due to the high number of omissions (see above) (**Fig. 4**). Analysis of the go-previous probability in the saline and muscimol groups (**Fig. 4A**) revealed a significant drug X stage X response interaction ($F_{(42, 1176)}$=2.235, p<0.001). Further analysis of this triple interaction revealed a significant drug X response interaction at R2 ($F_{(3.340, 93.078)}$=6.456, p<0.001), whereas, at R1 and R3, there was no main effect or interaction involving drug (largest $F_{(2.461, 68.910)}$=1.350, p=0.267). The drug X response interaction at R2 reflected that mPFC functional inhibition increased the go-previous probability compared to saline between responses 9-15 after reversal (smallest $F_{(1, 28)}$=5.165, p=0.031), with trends toward this difference at responses 8 ($F_{(1, 28)}$=3.980, p=0.056) and 16 ($F_{(1, 28)}$=4.090, p=0.053) after reversal, in line with the increased RTC and perseverative errors caused by mPFC muscimol at that stage (**Fig. 3A and B**).

Before rule reversal, during reminder trials, rats showed high lose-shift (**Fig. 4B**) and high win-stay (**Fig. 4C**) probabilities. The high lose-shift probabilities may partly reflect that rats made very few incorrect responses during reminder trials, and, therefore, the lose-shift probability was hardly updated by our Bayesian analysis, and remained at similarly high values as at the end of the preceding reversal session. The high win-stay probabilities indicated strong exploitation, consistent with the high percentage of correct responses during reminder trials. However, immediately after reversal, there was a steep decline in lose-shift behavior and a more gradual decline in win-stay behavior, similar to our previous findings following the switch from a spatial rule to a cue-based rule (Maggi et al., 2024), revealing that the rule reversal induced marked changes in exploitation and exploration. Analysis of lose-shift probabilities revealed a drug X stage X response interaction ($F_{(42, 1176)}$=1.436, p=0.037). This triple

interaction reflected that lose-shift behavior was reduced by muscimol, compared to saline, specifically at R2, between responses 8 to 16 after reversal (drug X response interaction, $F_{(1.803,\ 50.471)}$=3.569, p=0.040; main effect of drug at responses 8-16, smallest $F_{(1,\ 28)}$=6.795, p=0.014; main effect of drug for all other responses, largest $F_{(1,\ 28)}$=2.797, p=0.106). In contrast, at R1 and R3, there was no main effect or interaction involving drug (largest $F_{(1,\ 28)}$=1.250, p=0.273), only a main effect of response (both $F_{(2.094,\ 54.390)}$>40.771, p<0.001). In contrast to lose-shift behaviour, win-stay behaviour was not affected by drug, but declined similarly following both mPFC saline and muscimol and across all stages following reversal (only a main effect of response, $F_{(2.863,\ 80.174)}$=81.264, p<0.001, but no other main effects or interactions, largest $F_{(42,\ 1176)}$=1.289, p=0.104).

Bayesian analysis of strategies preceding learning: no effect of mPFC functional inhibition.
Contrasting with the group differences in strategy use around reversal, mPFC muscimol did not affect strategies leading up to the point of learning (data not shown). Analysis of go-new, lose-shift and win-stay probabilities during responses -5 to -1 leading up to learning only revealed an effect of response for each strategy (smallest $F_{(1.223,\ 34.246)}$=5.288, p=0.022), reflecting that these task-pertinent strategies increased leading up to reaching the performance criterion, without any other main effect or interaction involving drug (largest $F_{(1.387,\ 38.830)}$=1.667, p=0.203).

Experiment 2: **mPFC disinhibition, but not functional inhibition, impaired serial reversal learning**

mPFC disinhibition increased RTC during serial reversal learning
mPFC picrotoxin markedly impaired serial reversal performance, reflected by higher RTC (main effect of drug, $F_{(2,\ 24)}$=9.603, p<0.001),compared to saline (p=0.003) and muscimol (p=0.003), while saline and muscimol did not differ from each other (p=0.266) (**Fig. 5A**). There was a trend towards a main effect of infusion series ($F_{(1,\ 12)}$=4.255, p=0.058), which reflected that RTC tended to decline slightly from series 1 to 2 across all infusion conditions (data not shown), but no drug X infusion series interaction (F<1).

mPFC disinhibition increased both perseverative and regressive errors
mPFC picrotoxin increased perseverative errors (main effect of drug, $F_{(1.269,\ 15.231)}$=4.592, p=0.041) compared to saline (p=0.021) and muscimol (p=0.011) (**Fig. 5B**) and regressive

errors (main effect of drug, $F_{(2, 24)}$=5.427, p=0.011) compared to muscimol (p=0.003), with a trend toward an increase compared to saline (p=0.050) (**Fig. 5C**).There was no main effect of infusion series or drug X infusion series interaction (largest $F_{(1.199, 14.383)}$=1.199, p=0.140).

mPFC disinhibition impaired expression of the previous rule during reminder trials

mPFC disinhibition also markedly impaired the expression of the old rule prior to rule change, as reflected by a reduced percentage of correct responses during the reminder trials (main effect of drug, $F_{(2, 24)}$=24.673, p<0.001) (**Fig. 5D**). The percentage of correct reminder trials was reduced by picrotoxin compared to saline and muscimol infusions (both p<0.001), which did not differ from each other (p=0.333). No main effect or interaction involving infusion series was observed (main effect of infusion series, F<1; drug X infusion series interaction, $F_{(1.1912, 22.939)}$=2.778, p=0.085).

mPFC disinhibition increased omissions and response latencies

As in Experiment 1, and consistent with previous studies (see above), mPFC disinhibition markedly increased omissions and latencies of correct and incorrect responses across both reminder and reversal trials, whereas functional inhibition did not affect these measures (**Table 1**). However, compared to Experiment 1, omissions were overall lower, and the increase in omissions by prefrontal picrotoxin was not as detrimental to reminder trial responding. Nevertheless, a main effect of drug was still observed for reminder-trial omissions ($F_{(1.003, 12.04)}$=5.510, p=0.037), and reversal-trial omissions ($F_{(1.002, 12.02)}$=24.204, p<0.001), which, in both cases, was driven by increased omissions following mPFC picrotoxin compared to saline and muscimol (largest p=0.008), whereas saline and muscimol did not differ (both p=0.333). No main effect or interaction involving series was observed for omissions during reminder or reversal trials (largest $F_{(1, 12)}$=2.695, p=0.127). For reminder-trial response latencies, there was a significant drug X response-type interaction ($F_{(2, 22)}$=4.317, p=0.026), and no other main effect or interaction involving infusion series or response type (all F<1). Simple-main effects analysis showed a significant effect of drug for both correct- and incorrect-response latencies (smallest $F_{(2, 24)}$=7.279, p=0.004). Pairwise comparisons revealed that picrotoxin increased both correct and incorrect reminder-trial response latencies compared to saline and muscimol (all p<0.001), which did not differ (p=0.333). For reversal-trial response latencies, there was also a significant drug X response-type interaction ($F_{(2, 24)}$=3.902, p=0.034), alongside a significant drug X infusion series interaction ($F_{(1.078, 12.940)}$=4.747, p=0.046). Subsequent simple main effects analyses revealed significant main effects of drug across infusion series and response types (smallest $F_{(2, 24)}$=28.624, p<0.001). Pairwise comparisons revealed that this

effect was driven by increased response latencies following picrotoxin, compared to saline and muscimol for both series and response types (largest $p<0.001$); muscimol and saline did not differ regardless of infusion series or response type (all $p=0.333$).

Bayesian analysis of strategies around reversal: mPFC disinhibition changed adherence to old strategy (go-previous) before and after reversal and disrupted both exploration (lose-shift) and exploitation (win-stay)

During serial reversals, rats receiving mPFC saline or muscimol infusions showed high go-previous probabilities during the 6 last responses before the rule change; go-previous probabilities remained high until about the 7th response after rule change when they began to decrease (**Fig. 6A**), indicating that rats abandoned the previous rule and switched to the opposite lever. This decline in go-previous behavior was evident regardless of infusion series (i.e., during R5 to 7 and R8 to 10) and differed from the pattern during early reversals (R1 to R3) in Experiment 1, where the go-previous probability did not decline across the first 16 responses following rule reversal (**Fig. 4A**). Following mPFC disinhibition, however, the pattern differed: the go-previous probability tended to be reduced compared to saline and muscimol infusions during the reminder trials, at or below P=0.7 (in line with reduced reminder trial accuracy, **Fig. 5D**), but hardly declined following rule reversal, such that, towards the end of the first 16 responses after reversal, the go-previous probability was similar in all three infusion conditions (**Fig. 6A**). In support of these findings, ANOVA of the go-previous probabilities revealed a drug X response interaction ($F_{(42, 504)}=1.651$, $p=0.008$) and a trend towards a drug X infusion series interaction ($F_{(2, 24)}=2.965$, $p=0.071$), but no drug X infusion series X response interaction ($F_{(42, 504)}=1.213$, $p=0.174$). Simple main effects analysis of the drug X response interaction did not indicate a significant effect of drug for any response but indicated trends toward such an effect for responses -1, +2, and +8 around reversal (smallest $F_{(2, 24)}=2.586$, $p=0.096$); for all other responses, largest $F_{(2, 24)}=2.542$, $p=0.100$).

The time course of lose-shift behavior during repeated reversals (**Fig. 6B**) was similar to that during early reversals in Experiment 1 (**Fig. 4B**). Before reversal, lose-shift probability was high across all three infusion conditions, and, after reversal, there was a steep decline in this strategy, indicating a tendency to stick with the previous, but now incorrect, response for several trials. Following mPFC saline and muscimol infusions, this decline plateaued around 5 responses after reversal, coinciding with a decline in go-previous probability (**Fig. 6A**), which reflected that rats began to switch to the opposite lever. However, following mPFC picrotoxin, lose-shift probability continued to decline for all 16 responses that were analysed following reversal. These findings were supported by ANOVA of lose-shift probabilities, which revealed

a significant drug X response interaction ($F_{(42, 504)}$=2.250, p<0.001), with no further main effects or interactions (largest $F_{(42, 504)}$=1.096, p=0.319). Simple main effect analysis revealed a significant effect of drug at responses 15 and 16 (smallest $F_{(2, 24)}$=4.185, p=0.028) and a trend at responses 13 and 14 post-reversal (smallest $F_{(2, 24)}$=3.185, p=0.059) (all other responses: largest $F_{(2, 24)}$=2.285, p=0.123). *Post-hoc* pairwise comparisons showed that picrotoxin reduced lose-shift probabilities at both responses 15 and 16 compared to saline (largest p=0.043) and muscimol (largest p=0.029), which did not differ significantly from each other (p=0.600).

The probability of win-stay behaviour during repeated reversals remained high around reversal, contrasting with the steep decline following reversal during early reversals in Experiment 1 (**Fig. 4C**), but was reduced by mPFC picrotoxin compared to saline and muscimol during some of the responses around reversal (**Fig. 6C**). ANOVA of win-stay probabilities revealed a significant drug X response interaction ($F_{(42, 504)}$=2.904, p<0.001), with no further main effects or interactions (largest $F_{(21, 252)}$=1.196, p=0.254). Simple main effect analyses revealed a significant drug effect starting four responses before, and ending four responses after reversal, as well as at the 16th response after reversal (smallest $F_{(2, 24)}$=3.778, p=0.037). Additionally, there was a trend toward an effect of drug at the fifth response before and after reversal, as well as at the 15th response after reversal (smallest $F_{(2, 24)}$=2.591, p=0.096, but not drug effect at any other response (largest $F_{(2, 24)}$=2.311 p=0.121). *Post-hoc* pairwise comparisons, to analyze further the simple main effects of drug at responses -4 to +4 around reversal, showed that prefrontal picrotoxin significantly reduced win-stay probability compared to saline and muscimol (largest p=0.049), which did not differ (smallest p=0.260). On the other hand, the simple main effect of drug at response 16 after reversal was driven by significantly increased win-stay probability in the muscimol condition, compared to saline (p=0.019) and picrotoxin (p=0.036) which did not differ between each other at that stage (p=0.333).

Bayesian analysis of strategies preceding learning: mPFC disinhibition increased adherence to new rule (go-new) and functional inhibition increased win-stay and sticky behavior leading up to learning

Across the 5 responses leading up to the point of learning, the go-new probabilities were increased by mPFC picrotoxin compared to saline and muscimol infusions and gradually increased regardless of the infusion condition (**Fig. 7A**). The increased go-new probability, alongside increased RTC (**Fig. 5A**), suggests that, following mPFC picrotoxin, rats achieved

the learning 'breakthrough' (i.e., 10 consecutive correct responses) more gradually than following mPFC muscimol and saline infusion. ANOVA of go-new probabilities revealed a main effect of drug ($F_{(2, 24)}$=4.650, p=0.020) and of response ($F_{(1.526, 18.318)}$=16.264, p<0.001), reflecting a gradual increase, alongside a trend towards an effect of series ($F_{(1, 12)}$=3.718, p=0.078), but no interactions (largest $F_{(2.112, 25.343)}$=1.165, p=0.330). *Post-hoc* pairwise comparisons showed that, following picrotoxin infusions, go-new probabilities were higher than following muscimol infusions (p=0.006) and tended to be higher compared to saline infusions (p=0.063), with no differences between the latter (p=0.304).

Lose-shift probabilities were relatively stable during the last 5 responses preceding learning and were numerically higher in the saline condition (**Fig. 7B**), but this difference was not significant, with ANOVA not revealing any main effect or interaction involving drug (largest $F_{(3.139, 37.664)}$=2.180, p=0.104).

However, mPFC muscimol increased the probability of win-stay behavior (**Fig. 7C**) and of sticky behavior compared to saline and muscimol (**Fig. 7D**). ANOVA of win-stay probabilities revealed a main effect of drug ($F_{(2, 24)}$=5.922, p=0.008), alongside a trend for an effect of infusion series ($F_{(1, 12)}$=4.265, p=0.061) (all other main effects or interactions: largest $F_{(2.534, 30.408)}$=1.874, p=0.162). *Post-hoc* pairwise comparisons showed that the win-stay probabilities were higher following muscimol compared to saline (p=0.002) and tended to be higher compared to picrotoxin (p=0.095). Saline and picrotoxin did not differ (p=0.101). ANOVA of sticky behavior revealed a drug X response interaction ($F_{(2.849\ 34.186)}$=3.354, p=0.032), but no other main effects or interactions (largest $F_{(1,12)}$=2.757, p=0.123). Simple main effects analysis revealed an effect of drug at responses -5 and -4 before reaching criterion (smallest $F_{(2,24)}$=4.555, p=0.021) and a trend at responses -3 and -2 (largest $F_{(2,24)}$=2.902, p=0.074), but no drug effect at response -1 ($F_{(2.24)}$=2.012, p=0.156). *Post-hoc* pairwise comparisons indicated higher sticky probability in the muscimol condition compared to saline and picrotoxin at response -5 (largest p=0.044) and to saline (p=0.006), but not picrotoxin (p=0.119), at response -4. Saline and picrotoxin did not differ from one another at any point (smallest p=0.175). This suggests that, following mPFC muscimol, rats reached criterion during repeated reversal learning by using slightly different strategies than in the other two conditions.

DISCUSSION

During early reversals (R1 to R3) in Experiment 1, control rats showed a marked 'reversal cost' when they experienced a reversal for the first time, at R1. This was indicated by a clear

increase in RTC at R1 compared to the SD stage (**Fig. 3A**), and a high number of errors, especially perseverative errors (**Fig. 3B**), with both RTC and errors gradually declining again from R1 to R3, reflecting that rats were 'learning to reverse'. In contrast, during later serial reversals (R5 to R10) in Experiment 2, where rats had substantial experience of previous reversals and the opportunity to 'learn to reverse', control rats showed lower RTC than during the SD stage and fewer errors than during early reversals, with similar numbers of perseverative and regressive errors. Functional inhibition of the mPFC by muscimol impaired early-reversal performance, as reflected by increased RTC (**Fig. 3A**) and perseverative errors (**Fig. 3B**) during R2, but did not markedly affect serial-reversal performance (apart from a slight increase in win-stay and sticky behavior during responses leading up to criterion performance). This early-reversal impairment was associated with increased use of the old strategy (go-previous, **Fig. 4A**) and reduced exploration (lose-shift behavior, **Fig. 4B**), but not exploitation (win-stay, **Fig. 4C**), following reversal. In contrast, mPFC disinhibition by picrotoxin improved early reversal performance at R2, likely due to impaired expression of the previous response. However, disinhibition markedly disrupted serial-reversal performance, as reflected by increased RTC (**Fig. 5A**) and a rise in perseverative and regressive errors (**Fig. 5B & C**). This serial-reversal impairment by mPFC disinhibition was observed alongside the impaired expression of the old response, as reflected by reduced percentage of correct reminder-trial responses (**Fig. 5D**). It was associated with impairments in both lose-shift (**Fig. 6B**) and win-stay behavior (**Fig. 6C**), as well as an altered use of the old strategy around reversal (**Fig. 6A**).

mPFC functional inhibition impaired early, but not serial, reversals

Our finding that mPFC functional inhibition impaired early reversal learning, specifically at R2, but otherwise left reversal performance intact, confirms and extends several previous studies. In similar two-lever reversal procedures in rats, mPFC functional inactivation (by the sodium channel blocker bupivacaine) did not affect reversal performance at R1 (Floresco et al., 2008) and mPFC functional inhibition (by muscimol and baclofen) did not impair later serial-reversal performance (Hervig et al., 2020), but performance at R2 was not examined. Another study, testing the impact of cytotoxic mPFC lesions in rats performing two-odor serial reversals, reported findings strikingly similar to ours, namely a selective impairment at R2, but not at other reversal stages (testing was continued up to R8) (Kinoshita et al., 2008). Like mPFC functional inhibition in the present study, mPFC lesions did not impair performance beyond the 'normal' reversal cost at R1, but impaired rats' ability to overcome this reversal cost at R2. These findings are consistent with the idea that the mPFC is required to overcome prepotent

behavioral responses (Miller & Cohen, 2001; Haddon & Killcross, 2006; Marquis et al., 2007), in this case, pressing the previously correct lever.

In the present study, the reversal impairment at R2 reflected that mPFC functional inhibition increased perseverative errors and decreased lose-shift behavior. Previous studies reported that reversal learning impairments due to mPFC lesions on a three-lever spatial-discrimination serial-reversal task (Kosaki & Watanabe, 2012) or due to mPFC functional inhibition on a serial-spatial reversal task in a plus maze (Avigan et al., 2020) mainly reflected increased perseverative errors. In contrast, impairments on a touch-screen visual-pattern reversal task caused by mPFC lesions have consistently been reported to reflect mainly increased regressive errors (Bussey et al., 1997; Chudasama & Robbins, 2003; Brigman & Rothblat, 2008). Moreover, similar to the present study, reduced lose-shift behavior was reported following mPFC functional inactivation on a probabilistic two-lever spatial-reversal task (Dalton et al., 2016) and following mPFC lesions on bowl-digging set-shift or reversal tasks requiring olfactory and tactile discriminations (Wang et al., 2019). Interestingly, on the probabilistic two-lever spatial-reversal task, where only 80% of correct choices were rewarded, the reduced lose-shift behavior was associated with better reversal performance (Dalton et al., 2016), possibly reflecting that it was advantageous for rats to shift less readily following a non-rewarded response. In contrast, in the present study and on the bowl-digging set-shift and reversal tasks (Wang et al., 2019), where all correct responses were rewarded and incorrect responses were never rewarded, reduced lose-shift behavior was associated with impaired reversal or set-shift performance. These findings support that the mPFC can contribute to reversal performance by suppressing perseverative errors and sustaining lose-shift behavior. On tasks with deterministic reward contingencies, mPFC-dependent lose-shift behavior facilitates reversal performance (present study; Wang et al., 2019), whereas on tasks with probabilistic reward contingencies, such lose-shift behavior may hinder reversal performance (Dalton et al., 2016).

Previous studies suggested several other cognitive mechanisms by which the mPFC may contribute to reversal learning. First, the finding that mPFC lesions disrupt reversals on touch-screen visual pattern discrimination tasks, by increasing mainly regressive errors and only when difficult to discriminate visual stimuli were used, was suggested to reflect that the mPFC is important for visual attention (Bussey et al., 1997; Chudasama & Robbins, 2003; Brigman & Rothblat, 2008). However, attentional impairment is unlikely to play a main role in the reversal-learning impairments caused by mPFC functional inhibition in the present study. This is because mPFC functional inhibition selectively increased perseverative, but not regressive errors, and, although both mPFC functional inhibition and disinhibition caused attentional

impairments (Pezze et al., 2014; Auger et al., 2017), only functional inhibition impaired early reversal learning. Second, the requirement of the mPFC for serial spatial-reversal learning on plus-maze tasks was suggested to reflect that the mPFC is required for working memory (Kidder et al., 2024) and for reducing proactive interference based on hippocampus-dependent spatial memory (Guise & Shapiro, 2017; Avigan et al., 2020) (it should be noted that not all studies reported that mPFC was required for such spatial-reversal learning, Rich & Shapiro, 2007; Young & Shapiro, 2009). These cognitive mechanisms may also contribute to performance on our two-lever reversal task. However, it is difficult to see how disruption of these processes could account for the specific impairments at R2 that we observed following mPFC functional inhibition.

mPFC disinhibition impaired serial, but not early, reversals

mPFC disinhibition by picrotoxin markedly impaired serial reversal learning, manifested through increased RTC and increases in both perseverative and regressive errors, and impaired expression of the old rule on reminder trials. This dual impairment in both expressing the old rule and switching to the new rule was associated with impaired task-relevant exploration and exploitation, as reflected by reduced lose-shift and win-stay behavior. In contrast, mPFC disinhibition did not disrupt early reversals and even facilitated reversals at R2. The latter was at least partly due to a high number of omissions, which limited reminder-trial responses and reduced reinforcement of the old rule immediately before reversals. Our findings are consistent with a previous finding that mPFC disinhibition by bicuculline did not affect rats' reversal performance at R1 on a similar two-lever task, whereas the impact of mPFC disinhibition on further reversal stages were not examined (Enomoto et al., 2011).

Previous in vivo electrophysiological studies showed that neural disinhibition enhanced burst firing of mPFC neurons, whereas functional inhibition reduced firing, including burst firing, of mPFC neurons (Pezze et al., 2014). Only 'too little' mPFC activity caused by functional inhibition, but not 'too much' mPFC activity caused by disinhibition disrupted early-reversal performance, specifically at R2, whereas the opposite was true for serial reversal performance. In contrast, both functional inhibition and disinhibition of the mPFC caused attentional impairments (Pezze et al., 2014; Auger et al., 2017). This highlights that distinct cognitive and behavioral functions can show distinct relationships to mPFC activity, and that regional disinhibition can disrupt functions that do not require the disinhibited region (Bast et al., 2017). More specifically, whereas attentional performance requires balanced levels of mPFC activity, early reversals require mPFC activity, but not balanced levels of mPFC activity. Serial reversals do not require mPFC activity, but are disrupted by mPFC disinhibition,

presumably because disinhibition causes aberrant drive of mPFC projections to other brain regions and, thereby, disrupts processing in these projection sites (Bast et al., 2017).

Two mPFC projection sites that may be involved in mediating the serial-reversal impairments by mPFC disinhibition are the OFC (Sesack et al., 1989) and the dorsomedial striatum (McGeorge & Faull, 1989; Mailly et al., 2013). Lesion studies have long suggested that the OFC, especially the lateral OFC, is required for reversal learning, including serial-reversal learning (Murray et al., 2007). A more recent study reported that serial-reversal impairments caused by functional inactivation of the lateral OFC in rats were associated with impaired lose-shift behavior (Hervig et al., 2020). Lesions of the dorsomedial striatum also caused serial-reversal impairments due to increased perseveration (Clarke et al., 2008; Castane et al., 2010), whereas others reported early-reversal impairments due to increased regressive errors following functional inactivation of the dorsomedial striatum in rats on a T-maze task (Ragozzino & Choi, 2004). Moreover, Young et al. (2023) reported that increased activity within the dorsomedial striatum caused impairments on a probabilistic reversal task, specifically manifested as lose-shift impairments. Therefore, aberrant function within the lateral OFC and dorsomedial striatum may contribute to the serial-reversal impairments by mPFC disinhibition. The impact of mPFC disinhibition on mPFC projection sites remains to be examined in future studies.

Conclusion and clinical implications

Our findings suggest that reduced mPFC activation can impair early reversal learning by disrupting exploration necessary to overcome the early-reversal cost, whereas mPFC disinhibition can disrupt serial reversals by disrupting exploration and exploitation, possibly due to aberrant drive of mPFC projection sites, including OFC or striatum. Considering the functional-anatomical similarities between rodent mPFC and primate dlPFC (see *Introduction*), these results suggest that dlPFC disinhibition and hypofrontality could both contribute to reversal-learning deficits, which have been associated with many neuropsychiatric disorders (Izquierdo et al., 2017). This is particularly relevant to schizophrenia, which has been associated with both dlPFC disinhibition and hypofrontality (see *Introduction*).

Figure Legends

Figure 1

Reversal task and timelines of Experiment 1 and 2. A) Illustration of a reversal session, starting with 20 reminder trials followed by up to 200 reversal trials. During the 20 reminder trials at the start of each reversal stage, the same lever (either the left or the right lever) was rewarded as during testing on the previous day. After these 20 reminder trials, the reward associations of the levers were reversed, with the previously unrewarded lever now being rewarded and the previously rewarded lever being no longer rewarded. B) Sequence of procedures and operant test sessions and the timing of mPFC drug infusions, from the start of surgery until the end of each experiment. Red '^' indicates an mPFC microinfusion. In Experiment 1 (early reversals) the impact of mPFC saline (SAL), picrotoxin (PTX) and muscimol (MUS) infusions 10 min before the spatial discrimination (SD) stage and reversals 1 to 5 (R1 to R5) was compared between-subjects. In Experiment 2 (serial reversal), rats underwent SD training, followed by four reversals (R1 to R4) before the impact of mPFC SAL, PTX and MUS 10 min before five serial reversal sessions (reversals 5 to 10, R5 to R10) was tested within-subjects. In Experiment 2, retention days (Ret.) were included between reversal sessions, and rats completed two infusion series, where they underwent the same series of three different infusions (SAL, PTX and MUS) twice. Examples of between- (Experiment 1) and within-subjects (Experiment 2) infusion schedules for three example rats are included adjacent to the timelines.

Figure 2

Experiment 1 and 2: Infusion cannula placements in the mPFC. A) Cresyl-violet-stained coronal brain section depicting an injector tip placement in the prelimbic part of the mPFC. The photo shows a part of the guide cannula track and, underneath, the track left by the infusion cannula, with the white arrow indicating the tip. B) Approximate locations of infusion cannula tips (black dots) for Experiment 1 (separated by infusion group) and Experiment 2 shown on coronal plates adapted from the atlas by Paxinos and Watson (1998). Distance (in mm) anterior to bregma, according to the atlas, is indicated by numbers on the right.

Figure 3

Experiment 1: mPFC functional inhibition increased, whereas disinhibition reduced RTC and perseverative errors during reversal 2. RTC (A), perseverative errors (B), regressive errors (C) and percentage of correct responses during reminder trials (D) in the saline (white), muscimol (grey) and picrotoxin (black) groups across task stages (spatial discrimination, SD, and reversals 1 to 3, R1 to R3). Data are shown as mean+*SEM*. Significant differences and

trends as revealed by post-hoc pairwise comparisons are indicated: *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; #, $p=0.51$.

Figure 4

Experiment 1: Bayesian analysis of strategies around reversal: mPFC functional inhibition enhanced use of old strategy (go-previous) and disrupted exploration (lose-shift) at R2. Probabilities (mean+*SEM*) of different strategies are shown for the 6 reminder-trial responses preceding reversal and the 16 responses following reversal: go-previous (A), lose-shift (B), and win-stay (C). Grey-shaded region indicates reminder trials. Vertical dashed line indicates rule reversal. Asterisks indicate responses where saline and muscimol significantly differed ($p<0.05$).

Figure 5

Experiment 2: mPFC disinhibition, but not functional inhibition, impaired serial reversal learning. RTC (A), perseverative errors (B), regressive errors (C) and percentage of correct responses during reminder trials (D) following prefrontal infusions of saline (SAL), picrotoxin (PTX) or muscimol (MUS). Data are shown as mean+*SEM*. Significant differences (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$) and trends (#,$p=0.050$), as revealed by post-hoc pairwise comparisons, are indicated.

Figure 6

Experiment 2: Bayesian analysis of strategies around reversal: mPFC disinhibition reduced use of old strategy (go-previous) and disrupted both exploration (lose-shift) and exploitation (win-stay). Probabilities (mean+*SEM*) of different strategies are shown for the 6 reminder-trial responses preceding reversal and the 16 responses following reversal: go-previous (A), lose-shift (B), and win-stay (C) following mPFC infusions of saline (SAL), picrotoxin (PTX) or muscimol (MUS). Grey-shaded region indicates reminder trial performance. Vertical dashed line indicates rule reversal. Bracketed asterisk to the right of the line graphs in (A) indicates significant drug X response interaction ($p=0.008$). Asterisks above data points indicate significant ($p<0.05$) differences between infusion conditions at individual responses: (B) decrease in picrotoxin compared to muscimol and saline; (C) decrease in picrotoxin compared to muscimol and saline from response -4 to +4 and an increase in muscimol compared to picrotoxin and saline at response 16.

Figure 7

Experiment 2: Bayesian analysis of strategies preceding learning: mPFC disinhibition increased adherence to new rule (go-new) and functional inhibition increased win-stay and

sticky behavior. Go-new (A), lose-shift (B), win-stay (C), and sticky (D) probabilities (mean+*SEM*) are shown for the 5 responses preceding learning following mPFC infusions of saline (SAL), picrotoxin (PTX) or muscimol (MUS). Bracketed asterisks to the right of the line graphs indicate significant differences across all 5 responses ($p<0.05$): (A) increase in PTX compared to MUS condition; (C) increase in MUS compared to SAL condition. Asterisks above data points in (D) indicate significant differences ($p<0.05$) at individual responses: higher sticky probability in muscimol compared to both saline and picrotoxin conditions at response -5 and compared to saline at response -4

Table Legend

Table 1

Experiment 1 and 2: Omissions and response latencies.

Values (mean±*SEM*) are shown for the different task stages (spatial discrimination, SD; reversals 1 to 3, R1 to 3) in Experiment 1 and the two infusion series (series 1, reversals 5 to 7, R5 to 7; series 2, reversals 8-10, R8 to 10) in Experiment 2. Values during the 20 reminder trials at the beginning of each reversal stage are shown in brackets. Response latencies are indicated separately for correct responses (Lat_{Cor}) and incorrect responses (Lat_{Inc}). Sample sizes contributing to the mean values are indicated in the grey shaded rows. For reminder trials in Exp. 1, only one rat in the picrotoxin group completed at least one correct and one incorrect response; therefore, the picrotoxin group was excluded from the analysis of reminder-trial response latencies and no reminder trial response latencies are shown for this group. Moreover, in Experiment 1, two rats in the picrotoxin group did not make at least one incorrect response across all task stages and had to be excluded completely from the latency analysis, resulting in a reduced sample size of $n=12$ (instead of 14) for the latency means. For reminder trials in Experiment 2, one rat did not make at least one incorrect response during reminder trials in all three drug conditions in both series and, therefore, was excluded from the analysis of reminder trial response latencies. The reduced sample size contributing to reminder trial latency means in Experiment 2 is indicated in brackets.

References

Auger, M. L., & Floresco, S. B. (2015). Prefrontal cortical GABA modulation of spatial reference and working memory. *International Journal of Neuropsychopharmacology*, *18*(2), pyu013. https://doi.org/10.1093/ijnp/pyu013

Auger, M. L., Meccia, J., & Floresco, S. B. (2017). Regulation of sustained attention, false alarm responding, and implementation of conditional rules by prefrontal GABA(A) transmission: Comparison with NMDA transmission. *Psychopharmacology (Berl)*, *234*(18), 2777–2792. https://doi.org/10.1007/s00213-017-4670-1

Avigan, P. D., Cammack, K., & Shapiro, M. L. (2020). Flexible spatial learning requires both the dorsal and ventral hippocampus and their functional interactions with the prefrontal cortex. *Hippocampus*, *30*(7), 733–744. https://doi.org/10.1002/hipo.23198

Bast, T., & Feldon, J. (2003). Hippocampal modulation of sensorimotor processes. *Progress in Neurobiology*, *70*(4), 319–345. https://doi.org/10.1016/s0301-0082(03)00112-6

Bast, T., Pezze, M., & McGarrity, S. (2017). Cognitive deficits caused by prefrontal cortical and hippocampal neural disinhibition. *British Journal of Pharmacology*, *174*(19), 3211–3225. https://doi.org/10.1111/bph.13850

Benes, F. M., & Berretta, S. (2001). GABAergic interneurons: implications for understanding schizophrenia and bipolar disorder. *Neuropsychopharmacology*, *25*(1), 1–27. https://doi.org/10.1016/S0893-133X(01)00225-1

Boulougouris, V., Dalley, J. W., & Robbins, T. W. (2007). Effects of orbitofrontal, infralimbic, and prelimbic cortical lesions on serial spatial reversal learning in the rat. *Behavioural Brain Research*, *179*(2), 219–228. https://doi.org/10.1016/j.bbr.2007.02.005

Brady, A. M., & Floresco, S. B. (2015). Operant procedures for assessing behavioral flexibility in rats. *Journal of Visualized Experiments*, (96), e52387. https://doi.org/10.3791/52387

Brigman, J. L., & Rothblat, L. A. (2008). Stimulus-specific deficit on visual reversal learning after lesions of the medial prefrontal cortex in the mouse. *Behavioural Brain Research*, *187*(2), 405–410. https://doi.org/10.1016/j.bbr.2007.10.004

Brown, V. J., & Bowman, E. M. (2002). Rodent models of prefrontal cortical function. *Trends in Neurosciences*, *25*(7), 340–343. https://doi.org/10.1016/s0166-2236(02)02164-1

Bussey, T. J., Muir, J. L., Everitt, B. J., & Robbins, T. W. (1997). Triple dissociation of anterior cingulate, posterior cingulate, and medial frontal cortices on visual discrimination tasks using a touchscreen testing procedure for the rat. *Behavioral Neuroscience*, *111*(5), 920–936. https://doi.org/10.1037/0735-7044.111.5.920

Carter, C. S., Perlstein, W., Ganguli, R., Brar, J., Mintun, M., & Cohen, J. D. (1998). Functional hypofrontality and working memory dysfunction in schizophrenia. *American Journal of Psychiatry*, *155*(9), 1285–1287. https://doi.org/10.1176/ajp.155.9.1285

Castane, A., Theobald, D. E., & Robbins, T. W. (2010). Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behavioural Brain Research*, *210*(1), 74–83. https://doi.org/10.1016/j.bbr.2010.02.017

Chudasama, Y., & Robbins, T. W. (2003). Dissociable contributions of the orbitofrontal and infralimbic cortex to Pavlovian autoshaping and discrimination reversal learning: Further evidence for the functional heterogeneity of the rodent frontal cortex. *Journal of Neuroscience*, *23*(25), 8771–8780. https://doi.org/10.1523/JNEUROSCI.23-25-08771.2003

Clarke, H. F., Robbins, T. W., & Roberts, A. C. (2008). Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *Journal of Neuroscience*, *28*(43), 10972–10982. https://doi.org/10.1523/JNEUROSCI.1521-08.2008

Cunningham, M. G., Ames, H. M., Donalds, R. A., & Benes, F. M. (2008). Construction and implantation of a microinfusion system for sustained delivery of neuroactive agents. *Journal of Neuroscience Methods*, *167*(2), 213–220. https://doi.org/10.1016/j.jneumeth.2007.08.016

Dalley, J. W., Cardinal, R. N., & Robbins, T. W. (2004). Prefrontal executive and cognitive functions in rodents: Neural and neurochemical substrates. *Neuroscience and Biobehavioral Reviews, 28*(7), 771–784. https://doi.org/10.1016/j.neubiorev.2004.09.006

Dalton, G. L., Wang, N. Y., Phillips, A. G., & Floresco, S. B. (2016). Multifaceted contributions by different regions of the orbitofrontal and medial prefrontal cortex to probabilistic reversal learning. *Journal of Neuroscience, 36*(6), 1996–2006. https://doi.org/10.1523/JNEUROSCI.3366-15.2016

Deuel, R. K., Mishkin, M., & Semmes, J. (1971). Interaction between hemispheres in unimodal somesthetic learning. *Experimental Neurology, 30*(1), 123–132. https://doi.org/10.1016/0014-4886(71)90227-5

Dhawan, S. S., Tait, D. S., & Brown, V. J. (2019). More rapid reversal learning following overtraining in the rat is evidence that behavioural and cognitive flexibility are dissociable. *Behavioural Brain Research, 363*, 45–52. https://doi.org/10.1016/j.bbr.2019.01.055

Dias, R., Robbins, T. W., & Roberts, A. C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature, 380*(6569), 69–72. https://doi.org/10.1038/380069a0

Enomoto, T., Tse, M. T., & Floresco, S. B. (2011). Reducing prefrontal gamma-aminobutyric acid activity induces cognitive, behavioral, and dopaminergic abnormalities that resemble schizophrenia. *Biological Psychiatry, 69*(5), 432–441. https://doi.org/10.1016/j.biopsych.2010.09.038

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Floresco, S. B., Block, A. E., & Tse, M. T. (2008). Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behavioural Brain Research, 190*(1), 85–96. https://doi.org/10.1016/j.bbr.2008.02.008

Floresco, S. B., Ghods-Sharifi, S., Vexelman, C., & Magyar, O. (2006). Dissociable roles for the nucleus accumbens core and shell in regulating set shifting. *Journal of Neuroscience, 26*(9), 2449–2457. https://doi.org/10.1523/jneurosci.4431-05.2006

Goncalves, S., Hathway, G. J., Woodhams, S. G., Chapman, V., & Bast, T. (2023). No evidence for cognitive impairment in an experimental rat model of knee osteoarthritis and associated chronic pain. *Journal of Pain, 24*(8), 1478–1492. https://doi.org/10.1016/j.jpain.2023.04.002

Guise, K. G., & Shapiro, M. L. (2017). Medial prefrontal cortex reduces memory interference by modifying hippocampal encoding. *Neuron, 94*(1), 183–192.e8. https://doi.org/10.1016/j.neuron.2017.03.011

Haddon, J. E., & Killcross, S. (2006). Prefrontal cortex lesions disrupt the contextual control of response conflict. *Journal of Neuroscience, 26*(11), 2933–2940. https://doi.org/10.1523/JNEUROSCI.3243-05.2006

Hervig, M. E., Fiddian, L., Piilgaard, L., Bozic, T., Blanco-Pozo, M., Knudsen, C., Olesen, S. F., Alsio, J., & Robbins, T. W. (2020). Dissociable and paradoxical roles of rat medial and lateral orbitofrontal cortex in visual serial reversal learning. *Cerebral Cortex, 30*(3), 1016–1029. https://doi.org/10.1093/cercor/bhz144

Izquierdo, A., Brigman, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience, 345*, 12–26. https://doi.org/10.1016/j.neuroscience.2016.03.021

Kidder, K., Gillis, R., Miles, J., & Mizumori, S. J. Y. (2024). The medial prefrontal cortex during flexible decisions: Evidence for its role in distinct working memory processes. *Hippocampus, 34*(3), 141–155. https://doi.org/10.1002/hipo.23594

Kinoshita, S., Yokoyama, C., Masaki, D., Yamashita, T., Tsuchida, H., Nakatomi, Y., & Fukui, K. (2008). Effects of rat medial prefrontal cortex lesions on olfactory serial

reversal and delayed alternation tasks. *Neuroscience Research, 60*(2), 213–218. https://doi.org/10.1016/j.neures.2007.10.012

Kosaki, Y., & Watanabe, S. (2012). Dissociable roles of the medial prefrontal cortex, the anterior cingulate cortex, and the hippocampus in behavioural flexibility revealed by serial reversal of three-choice discrimination in rats. *Behavioural Brain Research, 227*(1), 81–90. https://doi.org/10.1016/j.bbr.2011.10.039

Krystal, J. H., & Anticevic, A. (2015). Toward illness phase-specific pharmacotherapy for schizophrenia. *Biological Psychiatry, 78*(11), 738–740. https://doi.org/10.1016/j.biopsych.2015.08.017

Laubach, M., Amarante, L. M., Swanson, K., & White, S. R. (2018). What, if anything, is rodent prefrontal cortex? *eNeuro, 5*(5), 1-14. https://doi.org/10.1523/ENEURO.0315-18.2018

Leeson, V. C., Robbins, T. W., Matheson, E., Hutton, S. B., Ron, M. A., Barnes, T. R., & Joyce, E. M. (2009). Discrimination learning, reversal, and set-shifting in first-episode schizophrenia: Stability over six years and specific associations with medication type and disorganization syndrome. *Biological Psychiatry, 66*(6), 586–593. https://doi.org/10.1016/j.biopsych.2009.05.016

Lilley, E., Stanford, S. C., Kendall, D. E., Alexander, S. P. H., Cirino, G., Docherty, J. R., George, C. H., Insel, P. A., Izzo, A. A., Ji, Y., Panettieri, R. A., Sobey, C. G., Stefanska, B., Stephens, G., Teixeira, M., & Ahluwalia, A. (2020). ARRIVE 2.0 and the British Journal of Pharmacology: Updated guidance for 2020. *British Journal of Pharmacology*, *177*(16), 3611–3616. https://doi.org/10.1111/bph.15178

Maggi, S., Hock, R. M., O'Neill, M., Buckley, M., Moran, P. M., Bast, T., Sami, M., & Humphries, M. D. (2024). Tracking subjects' strategies in behavioural choice experiments at trial resolution. *eLife*, *13*. https://doi.org/10.7554/eLife.86491

Mailly, P., Aliane, V., Groenewegen, H. J., Haber, S. N., & Deniau, J. M. (2013). The rat prefrontostriatal system analyzed in 3D: Evidence for multiple interacting functional units. *Journal of Neuroscience*, *33*(13), 5718–5727. https://doi.org/10.1523/JNEUROSCI.5248-12.2013

Marquis, J. P., Killcross, S., & Haddon, J. E. (2007). Inactivation of the prelimbic, but not infralimbic, prefrontal cortex impairs the contextual control of response conflict in rats. *European Journal of Neuroscience*, *25*(2), 559–566. https://doi.org/10.1111/j.1460-9568.2006.05295.x

McAlonan, K., & Brown, V. J. (2003). Orbital prefrontal cortex mediates reversal learning and not attentional set shifting in the rat. *Behavioral Brain Research*, *146*(1–2), 97–103. https://doi.org/10.1016/j.bbr.2003.09.019

McGeorge, A. J., & Faull, R. L. M. (1989). The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience*, *29*(3), 503–537. https://doi.org/10.1016/0306-4522(89)90128-0

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*(1), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, *66*(8), 811–822. https://doi.org/10.1001/archgenpsychiatry.2009.91

Murray, E. A., O'Doherty, J. P., & Schoenbaum, G. (2007). What we know and do not know about the functions of the orbitofrontal cortex after 20 years of cross-species studies. *Journal of Neuroscience*, *27*(31), 8166–8169. https://doi.org/10.1523/JNEUROSCI.1556-07.2007

Murray, G. K., Cheng, F., Clark, L., Barnett, J. H., Blackwell, A. D., Fletcher, P. C., Robbins, T. W., Bullmore, E. T., & Jones, P. B. (2008). Reinforcement and reversal learning in first-episode psychosis. *Schizophrenia Bulletin*, *34*(5), 848–855. https://doi.org/10.1093/schbul/sbn078

Paxinos, G., & Watson, C. (1998). *The rat brain in stereotaxic coordinates* (4th ed.). New York: Academic Press.

Pezze, M., McGarrity, S., Mason, R., Fone, K. C., & Bast, T. (2014). Too little and too much: Hypoactivation and disinhibition of medial prefrontal cortex cause attentional deficits. *Journal of Neuroscience*, *34*(23), 7931–7946. https://doi.org/10.1523/JNEUROSCI.3450-13.2014

Ragozzino, M. E., & Choi, D. (2004). Dynamic changes in acetylcholine output in the medial striatum during place reversal learning. *Learning & Memory*, *11*(1), 70–77. https://doi.org/10.1101/lm.65404

Rao, S. G., Williams, G. V., & Goldman-Rakic, P. S. (2000). Destruction and creation of spatial tuning by disinhibition: GABA(A) blockade of prefrontal cortical neurons engaged by working memory. *Journal of Neuroscience*, *20*(1), 485–494. https://doi.org/10.1523/JNEUROSCI.20-01-00485.2000

Renstrom, J. G. (2024). *Prefrontal GABAergic inhibition and reversal learning* [University of Nottingham].

Rich, E. L., & Shapiro, M. L. (2007). Prelimbic/infralimbic inactivation impairs memory for multiple task switches, but not flexible selection of familiar tasks. *Journal of Neuroscience*, *27*(17), 4747–4755. https://doi.org/10.1523/JNEUROSCI.0369-07.2007

Schoenbaum, G., Setlow, B., Nugent, S. L., Saddoris, M. P., & Gallagher, M. (2003). Lesions of orbitofrontal cortex and basolateral amygdala complex disrupt acquisition of odor-guided discriminations and reversals. *Learning & Memory*, *10*(2), 129–140. https://doi.org/10.1101/lm.55203

Schoonover, K. E., Dienel, S. J., & Lewis, D. A. (2020). Prefrontal cortical alterations of glutamate and GABA neurotransmission in schizophrenia: Insights for rational biomarker development. *Biomarkers in Neuropsychiatry*, *3*, 100015. https://doi.org/10.1016/j.bionps.2020.100015

Seamans, J. K., Lapish, C. C., & Durstewitz, D. (2008). Comparing the prefrontal cortex of rats and primates: Insights from electrophysiology. *Neurotoxicity Research*, *14*(2–3), 249–262. https://doi.org/10.1007/bf03033814

Sesack, S. R., Deutch, A. Y., Roth, R. H., & Bunney, B. S. (1989). Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: An anterograde tract-tracing study with *Phaseolus vulgaris leucoagglutinin*. *Journal of Comparative Neurology*, *290*(2), 213–242. https://doi.org/10.1002/cne.902900205

Smith, K. S., Bucci, D. J., Luikart, B. W., & Mahler, S. v. (2016). DREADDS: Use and application in behavioral neuroscience. *Behavioral Neuroscience*, *130*(2), 137–155. https://doi.org/10.1037/bne0000135

Tse, M. T., Piantadosi, P. T., & Floresco, S. B. (2015). Prefrontal cortical gamma-aminobutyric acid transmission and cognitive function: Drawing links to schizophrenia from preclinical research. *Biological Psychiatry*, *77*(11), 929–939. https://doi.org/10.1016/j.biopsych.2014.09.007

Uylings, H. B. M., Groenewegen, H. J., & Kolb, B. (2003). Do rats have a prefrontal cortex? *Behavioral Brain Research*, *146*(1–2), 3–17. https://doi.org/10.1016/j.bbr.2003.09.028

Wang, J., Tait, D. S., Brown, V. J., & Bowman, E. M. (2019). Exacerbation of the credit assignment problem in rats with lesions of the medial prefrontal cortex is revealed by Bayesian analysis of behavior in the pre-solution period of learning. *Behavioral Brain Research*, *372*, 112037. https://doi.org/10.1016/j.bbr.2019.112037

Wang, J., Tait, D. S., Brown, V. J., & Bowman, E. M. (2019). Exacerbation of the credit assignment problem in rats with lesions of the medial prefrontal cortex is revealed by Bayesian analysis of behavior in the pre-solution period of learning. *Behavioural Brain Research*, *372*, 112037. https://doi.org/10.1016/j.bbr.2019.112037

Young, J. J., & Shapiro, M. L. (2009). Double dissociation and hierarchical organization of strategy switches and reversals in the rat PFC. *Behavioral Neuroscience*, *123*(5), 1028–1035. https://doi.org/10.1037/a0016822

Young, M. K., Conn, K. A., Das, J., Zou, S., Alexander, S., Burne, T. H. J., & Kesby, J. P. (2023). Activity in the dorsomedial striatum underlies serial reversal learning

performance under probabilistic uncertainty. *Biological Psychiatry Global Open Science*, *3*(4), 1030–1041. https://doi.org/10.1016/j.bpsgos.2022.08.005

Figures

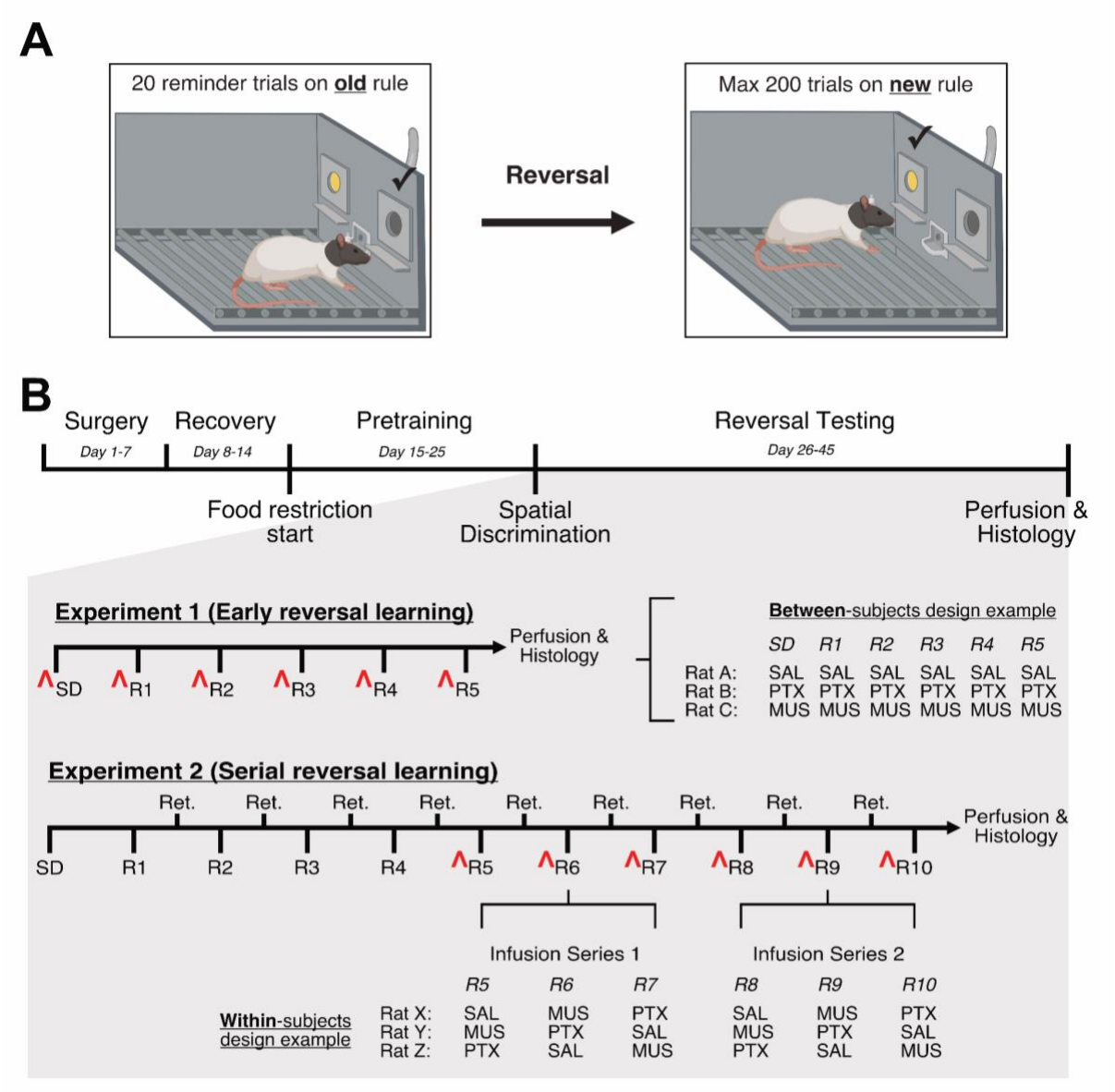

Figure 1. Reversal task and timeline of Experiment 1 and 2.

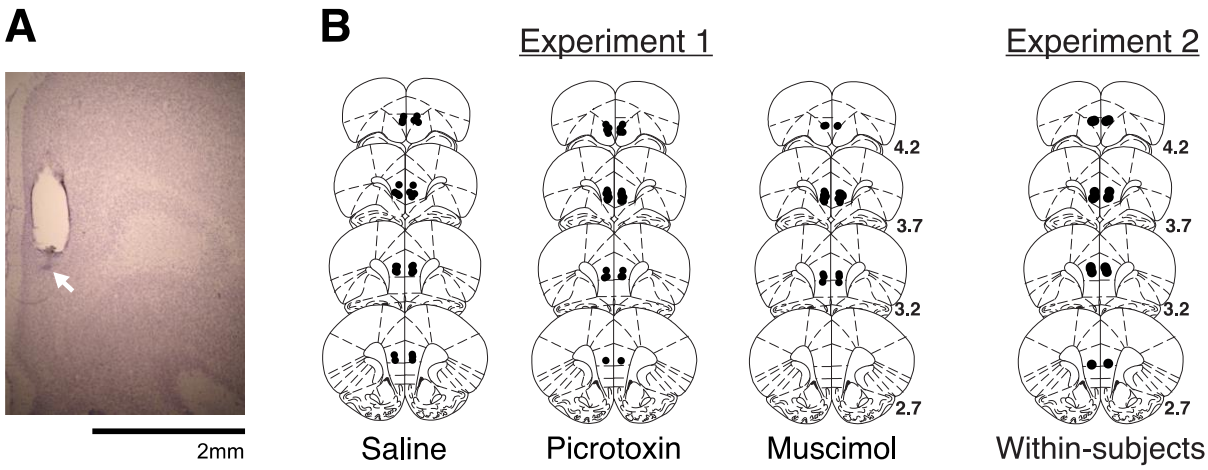

Figure 2. Experiment 1 and 2: Infusion cannula placements in the mPFC.

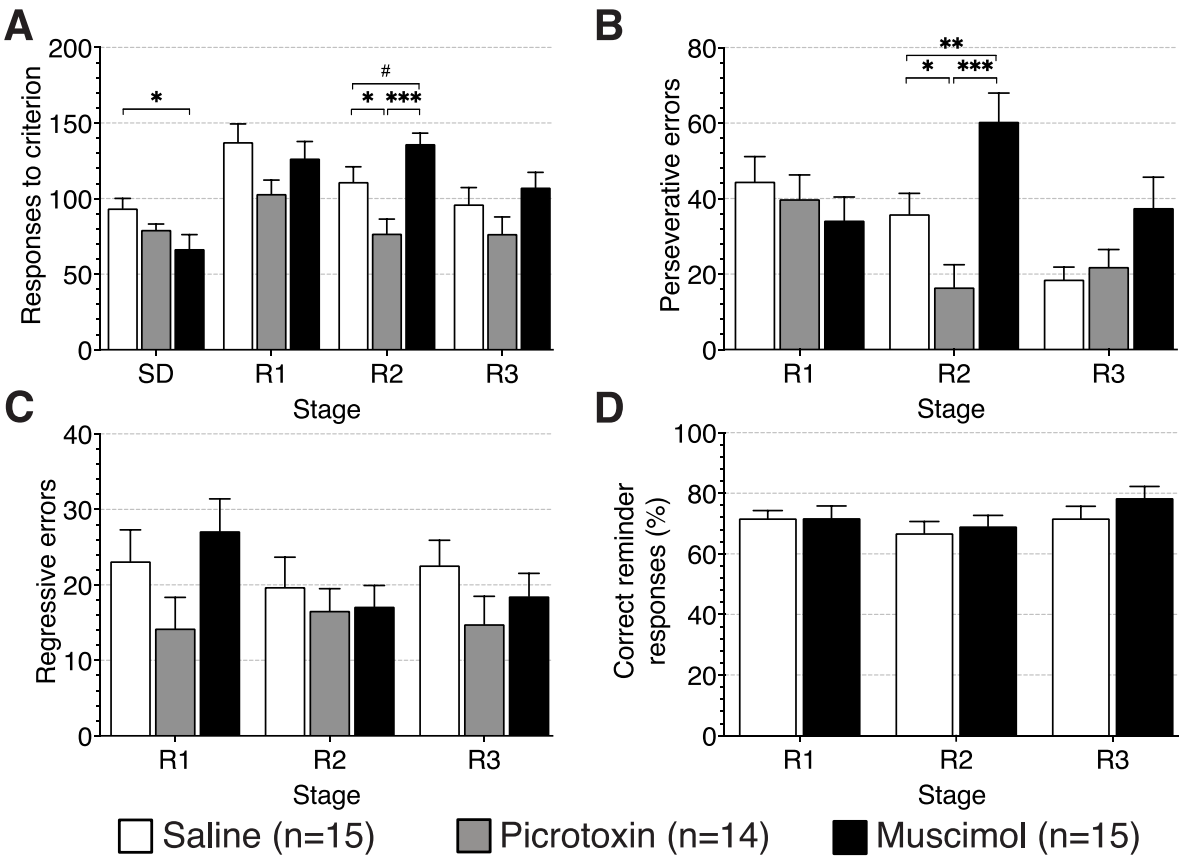

Figure 3. Experiment 1: mPFC functional inhibition increased RTC and perseverative errors during reversal 2.

	mPFC drug infusion								
	Saline (0.9%)			**Picrotoxin (300ng)**			**Muscimol (62.5ng)**		
	Omissions	Lat_{Cor} (s)	Lat_{Inc} (s)	Omissions	Lat_{Cor} (s)	Lat_{Inc} (s)	Omissions	Lat_{Cor} (s)	Lat_{Inc} (s)
Exp. 1	*n=15*	*n=15*	*n=15*	*n=14*	*n=12*	*n=12*	*n=15*	*n=15*	*n=15*
SD	9.00±5.66	0.99±0.07	1.31±0.09	108.92±17.91	2.64±0.18	2.94±0.19	2.40±0.84	1.17±0.09	1.32±0.10
R1	(0.33±0.16) 1.13±0.36	(1.08±0.14) 0.94±0.10	(1.15±0.23) 1.10±0.11	(14.5±1.41) 65.14.±13.15	1.98±0.25	2.75±0.21	(0.53±0.19) 3.20±0.86	(1.18±0.12) 1.15±0.09	(1.75±0.23) 1.24±0.12
R2	(0.07±0.07) 0.53±0.17	(0.99±0.15) 1.07±0.11	(0.96±0.09) 1.06±0.11	(17.7±1.22) 58.36±7.65	2.24±0.27	1.92±0.12	(0.13±0.09) 2.80±0.67	(1.06±0.17) 1.20±0.11	(1.08±0.10) 1.26±0.08
R3	(0.27±0.21) 1.07±0.55	(0.94±0.10) 0.96±0.10	(1.03±0.19) 0.99±0.10	(15.6±1.94) 49.50±11.24	1.65±0.19	1.81±0.23	(0.33±0.27) 1.87±0.92	(1.20±0.16) 1.20±0.13	(1.53±0.32) 1.57±0.22
Exp. 2	*(n=13)* *n=13*	*(n=12)* *n=13*	*(n=12)* *n=13*	*(n=13)* *n=13*	*(n=12)* *n=13*	*(n=12)* *n=13*	*(n=13)* *n=13*	*(n=12)* *n=13*	*(n=12)* *n=13*
Series 1	(0.15±0.10) 0.46±0.18	(0.66±0.08) 0.59±0.05	(0.87±0.17) 0.59±0.05	(13.62±6.77) 34.77±8.29	(2.65±0.27) 2.37±0.35	(2.42±0.15) 2.99±0.39	(0.23±0.17) 1.00±0.36	(0.67±0.10) 0.66±0.10	(0.76±0.14) 0.86±0.14
Series 2	(0.31±0.21) 0.38±0.21	(0.65±0.09) 0.50±0.04	(0.70±0.19) 0.66±0.10	(2.77±0.79) 23.54±7.64	(2.77±0.36) 1.44±0.18	(2.39±0.44) 1.95±0.21	(0.15±0.10) 0.31±0.13	(0.59±0.08) 0.69±0.06	(1.10±0.27) 0.63±0.08

Table 1. Experiment 1 and 2: Omissions and response latencies.

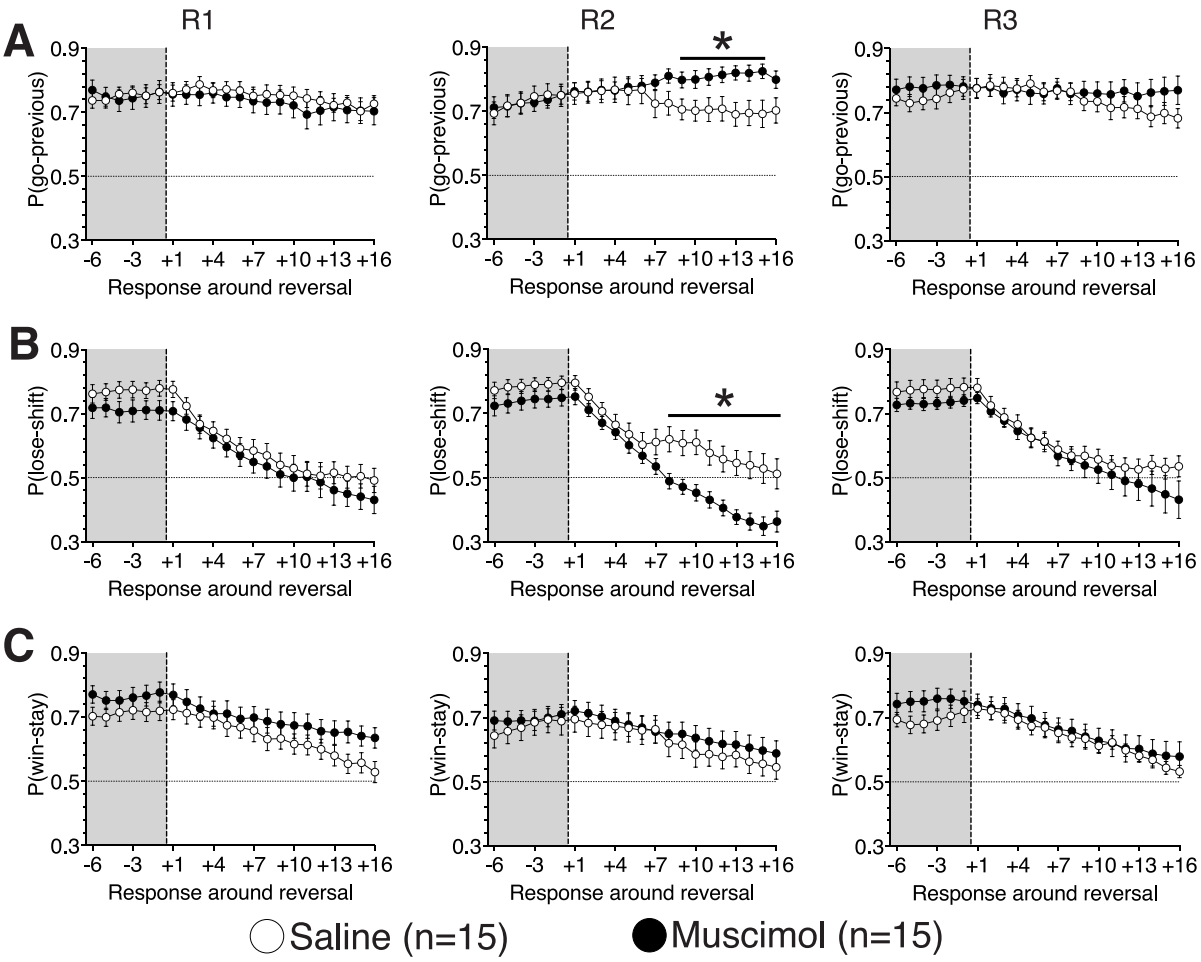

Figure 4. Experiment 1: Bayesian analysis of strategies around reversal: mPFC functional inhibition enhanced use of old strategy (go-previous) and disrupted exploration (lose-shift) at R2.

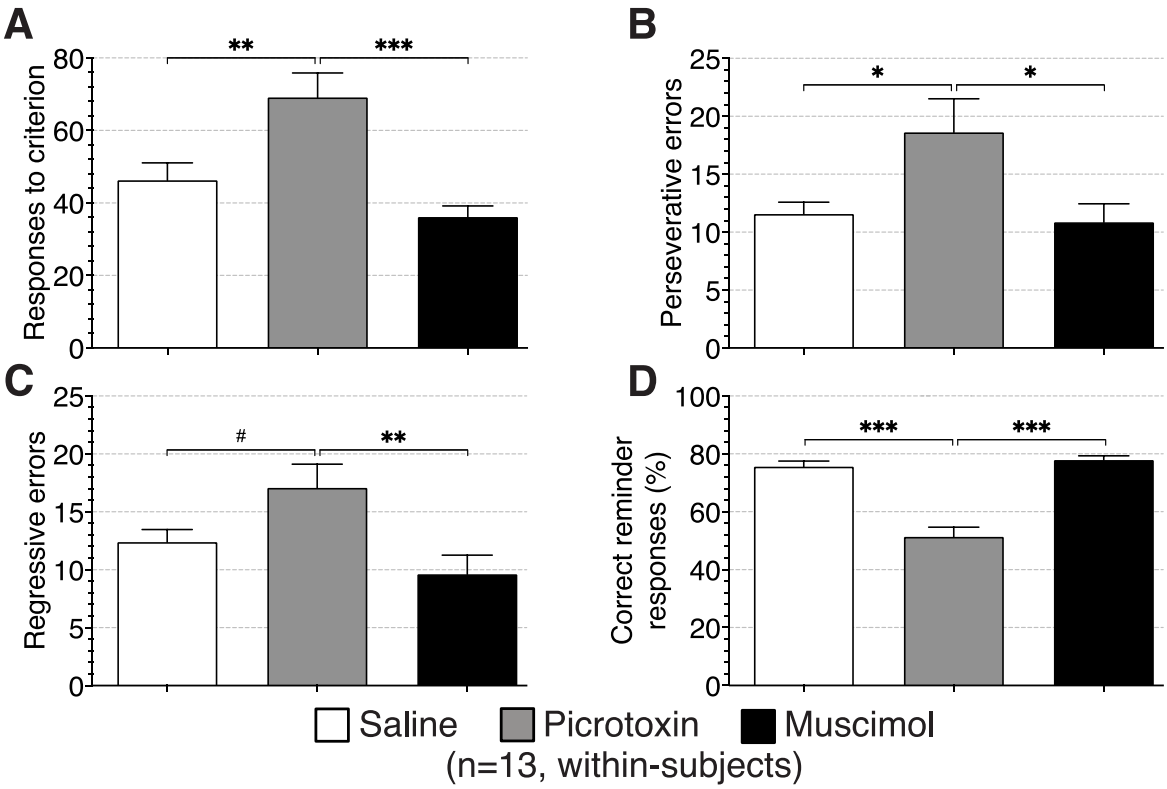

Figure 5. Experiment 2: mPFC disinhibition, but not functional inhibition, impaired serial reversal learning.

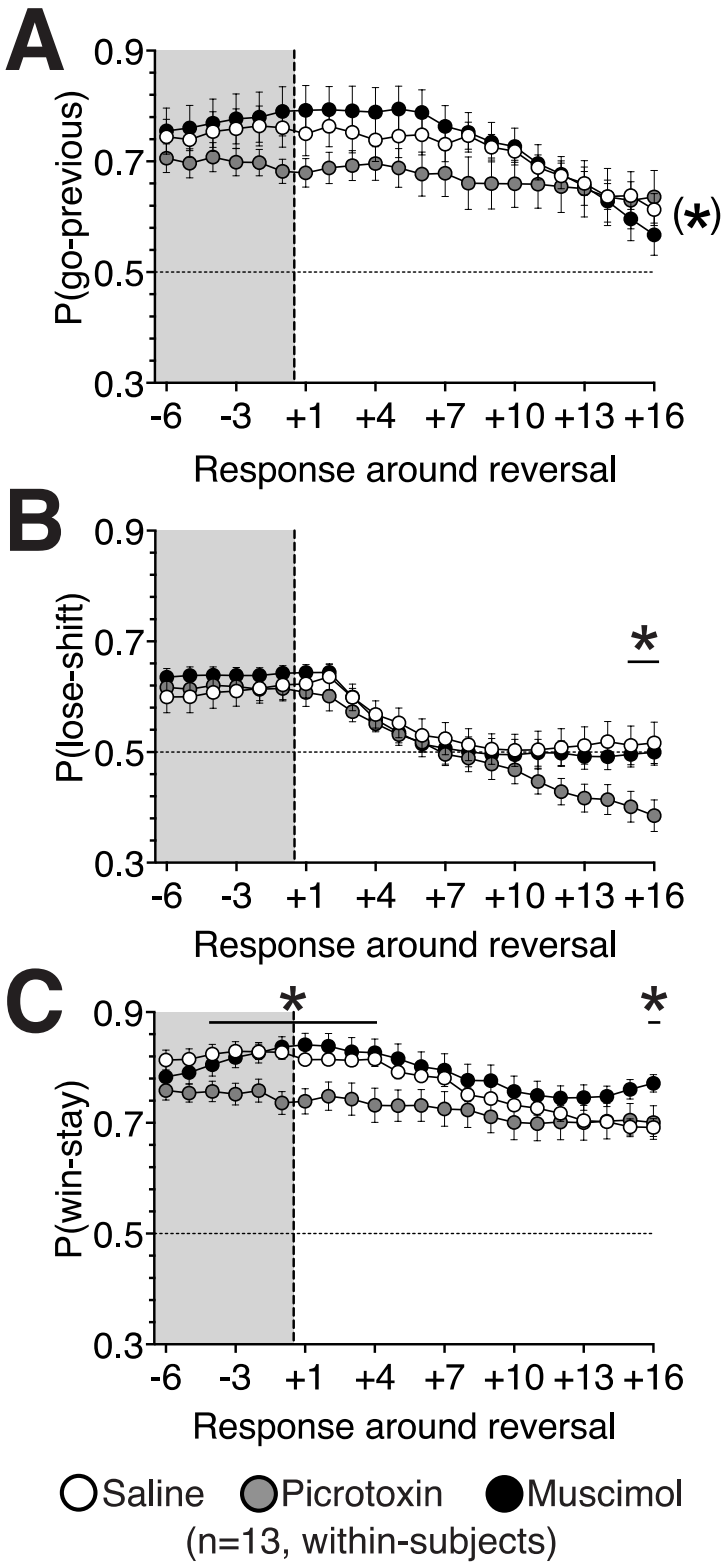

Figure 6. Experiment 2: Bayesian analysis of strategies around reversal: mPFC disinhibition reduced use of old strategy (go-previous) and disrupted both exploration (lose-shift) and exploitation (win-stay).

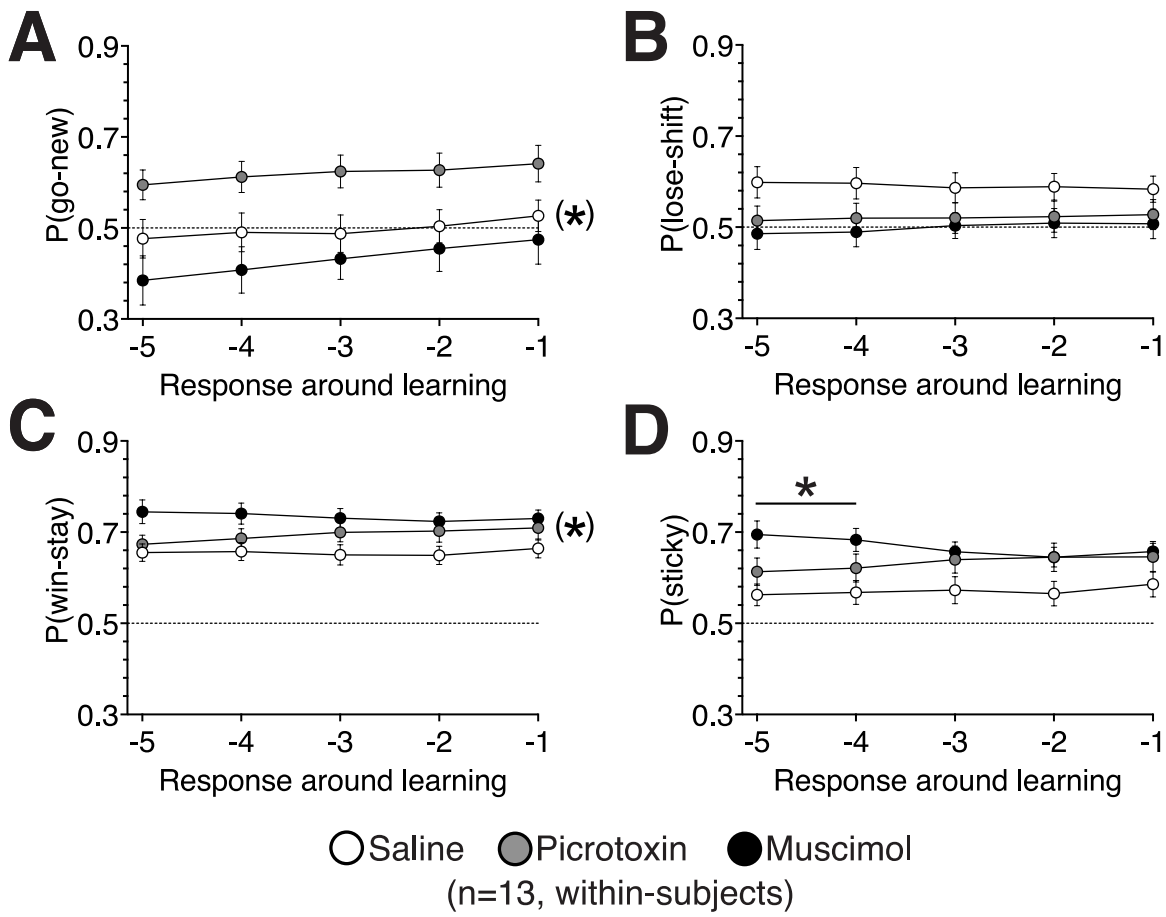

Figure 7. Experiment 2: Bayesian analysis of strategies preceding learning: mPFC disinhibition increased adherence to new rule (go-new) and functional inhibition increased win-stay and sticky behavior.