Neuroscience 555 (2024) 134–144

Contents lists available at ScienceDirect

Neuroscience

journal homepage: www.elsevier.com/locate/nsc

Research Article

"Blocking-like" effects in attentional set-shifting: Redundant cues facilitate shifting in male rats with medial prefrontal cortex inactivation

Tegan S. Knott [1], Alonzo J. Whyte [1], Sandeep S. Dhawan, David S. Tait, Verity J. Brown [*]

School of Psychology and Neuroscience, University of St Andrews, St Mary's Quad, South Street, St Andrews KY16 9JP, UK

ARTICLE INFO

Keywords:
Sensory gating
Aberrant salience
Kamin blocking

ABSTRACT

Without a functioning prefrontal cortex, humans and other animals are impaired in measures of cognitive control and behavioral flexibility, including attentional set-shifting. However, the reason for this is unclear with evidence suggesting both impaired and enhanced attentional shifting. We inhibited the medial prefrontal cortex (mPFC) of rats while they performed a modified version of an attentional set-shifting task to explore the nature of this apparent contradiction.

Twelve adult male Lister hooded rats received AAV5-CaMKIIa-hM4D(Gi)-mCherry viral vector bilaterally into mPFC to express inhibitory 'Designer Receptors Exclusively Activated by Designer Drugs' (iDREADDs). The receptors were activated by systemic clozapine *N*-oxide (CNO) to inhibit mPFC function. The rats were tested in the standard attentional set-shifting task four times: twice after i.p. administration and twice after oral administration of vehicle or CNO (10 mg/kg). They were then tested twice in a modified task, with or without oral CNO. The modified task had an extra stage before the extradimensional shift, in which the relevant exemplars remained relevant and new exemplars that were fully predictive but redundant replaced the previous irrelevant exemplars. These exemplars then became relevant at the subsequent ED stage.

In the standard task, mPFC inactivation impaired attentional set-shifting, consistent with previous findings. However, in the modified task, mPFC inactivation abolished ED shift-costs. The results support the suggestion that the mPFC is needed for the downregulation of attention that prevents learning about redundant and irrelevant stimuli. With mPFC inactivated, the rat learns more rapidly when previously redundant exemplars become the only relevant information.

Introduction

An organism's choices and decisions are based on expected outcomes, which are mental models of the world that have been built from prior experience of the outcomes that have followed actions or events in the past. The mental models need to be a good enough fit with reality to support adaptive choices, including having a representation of expected 'noise' because outcomes are not always fully predictable. This means they must be monitored for goodness-of-fit and updated when there is new information, but not necessarily abandoned immediately when there is a violation of a prediction. Prediction error (PE) refers to the mismatch between an expected state and reality, and neuronal encoding of PE is found throughout the brain (for review see Den Ouden et al., 2012). In models of reinforcement learning, PE drives new learning (Rescorla and Wagner, 1972), including reflecting valance: it indicates more than merely an unexpected outcome, but also whether the outcome is more or less than expected (Schultz and Dickinson, 2000). On the other hand, what is learned also depends upon the attention allocated to the to-be-conditioned cues (Mackintosh, 1975; Pearce and Hall, 1980; Le Pelley and McLaren, 2003). Additionally, the salience of cues is driven by PE, and this PE signal is also more than just a surprise signal but also has representational content (Den Ouden et al., 2012).

Although PE signals are ubiquitous throughout the brain, specific brain regions have been reliably associated with different aspects of the

Abbreviations: Cg, cingulate cortex; CNO, clozapine *N*-oxide; CD, compound discrimination; DREADDs, Designer Receptors Exclusively Activated by Designer Drugs; ED, extradimensional; ID, intradimensional; IL, infralimbic cortex; MO, medial orbitofrontal cortex; mPFC, medial prefrontal cortex; PBS, phosphate buffered saline; PE, prediction error; PrL, prelimbic cortex; REV, reversal; RS, redundant exemplar stage; SD, simple discrimination.

* Corresponding author.
E-mail addresses: dst@st-andrews.ac.uk (D.S. Tait), vjb@st-andrews.ac.uk (V.J. Brown).

[1] Joint first authors.

https://doi.org/10.1016/j.neuroscience.2024.07.034
Received 26 April 2024; Accepted 20 July 2024
Available online 25 July 2024

updating of mental models which give rise to cognitive and behavioral flexibility. A widely used test of cognitive flexibility in many different species is the intradimensional/extradimensional (ID/ED) attentional set-shifting task. This task involves a series of stages involving two-choice discriminations between complex, multidimensional stimuli (for example, visual stimuli that differ in both colour and shape or physical stimuli differing in odor and texture). Only one aspect of the stimuli is relevant for solving the discriminations and an attentional bias to the relevant features, referred to as an attentional set, forms during the initial stages of testing. The task typically includes reversal learning stages, an ID stage (where there is learning of novel exemplars in the same dimension as the current attentional set) and, at the ED stage of the test, a requirement to shift attentional set when the stimulus dimension relevant to solving the task changes. The number of additional trials to learn the ED discrimination, compared to learning at the ID stage, indicates the 'cost' of shifting attention.

The ID/ED task has been referred to as *the* attentional set-shifting task, particularly in the rodent literature, but there are many other tasks measuring behavioral flexibility, including rule switching, strategy shifting or response reversal tasks (Ghods-Sharifi et al., 2008; Gilmour et al., 2013; Brady and Floresco, 2015; Brown and Tait, 2015; Izquierdo et al., 2017). Across species, reversal learning impairments are often associated with damage to orbital prefrontal cortex, while deficits in attentional flexibility are frequently associated with impaired dorsolateral prefrontal cortex in monkeys or the medial prefrontal cortex (mPFC) in rats (for reviews see Chudasama and Robbins, 2006; Robbins, 2007, 2017; Keeler and Robbins, 2011). There is a long running debate about whether any part of rodent prefrontal cortex is homologous to primate dorsolateral prefrontal cortex (for review see Laubach et al., 2018). We have previously argued that a focus on functional similarities across species is not denying the clear anatomical differences (Brown and Bowman, 2002) and we are using the mPFC to denote where it is, rather than what it is.

There are, however, intriguing inconsistencies in both the human and animal literature. For example, although ED shifting deficits have been frequently observed in patients with schizophrenia (Elliott et al., 1995; Pantelis et al., 1997, 1999; Jazbec et al., 2007; Ceaser et al., 2008; Leeson et al., 2009; Waltz et al., 2013), it has been suggested that impairments in latent inhibition in patients with schizophrenia reflect 'hyperactive switching' (Weiner, 1990; Weiner and Feldon, 1997). Patients with first episode schizophrenia (Chu et al., 2021) and people scoring high on schizotypy (Le Pelley et al., 2010) have reduced learning benefit from previously relevant cues and increased learning cost with previously irrelevant cues, which was interpreted as reduced ability to ignore cues that are irrelevant. In a similar vein, it has been proposed that psychosis is associated with, and possibly arises from, 'aberrant salience' (Kapur, 2003; Roiser et al., 2009), or too much attention, assigned to neutral or irrelevant information. Aberrant salience has also been reported in people scoring high on schizotypy (Haselgrove et al., 2016). In the rodent literature, there is a similar apparent contradiction: rats with inactivation of the mPFC – the same brain area associated with impaired ED shifting – show more rapid conditioning to a previously blocked stimulus, interpreted as an inability to downregulate attention to irrelevant cues (Sharpe and Killcross, 2014, 2018), which has also been called learned inattention (Kruschke and Blair, 2000) or learned associability (Le Pelley and McLaren, 2003). However, the suggestion that there is an impairment of learned inattention, or an inability to downregulate attention, is not easily reconciled with an impairment in shifting attention *to* previously irrelevant cues in an ID/ED task. Sharpe and Killcross (2014) suggested that the downregulation impairment might be manifest in the ID/ED task only once an attentional set has formed: after learning, an inability to downregulate attention would result in continued focus on reinforced cues, impairing new learning when those cues are no longer relevant. On the other hand, it is possible that impairments in shifting aptitude at the ED stage of this test are due to the psychometric characteristics of this stage (Barch et al., 2009) such that the apparent contradiction between evidence suggesting more rapid, or 'hyper-,' shifting in some contexts, and impaired ED shifting in others, could be accounted for by a single context- dependent deficit. In this study, we tested these ideas by manipulating the psychometric characteristics of the ED stage.

We first tested rats in the standard, 7-stage ID/ED task, with and without inactivation of the mPFC by iDREADDs + CNO, to replicate the ED shift deficit reliably reported following cell-body lesions of this area. We then tested the same rats twice, with and without inactivation of mPFC, in a modified version of the task, which introduced redundant but predictive cues prior to the ED. Those redundant cues then became the sole predictor of reward at the ED shift. We hypothesised that this manipulation would enable us to observe more rapid learning about previously redundant cues, which should result in more rapid acquisition of the ED stage. By contrast, if the slower ED shift following mPFC inactivation is due to an inability to shift attention away from the relevant dimension, manipulations of the predictive value of the cues in the irrelevant dimension would have no effect.

Experimental procedures

Animals

We used adult male (n = 24) Lister hooded rats (Charles River, UK). The number of animals was determined by the requirement to counterbalance the order and direction of shifts. The rats were pair-housed in cages within 'Scantainer Classic' units (https://www.scanbur.com) maintained at 21 °C±2°C and a humidity of 55 % ± 5 %, with sawdust bedding and toys for environmental enrichment (e.g., wooden chew bar or wooden ball.) Behavioral testing was completed during the light phase of the light–dark cycle, 07:00–19:00 hr. Water was available *ad libitum* throughout, but laboratory chow was restricted to 15–20 g per rat per day. At time of surgery, rats weighed 395 to 430 g; at perfusion, they weighed 445 to 525 g. Although we used only male rats in this experiment, we and others (Mohamed et al., 2011; Snigdha et al., 2011; Murphy et al., 2017) have observed consistencies in the pattern of data across strains and sex in this task.

All experiments were conducted in accordance with the regulations laid down in the United Kingdom Animals (Scientific Procedures) Act 1986 and performed with the authority of a UK Home Office Project Licence. We followed the 'Essential 10′ ARRIVE guidelines (du Sert et al., 2020) in the design and reporting of this study. We also adhered to the Recommendations, apart from Recommendation 19: although we did not register the protocol, the protocol was approved in advance by the University of St Andrews Animal Welfare and Ethics Committee.

Chemogenetic manipulation

Rats were anaesthetised with isoflurane (5 % induction; 2 % maintenance) in oxygen. A 0.05 ml dose of the non-steroidal anti-inflammatory drug Carprieve® (carprofen; Pfizer, UK) was administered via subcutaneous injection. They were placed into a stereotaxic frame (Kopf, CA, USA) using atraumatic ear bars, and tooth bar set at − 3.3 (level skull). A burr hole was drilled over two injection sites per hemisphere to allow infusions into the mPFC at (coordinates with respect to bregma): AP+3.9, ML±0.5 and DV−3.1 (from dura); AP+2.9, ML±0.5, DV−3.1 (from dura). Using a Hamilton syringe, the infusions were made over 5 min, with the syringe left *in situ* for 5 min post-injection. Microinjections of 1 µl AAV5-CamKII-hM4Di-mCherry Designer Receptors Exclusively Activated by Designer Drugs (DREADDs; 3.4×10^{12} µg/ml; pAAV-CaMKIIa-hM4D(Gi)-mCherry) were administered at each of the four sites, with the dose and location selected based on our preliminary studies (data not shown) to transfect cells across an area as large as our previous cell-body lesions. The DREADDs were a gift from Bryan Roth (Addgene viral prep # 50477-AAV5; https://n2t.net/addgene:50477; RRID:Addgene 50477). Eighteen rats received virus injections, and six

rats received a sham procedure wherein the needle was lowered, but nothing was injected. Following surgery, rats were single housed for 24 hr. Twelve of the virus-administered rats then undertook the attentional set-shifting task as described below. The remaining six virus-administered rats and the six sham-administered rats were sacrificed to investigate Fos expression as described in section 2.5 below.

Rats will rapidly consume a gelatine 'gummy' to which drugs can be added and this method is less stressful than oral gavage and successfully delivers a pharmacokinetic profile suitable for behavioral testing (Dhawan et al., 2018; Ferrari et al., 2022). To establish if it would be possible to deliver CNO orally using this method, we compared the behavioral profile following intraperitoneal (i.p.) injection and oral administration.

CNO (Sequoia Research Product Ltd, UK) was administered at 10 mg/kg, which is at the higher end of the effective range and active for approximately 5 hrs (Roth, 2016), which was sufficient time to complete testing.

For two days prior to the first test, rats were injected with saline vehicle to habituate them to the i.p. procedure. CNO was dissolved in saline at 10 mg/ml and administered 10 mg/kg i.p. 30 min prior to behavioral testing. We created gelatine tablets for individual rats, based on their weight to contain CNO at 10 mg/kg. We suspended the appropriate dose in 1.5 ml of the vehicle solution (60 ml of sugar-free blackcurrant flavored juice (Robinsons, Britvic, UK) and 12 g of gelatine powder (Dr. Oetker, UK) pipetted into a plastic mould, then refrigerated until set). Rats were habituated to vehicle gelatine gummies over several days, until they were consuming them within 1 min. The gelatine gummies were presented for the rats to consume 30 min prior to the start of behavioral testing. The tester was not blind to the treatment condition.

Behavioral testing

The ID/ED attentional set-shifting task training and testing have been described previously (Tait et al., 2018). Briefly, a modified plastic housing-cage (69.5 × 40.5 × 18.5 cm) had two individually compartmented chambers in which ceramic bowls containing digging material and reward were placed. At least 12 hr prior to training, each rat was given a bowl filled with home-cage sawdust and ~ six pieces of food reward (a Honey Loop cereal piece; Kellogg, UK) in the home-cage. Following exposure to the reward, rats were trained to dig in bowls filled with sawdust to obtain a food reward (half a Honey Loop) within the testing chambers of the arena. To shape this response, the reward was placed on top of the sawdust of each of two bowls. After the rat retrieved both rewards, the reward was placed slightly deeper in the sawdust on each subsequent presentation. This continued until the reward was completely buried at the bottom of the bowls and the rat was reliably retrieving it on each presentation. This 'digging' training regime was typically completed in six presentations. All rats were then trained on two simple discriminations (SD) using the same exemplars for both odor (sawdust scented with mint or oregano) and medium (shredded paper and polystyrene) discriminations. These exemplars were not used again.

On each testing day, the rats started with a simple discrimination and then completed a series of stages, each involving learning a novel discrimination or reversal of that learning, of compound discriminations. The compound stimuli were a pair of bowls containing distinctive digging media (medium exemplars, M) and added odors (odor exemplars, O). The reward was associated with one of the four exemplars (i.e., one of the odors or one of the digging media) and the exemplars in the other dimension were irrelevant, being pseudo-randomly associated with the reward.

For the first four trials of each stage, rats were permitted to obtain the food reward from the baited bowl after an incorrect dig. These four trials were presented in a standard order, albeit with the exemplar associated with reward counterbalanced across rats. In trials 1 and 2, the same exemplar pairings were used, with the rewarded bowl first on one side and then on the other. In trials 3 and 4, the second exemplar pairing was presented, again with the sides alternating. For example, two bowls containing M1 + O1 or M2 + O2 were presented, with M1 + O1 on the right (trial 1) and then on the left (trial 2). For the next two trials, bowls containing M1 + O2 or M2 + O1 were presented, with M1 + O2 on the right (trial 3) and then on the left (trial 4)). This means that trials 1 and 3 were always the first exposure to each novel combination of exemplars. Meanwhile, for trials 2 and 4, the rats had a second exposure to the same exemplar combinations as the previous trial, but with the rewarded location changed. The first exposure to completely novel exemplars (i.e., SD, ID and ED stages in the standard task) must be 'guessed' therefore the group would be expected to have a 50 % success rate on trial 1. For reversal stages, which are only signalled by the lack of a reward in the previously correct bowl, the group success rate on trial 1 is expected to be 0 %.

Completion of a stage required rats to reach a criterion performance of consecutive correct responses which was greater than that predicted by chance (6 consecutively correct responses, $p = 0.0156$). Correct responses within the first four trials were included in this measure, thus it is possible for rats to complete a stage in just six trials, even while the first trial has only a 50 % chance of being correct.

The first test commenced the day following training. Subsequent tests were completed without the need for further training and were pseudo-randomly counterbalanced for exemplar pairing presentation-order and the dimension (O or M) rewarded. The exemplars used are shown in Table 1.

The 7-stage attentional set-shifting (ID/ED) task

We first tested rats in the standard 7-stage ID/ED task. The seven stages are: a simple discrimination (SD), in which the rat discriminates either between two odors in sawdust, or between two unscented digging media; a compound discrimination (CD) in which the SD exemplars remain relevant, but are paired with irrelevant exemplars from the other dimension; the first reversal (REV1), in which the incorrect exemplar at the SD and CD stages is now correct and vice versa; an intra-dimensional shift (ID), where new compound exemplars are presented and the discrimination is within the same dimension as for the preceding stages; a reversal of the correct and incorrect ID exemplars (REV2); an extra-dimensional (ED) shift, where new compound exemplars are presented and the discrimination is between exemplars in the previously irrelevant dimension; and a final reversal (REV3) of those exemplars. All rats completed all testing stages within 2–3 hrs after treatment with vehicle or CNO.

Task variant with novel redundant exemplars

For the modified version of the task, the first four stages (SD, CD, REV1, ID) followed the same method as the 7-stage task. After reaching criterion at the ID, rather than a reversal, rats were given an additional stage with redundant exemplars (RS): the rewarded exemplar from the ID was still rewarded but novel exemplars were introduced in the irrelevant dimension. The same pair of bowls was used on every trial, however, so that the novel exemplars in what had been the irrelevant

Table 1
Exemplars were presented in pairs of two odours and two media. The training pair were always the same and never used for testing. New exemplars were introduced at the SD/CD; ID; and ED stages. The three other pairings were used in a counter-balanced order between rats and across test sessions. The rewarded exemplar was similarly counterbalanced.

Dimension	Training pairs	Pair 1	Pair 2	Pair 3
Odor	Mint	Cinnamon	Sage	Turmeric
	Oregano	Ginger	Paprika	Cloves
Media	Polystyrene	Coarse tea	Sand	Coarse sawdust
	Shredded paper	Fine tea	Grit	Fine sawdust

dimension were also reliable predictors of reward. For example, if the rat was correctly responding to M3 (and not M4, with O3 and O4 irrelevant), a novel exemplar, O5, would be consistently paired with M3 and another novel exemplar, O6, would be consistently paired with M4. Thus, in this example, the rewarded bowl is always M3 + O5 and the unrewarded bowl is always M4 + O6. As the same exemplars of M continue to be predictive of reward / no reward, the new exemplars of O are providing redundant information. It is well established that there is reduced learning about redundant cues, a phenomenon known as blocking (Kamin, 1968). For the subsequent ED stage, these redundant exemplars became the sole predictors of reward and the exemplars that had been relevant were replaced with novel exemplars that were now irrelevant. A schematic of the modified task stages is shown in Fig. 1.

We predicted that, with a normally functioning mPFC, attention to the irrelevant dimension would be downregulated and there would be no attention to the redundant cues, resembling Kamin-blocking. Thus, rats would not learn about the redundant novel exemplars and the magnitude of the shift-cost at the ED stage would not change. However, with mPFC inhibited by CNO-treatment, a failure to downregulate attention to both irrelevant and redundant information would result in more rapid learning of the ED shift.

Order of testing

Two weeks following surgery, all rats underwent the standard ID/ED task (data not shown) to familiarise them with exemplars and discriminations and to ensure they could complete the task within the time course of CNO-mediated DREADDs activation (Roth, 2016). For all subsequent testing, the rats were tested in a counterbalanced AB design with half of the rats receiving 10 mg/kg CNO and the remaining receiving the vehicle without CNO. For the 7-stage testing, the rats were tested twice, one week apart, with an i.p. injection of 10 mg/kg CNO or saline vehicle. Two weeks later the rats were retested twice (one week apart) in the standard ID/ED task following oral administration of 10 mg/kg CNO suspended in a gelatine gummy or a gummy without CNO. Three weeks later, the rats were all tested twice (one week apart) in the task variant, with the novel redundant exemplars, following oral administration of 10 mg/kg CNO or vehicle.

Histology

Following behavioral testing rats were anaesthetised with 0.8 ml pentobarbital (i.p.; Pharmasol, Ltd, UK), then transcardially perfused with 4 % paraformaldehyde in 0.1 M phosphate buffer. Brains were stored in 20 % sucrose for 24 hrs at 4 °C. Brains were washed in distilled water, dried, placed in wells, covered in egg yolk, and placed in a 40 % formaldehyde bath for 72 hrs. Afterwards, brains were sectioned at 50 μm on a freezing stage microtome (Jung Histoslide 2000, Reichert-Jung, Cambridge Instruments). Sections were stored in glycerol solution at − 20° C.

For immunofluorescent detection, every 4th section of the frontal cortex was collected. Sections were placed in 9-hole netwells and petri dishes and washed four times for 5 min each in 0.1 M phosphate buffered saline (PBS) on an automated rotator. Next the netwells were placed in blocking solution (1:5 normal goat serum, 1:100 10 % Triton, in PBS) and rotated for 1 hr at room temperature. Sections were then washed in PBS three times for 5 min each. Sections were placed in histology pots before being incubated with 5 ml of anti-mCherry (rabbit anti-mCherry 1:2000; Abcam Cat# ab167453 RRID:AB_2571870) in antibody diluting solution (ADS; 1:100 normal goat serum, 1:100 10 % triton, in PBS) overnight at room temperature. The following day sections were washed in PBS three times for 5 min each. Sections were then switched into foil-covered histology pots and incubated in the dark for 1 hr with the secondary (1:500 goat anti-rabbit, Abcam, #ab150084) in ADS. After secondary incubation, sections were washed three times for 5 min in PBS, then mounted to slides. Vectashield anti-fade mounting medium with DAPI (Vector Laboratories Cat# H-1200 RRID:AB_2336790) was applied, then slides were cover-slipped and sealed.

Fos expression in mPFC

To examine the effect of iDREADDs + CNO on expression of Fos protein in mPFC, six rats expressing AAV5-CamKII-hM4Di-mCherry within the mPFC, and six surgical controls were habituated to i.p. injections of 2 ml/kg saline, for two days prior to testing. We have previously reported no difference in Fos expression when either lesioned or control rats were performing the ID/ED task compared to their 'yoked' controls (Tait et al., 2009). Therefore, we used a modified "active-wake" condition (Gompf et al., 2010). Non-cage mate pairs of a control or iDREADDs-transfected rats were treated with i.p. injections of either 10 mg/kg CNO or vehicle (administered at 5 mg/ml CNO in 2 ml/kg saline). Thirty minutes following injections the rats were placed into a large wooden enclosure (66 cm × 66 cm × 40 cm) filled with sawdust, and environmental enrichment (cardboard houses, wooden chew-bars, and shredded cardboard) and allowed to explore for 2 hrs. Immediately, at the end of the exploration period, the rats were removed from the apparatus, perfused, and the brains were stored for immunohistochemistry. For quantification of Fos, tissue sections were prepared as above and incubated in rabbit anti-c-Fos (1:10000; Calbiochem, San Diego, USA), in ADS overnight at room temperature. Sections were washed three times for 5 min in PBS then incubated in 5 ml biotinylated

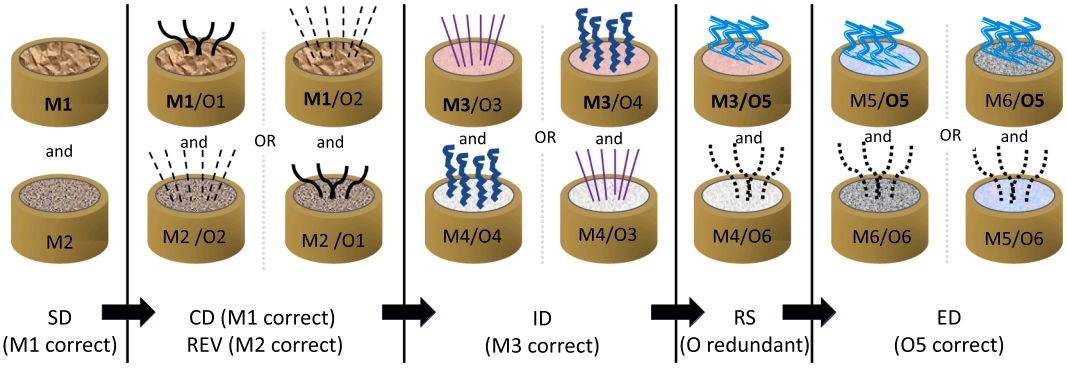

Fig. 1. Schematic showing the stages (in order of testing from left to right) of the modified set-shifting task, with an example of a medium (M) to odor (O) shift. The initial stages are the same as the standard ID/ED task: namely, simple discrimination (SD); a compound discrimination (CD); reversal (REV1); and an intradimensional discrimination (ID). After the ID, there was a stage in which the correct (M3) and incorrect (M4) exemplars did not change but the irrelevant exemplars (O3 and O4) were replaced with novel exemplars (O5 and O6). In addition, a single pair of bowls was used (M3/O5 and M4/O6), such that the novel exemplars were also fully predictive of reward but provided redundant information. In the final extradimensional shift (ED), new exemplars (M5 and M6) were introduced as irrelevant. The previously redundant odours (O5 and O6) were now the only predictors of reward/non-reward.

anti-rabbit antibodies (goat anti-rabbit, 1:500, Vectorlabs, in ADS) for 1 hr at room temperature. After incubation, sections were washed three times for 5 min in PBS, then placed in histology pots containing 5 ml of the biotinylation solution (Vectastain ABC KIT, Vectorlabs) and incubated for 1 hr. Sections were then washed three times for 5 min in PBS and Fos protein immunoreactivity was detected by staining with 3,3'-Diaminobenzidine (DAB; one tablet per 20 ml, Sigma, in distilled H_20). The sections were determined to be stained when landmark anatomical structures were clearly identifiable, which was always within 10 min. Following DAB staining, sections were washed three times for 5 min in PBS and stored in 9-hole netwells at 4 °C until (up to 72 hrs) they were mounted to gelatine-treated glass slides and cover-slipped with DPX. Fos positive cells in six discrete mPFC sections between 4.7–2.2 mm were imaged on a Zeiss Axio Imager M2 with ZEN software. Fos immunostaining was quantified from grayscale transformed images using an ImageJ particle analyser. The quantification was done blind to condition.

Statistical analysis

The standard dependant variable collected in the ID/ED task is the number of trials to reach the learning criterion of six consecutively correct trials. We have previously shown (Dhawan et al., 2019) that after a rat has made six consecutively correct trials, most rats make no subsequent errors if tested for 30 additional trials.

We also applied Bayes' rule to the profile of the sequence of response choices to estimate, for a given trial, the posterior probability that the response on that trial was consistent with one of eight hypothetical patterns. Four of these were perceptual patterns (i.e., a response to any one of the four exemplars: the rewarded exemplar; the unrewarded exemplar; or either of the two exemplars in the irrelevant dimension) and the other four were spatial patterns (i.e., win-stay; win-shift; perseverate to location; alternate locations). At the start of the SD stage, the prior probabilities were set to 1/6 (0.167) as there was no irrelevant dimension, so only six potential patterns. At the CD, the priors from the final trial of the SD were carried forward but adjusted proportionally to assign priors to the new irrelevant dimension exemplars. All priors were reset to one 8th (0.125) at the ID and ED stage. At the reversal stages, the priors were not reset. See Wang et al. (2019) for a full description of the methodology. We refer to these values as 'b-values'.

Multi-factorial repeated measures ANOVA (SPSS v28) were performed on the dependent variables (e.g., trials to criterion; b-values) with the relevant independent variables as the repeated factors (e.g., Stage; Trial; Treatment). When the repeated measures ANOVA revealed interactions that were significant ($p < 0.05$), planned contrasts or analysis of simple main effects were performed.

For Fos data, a count of labelled cells within a region of interest of 1 mm^2 in each of six discrete sections (Bregma + 4.70, +4.20, +3.70, +3.20, +2.70, +2.20). The mean count from each section was analysed using between-Ss measures ANOVA with factors being Treatment (Vehicle or CNO) and Group (controls or mPFC iDREADDs). Significant interactions were further analysed as simple main effects. All rats

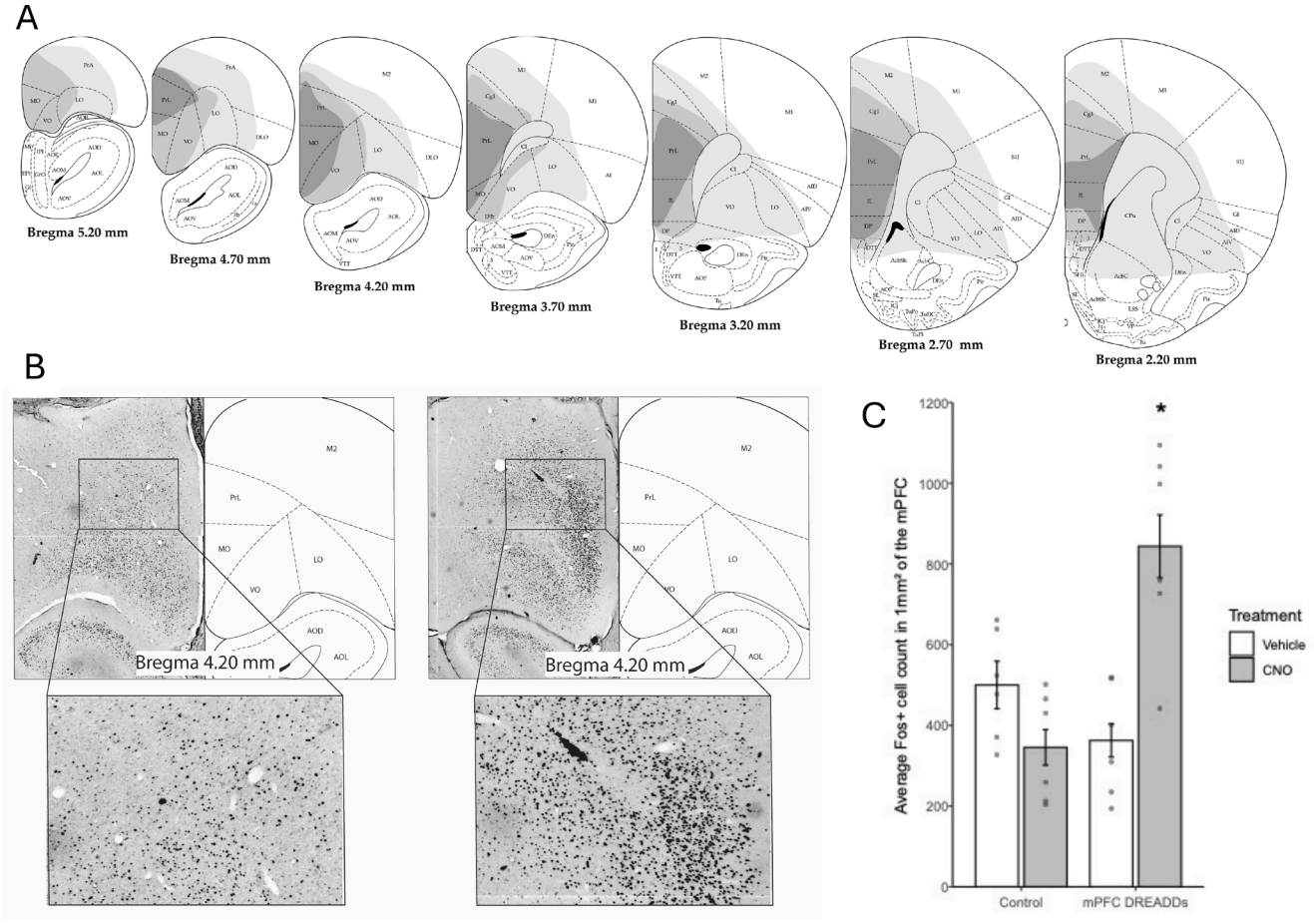

Fig. 2. A: Schematic representation of the area of mPFC neurons infected by microinjection of AAV5-CamKII-hM4Di-mCherry as detected by anti-mCherry immunohistochemistry drawn on sections from Paxinos and Watson (1997). Dark gray shading indicates the area common to all rats; mid-gray shading is the typical extent; light gray shading is the maximum extent in any rat. B: Representative images of c-Fos immunostaining within the mPFC; C: Box-and-whisker plots of average c-Fos counts in the six sections. iDREADDs + CNO was associated with increased c-Fos immunostaining compared to iDREADDs + vehicle; controls + CNO; or controls + vehicle (* $p < 0.05$).

completed all the behavioral tests, and it was not necessary to exclude any from the analysis.

Results

All rats exhibited iDREADDs expression within mPFC

Fig. 2a shows the extent of area with mPFC neurons transfected with AAV5-CamKII-hM4Di-mCherry and Fig. 2b is a representative fluorescent image. All rats (n = 18) showed mCherry-tagged iDREADDs expression within the mPFC target region (approximate distance from bregma 4.2–2.2 mm), centred on the prelimbic (PrL) and infralimbic (IL) cortices (Fig. 2a, dark gray shading). Additional DREADDs expression was evident in the medial orbitofrontal (MO) and cingulate (Cg) cortex in most rats (Fig. 2a, mid-gray shading). Light gray shading in Fig. 2a indicates the maximum extent of any transfected cells. DREADDs expression was also observable within axons consistent with projections from the mPFC (Vertes, 2004).

CNO-treatment increased Fos activity within the mPFC

As shown in Fig. 2b and 2c, compared to controls and vehicle-treated animals, treatment with 10 mg/kg CNO caused a significant increase in Fos expression within the mPFC. This effect was confirmed by a significant Treatment × Group interaction ($F(1, 8) = 15.03$, $p < 0.05$, $\eta_p^2 = 0.65$). Simple main effects analysis confirmed that average Fos expression was higher when rats with mPFC iDREADDS were treated with CNO (mPFC-DREADDs + CNO=843.89 (95 % CI±71.05)) rather than vehicle (mPFC-DREADDs + vehicle = 362.72 (95 % CI±44.88)) ($F(1, 8) = 18.81$, $p < 0.05$, $\eta_p^2 = 0.70$). Also, Fos expression was higher following CNO treatment in rats with mPFC iDREADDs than those without (control + CNO=345.67 (95 % CI±42.71); $F(1, 8) = 19.03$, $p < 0.05$, $\eta_p^2 = 0.70$).

mPFC inactivation resulted in an ED set-shifting deficit

In a within Ss design, we tested the hypotheses that (1) CNO administration, to inactivate mPFC would impair performance at the ED stage of the standard ID/ED task, just as do cell body lesions of mPFC and (2) that oral administration of CNO would produce the same profile of behavior as i.p. injection.

The mean trials to criterion across all stages was overall slightly higher in the first two tests, with i.p. injections (mean across all stages = 17.18, 95 % CI±1.3) compared to the second two tests, with oral administration (mean across all stages = 15.72, 95 % CI±0.9) (main effect of Test: $F(1, 11) = 16.7$, $p < 0.05$, $\eta_p^2 = 0.605$). However, the effect was equivalent for all stages of the test and irrespective of whether CNO or the vehicle was administered (all interactions with Test were not significant and, in particular, Test × Treatment × Stage: $F(6, 66) < 1$). We concluded that oral administration by voluntary ingestion of a gummy was an effective method to administer CNO and therefore was preferred over i.p. injection. Even if there is a different rate of metabolism as a function of route of administration, the results indicate that the testing was within the 'window' of effect for each route and any difference in metabolic rate is thus not relevant with these doses and in the context of this behavioral protocol. Subsequent testing used only oral administration.

The rats formed an attentional set, indicated by the fact that, regardless of whether vehicle- or CNO-treated, they required more trials to reach criterion at the ED shift compared to the ID (planned pairwise comparison ID vs ED: $F(1, 11) = 297.5$, $p < 0.05$, $\eta_p^2 = 0.964$; control mean ED increase = 4.33; CNO mean ED increase = 13.9, both $p < 0.05$). However, when CNO-treated, the rats required significantly more trials to complete the ED stage than when they were vehicle-treated (ED trials to criterion when CNO-treated = 26.04 (95 % CI±2.82); when vehicle-treated = 15.21 (95 % CI±1.46). See Fig. 3).

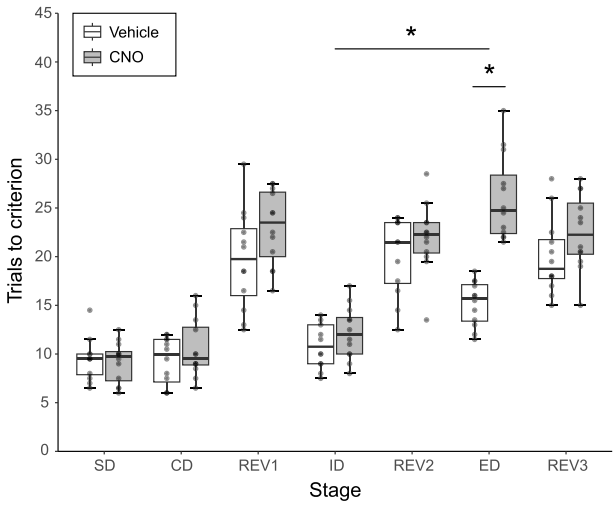

Fig. 3. Box-and-whisker plots (median, 5th, 25th, 75th and 95th percentiles, with individual rat data points) for trials to criterion in the standard 7-stage ID/ED task. Rats required a greater number of trials to learn a novel ED shift, compared to an ID shift. With the iDREADDs activated by CNO, rats required significantly more trials at the ED stage, compared to when vehicle treated (* $p < 0.05$).

In terms of patterns of consistent responding during the ED stage (defined as when the b-values from the Bayesian analysis were ≥ 0.6, when CNO-treated, rats were more likely to respond to one of the exemplars in the irrelevant dimension (that had previously been relevant) compared to when vehicle-treated (Treatment × Task interaction: $F(1,11) = 7.35$, $p = 0.02$, $\eta_p^2 = 0.401$; see Fig. 4).

In the trials to criterion data, there were no significant differences between the CNO-treated and vehicle conditions at any other stage of the test, replicating the effect of cell-body lesions of the same region.

There was no effect of mPFC inactivation in the stages prior to ED

We looked for patterns of consistent responding in the four stages that were equivalent in each task, namely, SD, CD, REV1 and ID (i.e., those prior to REV2 in the 7-stage task or RS in the task variant with novel redundant exemplars). For the 7-stage task, we used the data from the tests with oral administration of CNO or vehicle. This gave a total of four repeats, two with and two without CNO-treatment.

Before rats started responding consistently to the correct exemplar, any consistent behavioral patterns tended to be spatial: all rats in each of these four initial stages had some trials in which patterns of either spatial perseveration or spatial alternation predominated (b-values ≥ 0.6), with neither pattern more likely than the other. There were no differences in the likelihood of spatial response patterns between Treatment groups for these stages of the test (main effect of Treatment: $F(1,11) = 0.16$, ns; interaction of Treatment and Stage: $F(2,22) = 0.17$, ns). Although we observed patterns of spatial responding, this is not to suggest the rat is 'testing the hypothesis' that the reward is determined by spatial location. Rather, the rat cannot approach both bowls simultaneously and does not know which is the correct bowl until it digs to find out. Therefore, a pattern of spatial perseveration or alternation may merely indicate a systematic choice of where to start its exploration.

A pattern of consistent responding to exemplars in the irrelevant dimension was detected in only one (vehicle-treated) rat at the CD and in three (two vehicle-treated; one CNO-treated) at the ID. Irrelevant dimension responding was slightly more likely at REV1, but it was still not seen in all rats and this was without respect to treatment (seven and five rats when vehicle and CNO-treated, respectively).

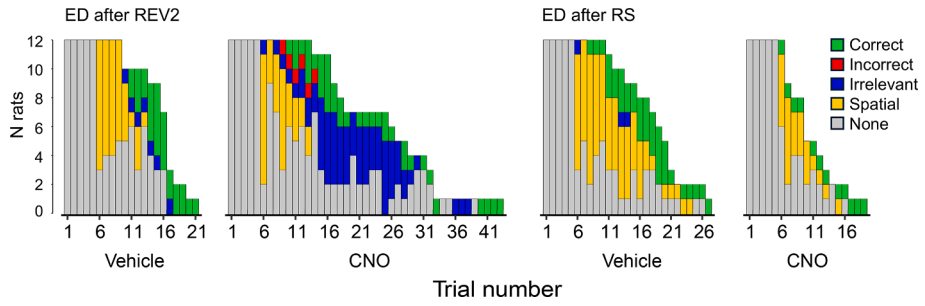

Fig. 4. Histograms showing the number of rats with different response patterns (where the b-value is > 0.6) across trials. The four histograms are for the ED stages, either following a reversal (left two histograms) or following exposure to redundant stimuli (RS) (right two histograms). For the first five trials, there is insufficient evidence to identify a consistent pattern of responding. If a pattern does emerge, it is initially most likely to be one of the spatial patterns (yellow). The histogram drops away as rats reach criterion (six consecutive correct responses) and at this point there is often (but not necessarily) sufficient evidence of responding consistent with the correct hypothesis (green) – and inconsistent with other hypotheses – to bring the b-value above 0.6. The obvious difference between the histograms is the evidence of response patterns consistent with responding to one or other of the irrelevant dimension exemplars (blue), but only when CNO-treated and when the ED follows a reversal (REV). This pattern is not seen when CNO-treated if the ED follows RS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Inactivation of mPFC facilitated an attentional shift when preceded by redundant exemplars in the to-be-shifted-to dimension

The modified ID/ED task had an additional stage (RS) that provided reward-contingent but redundant information in what had been the irrelevant dimension, before the ED. Most rats (8/12) rats made no errors in the RS stage, either when CNO– or vehicle-treated: they completed the stage both times in six trials. Four rats did make errors in one or both treatment conditions. When CNO treated, two rats each made two errors (trials 1 and 2) and a third made four errors (trials 1, 4, 5 and 6). When vehicle treated, two of the same rats made errors, on trials 1 and 2 and trials 2 and 3 respectively. A fourth rat made a single error, on trial 1, when tested with vehicle. These errors did not relate to any differences in the trials to criterion on the preceding ID or the subsequent ED.

In the subsequent ED stage, when vehicle-treated, rats showed the expected increase in trials to criterion (planned contrast, ID vs ED for vehicle treated: mean ED increase = 7.92, $p < 0.05$). There was no statistically significant difference between this ED performance compared to when the ED followed a reversal (7-stage task) (mean difference = 2.92, 95 % CI = ±3.7; t = 1.74, df = 11, ns).

When CNO-treated, there was not only no statistically significant increase in ED over ID, but there was also no numerical increase (mean ED increase = − 0.17, ns). There was also no difference in the trials to criterion between Treatment condition at any stage other than the ED, where the CNO treatment decreased the ED trials relative to vehicle treatment, and to the equivalent of ID performance with or without CNO (see Fig. 5; Treatment × Stage interaction $F(5, 28.1) = 3.12$, $p < 0.05$, $\eta_p^2 = 0.62$). As can be seen in Fig. 4, CNO-treated rats did not respond to exemplars in the irrelevant dimension.

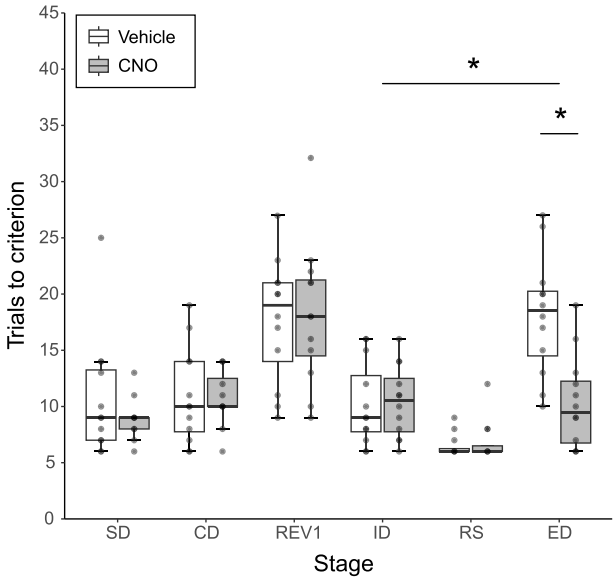

Fig. 5. Box-and-whisker plots (median, 5th, 25th, 75th and 95th percentiles, with individual rat data points) for trials to criterion in the modified task where the second reversal is replaced with a stage with redundant, but fully predictive, exemplars (RS). These redundant exemplars became the relevant exemplars at the subsequent ED stage. When vehicle-treated, the rats showed a typical ED shift cost. However, when CNO-treated, the subsequent ED was learned in significantly fewer trials compared to vehicle (* $p < 0.05$) and the shift cost was abolished.

Evidence to suggest more rapid learning of blocked cues

It was expected that, when treated with vehicle, rats exposed to novel redundant exemplars would exhibit no learning, due to blocking (Kamin, 1968), and therefore there would be no effect on ED performance. If there had been any learning about the new exemplars, rats would be more likely than chance to make a correct response on the first trial. Therefore, we looked at the binomial probability that the rats responded correctly greater than by chance on the trials they encountered the novel combination of exemplars for the first (trials 1 and 3) and the second encounter of each bowl (trials 2 and 4). Fig. 6 shows the percentage of the group responding correctly when the ED followed a reversal or when it followed exposure to redundant exemplars. When the ED followed a reversal, all the exemplars were novel, and the rats' correct responding was no greater than chance on either the first encounter or the second, regardless of treatment condition. In the modified task, after prior exposure to the now-correct exemplars as fully predictive but redundant, when CNO– or vehicle-treated, rats were still no more likely to respond correctly on the first encounter with each novel combination of exemplars, indicating that the rats had not learned specifically about the predictive redundant exemplars. We can rule out the possibility that the novel irrelevant exemplars were a distraction on this first trial: in both the CD and the RS stages, novel irrelevant exemplars were introduced while the correct exemplar was unchanged and, regardless of treatment condition, the percentage of correct responses on the first trial of the CD and RS stages was > 75 %, which greater than chance.

However, on the second encounter of the bowls with exemplars having been redundant, the CNO-treatment resulted in significantly greater likelihood of a correct response, with 10/12 (83 %) rats correct

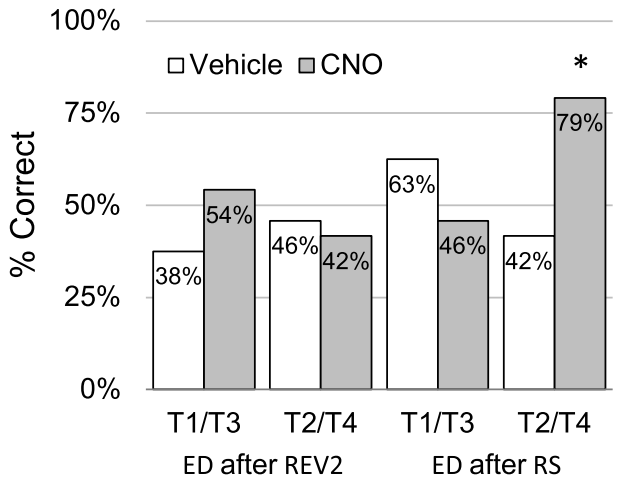

Fig. 6. Bar graph showing the percentage of responses that were correct when a novel combination of exemplars was encountered for the first time (Trials 1 and 3) or the second time (Trials 2 and 4) in the ED. The rats responded at chance when encountering novel bowls for the first time, as expected. They also responded at chance on the second encounter with the bowls in the ED of the 7-stage task, regardless of CNO treatment. However, when they were encountering bowls for the second time having experienced the now correct exemplars as redundant, when CNO-treated they were more likely than chance to respond correctly (* $p < 0.05$).

on trial 2 and 9/12 (75 %) correct on trial 4. Overall, 79 % of trial 2 and 4 responses were correct ($p < 0.006$; the corrected criterion p-value for 8 tests and overall familywise error of 0.05 for a one-tailed binomial test with H_0 p(correct) = 0.5 and H_A p(correct) > 0.5.) This pattern suggests that rather than a failure of blocking, the improved ED performance is because learning about the redundant cues is more rapid.

The behavioral data are available online (Knott et al., 2024).

Discussion

It is well established, in many species, that additional trials are required to learn a discrimination if the focus of attention is elsewhere due to prior experience of the relevance of information (Keeler and Robbins, 2011). The construct of cognitive flexibility is invoked to describe the facility with which an organism shifts attention to learn about newly relevant information. If new learning occurs more slowly under these circumstances, such as is seen following inactivation of the mPFC, a deficit in cognitive flexibility is inferred. Here, we have replicated the mPFC lesion-induced deficit in learning at the ED stage with the temporary inactivation of DREADDs-transfected cells in the mPFC. We have additionally demonstrated that learning at the ED stage may be more rapid if exemplars are presented prior to the shift as redundant information. We conclude that both the impaired and the facilitation of learning at the ED shift is due to a common deficit in the downregulation of attention.

Did the iDREADDs + CNO inactivate mPFC?

Before discussing the behavioral results of the experiment in detail, it is first necessary to discuss whether iDREADDs + CNO inactivated mPFC and whether this inactivation, as opposed to other non-specific off-target effects of CNO, can account for the behavioral results we report here.

In all cases, bilateral expression of iDREADDs was confirmed histologically. The volume injected (2ul/hemisphere), which was large but not unprecedented (e.g., Robinson et al. (2019) used a comparable volume), and location (two sites per side) resulted, as intended, in the expression of iDREADDs in a large area of mPFC, encompassing the majority of PrL in all rats and extending to MO and IL in most.

CNO is metabolized to clozapine, which binds to the DREADD but clozapine also has effects that are not mediated via the DREADDs (MacLaren et al., 2016; Gomez et al., 2017; Manvich et al., 2018). Ferrari et al. (2022) directly compared CNO with Compound 21 and deschloroclozapine as alternative DREADDs ligands and found the effects of all three were equivalent. They suggested, because they are not metabolised to clozapine, the alternatives are preferrable to CNO. Nevertheless, the similarity of all three suggests that the effects are more likely to be via their action on the DREADDs.

In the behavioral experiments, we did not include a control group administered CNO without iDREADDs. However, we have previously compared CNO and clozapine directly in this task in rats without DREADDs: there was no effect of either 1 mg/kg clozapine or of 10 mg/kg CNO (giving an equivalent brain-availability of clozapine (Gomez et al., 2017)) at any specific stage of the task (data available on request). Although it is not possible to definitively conclude that there were no off-target actions of CNO, the fact that mPFC-iDREADDs + CNO mimics the effect of a cell-body lesion of mPFC, resulting in precisely the same pattern of behavior, is strong circumstantial evidence that the effects of CNO that we observe are due to the inactivation of DREADDs transfected cells in mPFC.

In the Fos experiment, we did include a non-iDREADDs + CNO control group. Only the rats with iDREADDs + CNO had increased Fos: control-vehicle, control + CNO and iDREADDs + vehicle all had equivalent Fos expression. This suggests that it is the inactivation of the cells expressing the iDREADDs by CNO that results in the increased Fos expression in other cells and it is not due to a direct effect of CNO/clozapine on non-transfected cells. The elevation in Fos might seem paradoxical: iDREADDs + CNO *must* reduce neuronal activity (and therefore Fos expression) where the iDREADDs are expressed. However, this is only necessarily so if all the cells in a region are transfected, which is known not to be the case (Smith et al., 2016). Chang et al. (2015) recorded from the rat ventral pallidum that had been transfected with iDREADDs and, while most of the cells were inhibited by CNO, there was a proportion that were excited. An increase in Fos expression was also reported in the mPFC of rats with mPFC lesions and, furthermore, the Fos expression was greater still when the lesioned rats were given Asenapine (Tait et al., 2009). Obviously, there is no mechanism by which 'silenced' cells, whether due to cell-body lesion or by iDREADDs + CNO, can themselves express Fos. Tait et al. (2009) interpreted the increased Fos expression in the mPFC of mPFC-lesioned rats as compensatory activity of spared cells. Here we have shown that not only the behavioral effects of mPFC lesions are replicated by iDREADDs + CNO, the increase in mPFC Fos expression is also replicated. Both the lesions and the iDREADDs transfection were confirmed by histology. We acknowledge that saying that mPFC is 'inactivated' by iDREADDs + CNO, or indeed by a cell-body lesion, even while c-Fos activity is increased in the region, may seem contradictory. What the increased c-Fos activity signifies is unclear and beyond the scope of this manuscript to establish. Nevertheless, we conclude that the silencing of some mPFC cells must cause an increase in activity of, and Fos expression in, other non-lesioned/non-transfected cells. Clearly, in this brain area, the level of Fos expression cannot be taken as indicative of either the presence of a lesion or cellular inactivation by iDREADDs + CNO.

Did the task manipulation improve cognitive flexibility?

Barch et al. (2009) suggested that there would be value in identifying a task modification in the ID/ED attentional set-shifting task as an additional control, giving as an example "*one in which set shifting deficits convey a performance advantage*" (*ibid.*, p 120). We have shown here that the same intervention that impaired ED shifting leads to a performance advantage at the ED stage if predictive, redundant, information is provided prior to the shift. We are not suggesting that the task manipulation (i.e., the introduction of the predictive, redundant, exemplars prior to the ED shift stage) has in some sense improved flexibility. Rather, the

attentional impairment that results in a typical set-shifting deficit conveys a performance advantage at the ED stage under these circumstances. This provides additional information about how both intact and impaired animals solve these tasks.

At the ID and ED stages of the 7-stage task, the total change of exemplars is a signal of a requirement for new learning, although the change does not signal whether the rule has changed. The task manipulation involved a new stage, the RS, where there was a partial change of exemplars, albeit not signalling a requirement for new learning. At the ED stage, there was also a partial, not total, change of exemplars: the redundant exemplars from the RS stage remain and become relevant and only the previously rewarded / unrewarded exemplars change. Arguably, a partial change of exemplars is less salient (i.e., it will generate less PE) than a total change. However, if that were the case, one would predict slower learning of the partial change ED shift, which we did not see: with mPFC functioning normally, there was no significant difference in learning rate of the two ED stages, suggesting that PE driving new learning was equivalent in both conditions.

In ID/ED tests with healthy humans, the shift-cost (i.e., number of errors at the ED stage) has been reported to be greater in a condition where the opportunity to perseverate is removed and so only learned irrelevance can slow the shift, compared to a condition when perseveration is possible but learned irrelevance cannot impact performance (Gauntlett-Gilbert et al., 1999; Maes et al., 2004; Maes and Eling, 2009). Similarly, in both healthy humans (as well as in patients with Parkinson's disease) when an irrelevant dimension was made partially relevant, the ED shift-cost was reduced (Slabosz et al., 2006). Together, these results are consistent with the suggestions that the ED shift-cost in healthy controls is due to a reduction in attention to irrelevant non-predictive cues, which has been termed learned inattention (Kruschke and Blair, 2000; Le Pelley and McLaren, 2003) or reduced associability (Le Pelley, 2004). If learned inattention is the reason there is a shift cost in an intact subject, it might be surmised that an increase in shift-cost is most easily explained by postulating increase in learned inattention. As Maes and Eling (2009) speculated, if a subject is not ignoring irrelevant cues, one would predict a reduction, rather than an increase, in shift-costs when those cues become relevant. On the other hand, if associability (salience) of the irrelevant cues remains high, but associative strength is reduced, a failure to ignore those cues might enhance learning about their irrelevance, which in turn would increase perseverative errors and thus shift-cost. Indeed, Castro and Wasserman (2016) argue that, in their experiment, attending to relevant cues "*protected the irrelevant cues from learned irrelevance*" (*ibid.*, p.71). They cite the earlier work of Winefield (1978), who showed that irrelevant brightness cues that were present during learning of a solvable spatial discrimination were more readily learned about subsequently than when those cues were present during an unsolvable discrimination. The suggestion is that, because the discrimination was unsolvable, attention to the cues was maintained and thus removed the protection from learned irrelevance (Castro and Wasserman, 2016). Similarly, (Maes and Eling, 2009) interpret the correlation between errors during initial learning and ED shift performance as being due to more rapid acquisition (i.e., fewer learning trials) protecting participants from the opportunity to learn about the irrelevance of other cues.

It follows from this that the attentional set-shifting effect should only be seen when there are irrelevant cues during initial learning, and there is good evidence that this is the case. For example, an impairment following mPFC lesions in shifting from an "easy" (more rapidly learned) visual cue discrimination to a "more difficult" (more slowly learned) response side discrimination was not seen when the rats were required to shift in the opposite direction, from the response side to the visual discrimination (Experiment 1; Floresco et al. (2008)). Crucially, however, the training regime did not introduce the lights until after the response discrimination had been learned and when the rats were required to shift. Thus, unlike the 'spatial dimension,' which was always (necessarily) present, the visual cues had never been present and irrelevant. However, when the cues were presented as irrelevant stimuli throughout pre-training and during initial response training (Experiment 2; Floresco et al. (2008)), mPFC inactivation did impair the shift to visual cue discrimination. These results together demonstrate that mPFC is not required to shift attention between different cues, or rules or strategies, *per se*, but rather to learn about cues, or rules or strategies, that had previously been learned to be irrelevant.

In the introduction, we highlighted the apparent contradiction in the schizophrenia literature between reports of impaired shifting in the ID/ED task and the suggestion that hyperactive, rather than impaired, switching might underlie impairments in latent inhibition (Weiner, 1990; Weiner and Feldon, 1997). The results here suggest that the parsimonious explanation for both phenomenon in schizophrenia may be a deficit in learned inattention (Mackintosh, 1975; Lubow, 1997; Gray and Snowden, 2005). In the rat, latent inhibition is enhanced following large lesions of a comparable area of mPFC, however smaller lesions revealed that it was ventral (IL), and not dorsal (PrL), mPFC that was implicated (George et al., 2010). This does not exclude the possibility that dorsal mPFC may also be involved if latent inhibition were measured using appetitive reinforcement rather than conditioned suppression of responding.

Kamin (1968) wrote about "attention-like" processes in classical conditioning, suggesting that there is not merely reduced learning (decreased associative strength) of a blocked cue, but attention to the cue is also attenuated (decreased associability or salience). Here we argue that "blocking-like" processes account for shift-costs in the attentional set-shifting task. Irrelevant cues have low associative strength because they are non-predictive and their associability, or salience, correspondingly reduces. Modulation of salience is a function of the mPFC, regulating attention to focus on informative (i.e., contingent and thereby predictive) cues over non-informative (i.e., non-contingent, or contingent but redundant) cues. For this reason, a rat with an intact mPFC learns an ED discrimination more slowly because attention is focussed on that which had been informative, and attention is downregulated to previously uninformative cues. This "blocking-like" effect (i.e., learned inattention to non-informative cues) prevents excessive learning about their irrelevancy. The same "blocking-like" effect is evident when cues become redundant predictors of reward: the intact, non-attending, rat does not learn about these cues and the ED shift-cost is still determined by the same requirement to overcome the learned inattention. By contrast, a rat with an inhibited mPFC continues to attend to cues with low associative strength (e.g., irrelevant exemplars), resulting in additional reinforcement of their irrelevancy. This makes an ED discrimination yet more slowly learned compared to when the mPFC is intact. However, when the cues cease to be irrelevant and become redundant predictors of reward, their maintained saliency means the rat learns that they may not be irrelevant, and this facilitates learning of the subsequent ED. Notably, while the cues are redundant, the prior learning is still valid and so there is no PE to drive learning about the specific redundant exemplars: responses to the first encounter of each novel combination of exemplars (trials 1 and 3) are at chance. However, because the cues are predictive, they are known to be not irrelevant and consequently subsequent learning is very rapid: responses to the second encounter of each bowl configuration (trials 2 and 4) are significantly more likely to be correct.

The manipulation we have reported here satisfies the challenge suggested by Beesley and Le Pelley (2011), namely that if there were a manipulation that impaired the ability to selectively attend to stimuli, and that same manipulation reduces the influence of blocking on novel learning, this would provide support for the suggestion that attention mediates the blocking effect. It also satisfies the challenge suggested by Barch et al. (2009), being a modification in the ID/ED attentional set-shifting task which conveys a performance advantage.

CRediT authorship contribution statement

Tegan S. Knott: Writing – review & editing, Visualization, Software, Investigation, Formal analysis. **Alonzo J. Whyte:** Writing – review & editing, Writing – original draft, Investigation, Formal analysis. **Sandeep S. Dhawan:** Writing – review & editing, Investigation. **David S. Tait:** Writing – review & editing, Supervision, Resources, Project administration, Data curation, Conceptualization. **Verity J. Brown:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the technical staff of the University of St Andrews' St Mary's Animal Unit for their expert care of the animals used for this research. Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Alonzo Whyte was supported by a 600th Anniversary University of St Andrews Doctoral Scholarship.

References

Barch, D.M., Braver, T.S., Carter, C.S., Poldrack, R.A., Robbins, T.W., 2009. CNTRICS final task selection: Executive control. Schizophr. Bull. 35, 115–135.
Beesley, T., Le Pelley, M.E., 2011. The Influence of Blocking on Overt Attention and Associability in Human Learning. J. Exp. Psychol. Anim. Behav. Process. 37, 114–120.
Brady, A.M., Floresco, S.B., 2015. Operant Procedures for Assessing Behavioral Flexibility in Rats. Available at: J. Visualized Exp. 1–13 http://www.jove.com/video/52387/operant-procedures-for-assessing-behavioral-flexibility-in-rats.
Brown, V.J., Tait, D.S., 2015. Behavioral Flexibility: Attentional Shifting, Rule Switching, and Response Reversal Stolerman IP, Price LH, eds. Encyclopedia Psychopharmacol. 264–269. Available at: 10.1007/978-3-642-36172-2_340.
Brown, V.J., Bowman, E.M., 2002. Rodent models of prefrontal cortical function. Trends Neurosci. 25.
Castro, L., Wasserman, E.A., 2016. Attentional shifts in categorization learning: Perseveration but not learned irrelevance. Behav. Process. 123, 63–73.
Ceaser, A.E., Goldberg, T.E., Egan, M.F., McMahon, R.P., Weinberger, D.R., Gold, J.M., 2008. Set-Shifting Ability and Schizophrenia: A Marker of Clinical Illness or an Intermediate Phenotype? Biol. Psychiatry 64, 782–788.
Chang, S.E., Todd, T.P., Bucci, D.J., Smith, K.S., 2015. Chemogenetic manipulation of ventral pallidal neurons impairs acquisition of sign-tracking in rats. Eur. J. Neurosci. 42, 3105–3116.
Chu, R.S.T., Ng, C.M., Chan, K.N., Chan, K.W., Lee, H.M., Hui, L.M., Chen, E., Chang, W. C., 2021. Aberrant learned irrelevance in patients with first-episode schizophrenia-spectrum disorder. Brain Sci. 11, 1370.
Chudasama, Y., Robbins, T.W., 2006. Functions of frontostriatal systems in cognition: Comparative neuropsychopharmacological studies in rats, monkeys and humans. Biol. Psychol. 73, 19–38.
Den Ouden, H.E.M., Kok, P., de Lange, F.P., 2012. How prediction errors shape perception, attention, and motivation. Front. Psychol. 3, 36523.
Dhawan, S.S., Xia, S., Tait, D.S., Bundgaard, C., Bowman, E., Brown, V.J., 2018. Oral dosing of rodents using a palatable tablet. Psychopharmacology (Berl) 235, 1527–1532.
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. Behav. Brain Res. 363, 45–52. Available at: 10.1016/j.bbr.2019.01.055.
du Sert NP et al., 2020. The arrive guidelines 2.0: Updated guidelines for reporting animal research. PLoS Biol. 18.
Elliott, R., McKenna, P.J., Robbins, T.W., Sahakian, B.J., 1995. Neuropsychological evidence for frontostriatal dysfunction in schizophrenia. Available at: Psychol. Med. 25, 619–630 http://journals.cambridge.org/abstract_S0033291700033523.
Ferrari, L.L., Ogbeide-Latario, O.E., Gompf, H.S., Anaclet, C., 2022. Validation of DREADD agonists and administration route in a murine model of sleep enhancement. J. Neurosci. Methods 380.
Floresco, S.B., Block, A.E., Tse, M.T.L., 2008. Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. Behav. Brain Res. 190, 85–96.
Gauntlett-Gilbert, J., Roberts, R.C., Brown, V.J., 1999. Mechanisms underlying attentional set-shifting in Parkinson's disease. Neuropsychologia 37, 605–616.
George, D.N., Duffaud, A.M., Pothuizen, H.H.J., Haddon, J.E., Killcross, S., 2010. Lesions to the ventral, but not the dorsal, medial prefrontal cortex enhance latent inhibition. Eur. J. Neurosci. 31, 1474–1482.
Ghods-Sharifi, S., Haluk, D.M., Floresco, S.B., 2008. Differential effects of inactivation of the orbitofrontal cortex on strategy set-shifting and reversal learning. Neurobiol. Learn. Mem. 89, 567–573.
Gilmour, G., Arguello, A., Bari, A., Brown, V.J., Carter, C., Floresco, S.B., Jentsch, D.J., Tait, D.S., Young, J.W., Robbins, T.W., 2013. Measuring the construct of executive control in schizophrenia: Defining and validating translational animal paradigms for discovery research. Neurosci. Biobehav. Rev. 37, 2125–2140. Available at: 10.1016/j.neubiorev.2012.04.006.
Gomez, J.L., Bonaventura, J., Lesniak, W., Mathews, W.B., Sysa-Shah, P., Rodriguez, L. A., Ellis, R.J., Richie, C.T., Harvey, B.K., Dannals, R.F., Pomper, M.G., Bonci, A., 2017. Michaelides M (2017) Chemogenetics revealed: DREADD occupancy and activation via converted clozapine. Available at: Science (1979) 357, 503–507 https://www.science.org/doi/abs/10.1126/science.aan2475.
Gompf, H.S., Mathai, C., Fuller, P.M., Wood, D.A., Pedersen, N.P., Saper, C.B., Lu, J., 2010. Locus ceruleus and anterior cingulate cortex sustain wakefulness in a novel environment. J. Neurosci. 30, 14543–14551.
Gray, N.S., Snowden, R.J., 2005. The relevance of irrelevance to schizophrenia. Neurosci. Biobehav. Rev. 29, 989–999.
Haselgrove, M., Le Pelley, M.E., Singh, N.K., Teow, H.Q., Morris, R.W., Green, M.J., Griffiths, O., Killcross, S., 2016. Disrupted attentional learning in high schizotypy: Evidence of aberrant salience. Br. J. Psychol. 107, 601–624.
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. Available at: 10.1016/j.neuroscience.2016.03.021.
Jazbec, S., Pantelis, C., Robbins, T., Weickert, T., Weinberger, D.R., Goldberg, T.E., 2007. Intra-dimensional/extra-dimensional set-shifting performance in schizophrenia: Impact of distractors. Schizophr. Res. 89, 339–349.
Kamin, L.J., 1968. "Attention-like" processes in classical conditioning. In: Miami Symposium on the prediction of behavior: Aversive stimulation. University of Miami Press, Miami, pp. 9–31. Available at: https://ntrs.nasa.gov/api/citations/19680013711/downloads/19680013711.pdf [Accessed March 20, 2024].
Kapur, S., 2003. Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. Am. J. Psychiatry 160, 13–23.
Keeler, J.F., Robbins, T.W., 2011. Translating cognition from animals to humans. Biochem. Pharmacol. 81, 1356–1366. https://doi.org/10.1016/j.bcp.2010.12.028.
Knott, T.S., Whyte, A.J., Dhawan, S.S., Tait, David S., Brown, Verity J., 2024. Dataset for "Blocking-like" effects in attentional set-shifting: redundant cues facilitate shifting in male rats with medial prefrontal cortex inactivation. Mendeley Data V1.
Kruschke, J.K., Blair, N.J., 2000. Blocking and backward blocking involve learned inattention. Psychon. Bull. Rev. 7, 636–645. https://doi.org/10.3758/BF03213001.
Laubach, M., Amarante, L.M., Swanson, K., White, S.R., 2018. What, if anything, is rodent prefrontal cortex? eNeuro 5.
Le Pelley, M.E., 2004. The role of associative history in models of associative learning: A selective review and a hybrid model. Quart. J. Exp. Psychol. Sect. B: Compar. Physiol. Psychol. 57, 193–243.
Le Pelley, M.E., McLaren, I.P.L., 2003. Learned associability and associative change in human causal learning. In: Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology, pp. 68–79.
Le Pelley, M.E., Schmidt-Hansen, M., Harris, N.J., Lunter, C.M., Morris, C.S., 2010. Disentangling the attentional deficit in schizophrenia: Pointers from schizotypy. Psychiatry Res. 176, 143–149. https://doi.org/10.1016/j.psychres.2009.03.027.
Leeson, V.C., Robbins, T.W., Matheson, E., Hutton, S.B., Ron, M.A., Barnes, T.R.E., 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. Biol. Psychiatry 66, 586–593. https://doi.org/10.1016/j.biopsych.2009.05.016.
Lubow, R.E., 1997. Latent inhibition as a measure of learned inattention: some problems and solutions.
Mackintosh, N.J., 1975. A theory of attention: Variations in the associability of stimuli with reinforcement. Psychol. Rev. 82, 276–298.
MacLaren, D.A.A., Browne, R.W., Shaw, J.K., Krishnan Radhakrishnan, S., Khare, P., Espana, R.A., Clark, S.D., 2016. Clozapine N-Oxide Administration Produces Behavioral Effects in Long-Evans Rats: Implications for Designing DREADD Experiments. eNeuro 3.
Maes, J.H.R., Eling, P.A.T.M., 2009. Do learned irrelevance and perseveration play a role during discrimination learning? Learn Motiv. 40, 274–283. https://doi.org/10.1016/j.lmot.2009.02.001.
Maes, J.H.R., Damen, M.D.C., Eling, P.A.T.M., 2004. More learned irrelevance than perseveration errors in rule shifting in healthy subjects. Brain Cogn. 54, 201–211.
Manvich, D.F., Webster, K.A., Foster, S.L., Farrell, M.S., Ritchie, J.C., Porter, J.H., Weinshenker, D., 2018. The DREADD agonist clozapine N-oxide (CNO) is reverse-metabolized to clozapine and produces clozapine-like interoceptive stimulus effects in rats and mice. Sci. Rep. 8.
Mohamed, W.M.Y., Unger, E.L., Kambhampati, S.K., Jones, B.C., 2011. Methylphenidate improves cognitive deficits produced by infantile iron deficiency in rats. Behav. Brain Res. 216, 146–152.
Murphy, K.L., McGaughy, J., Croxson, P.L., Baxter, M.G., 2017. Exposure to sevoflurane anesthesia during development does not impair aspects of attention during adulthood in rats. Neurotoxicol. Teratol. 60, 87–94. https://doi.org/10.1016/j.ntt.2016.11.010.
Pantelis, C., Barnes, T.R., Nelson, H.E., Tanner, S., Weatherley, L., Owen, A.M., Robbins, T.W., 1997. Frontal-striatal cognitive deficits in patients with chronic schizophrenia. Available at: Brain 120, 1823–1843 http://brain.oxfordjournals.org/

content/120/10/1823%5Cnhttp://brain.oxfordjournals.org/content/120/10/1823.full.pdf%5Cnhttp://brain.oxfordjournals.org/content/120/10/1823.short%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/9365373.
Pantelis, C., Barber, F.Z., Barnes, T.R.E., Nelson, H.E., Owen, A.M., Robbins, T.W., 1999. Comparison of set-shifting ability in patients with chronic schizophrenia and frontal lobe damage. Schizophr. Res. 37, 251–270.
Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev. 87, 532–552.
Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Available at: Classical Conditioning II Current Res. Theory 21, 64–99 http://homepage.mac.com/sanagnos/rescorlawagner1972.pdf.
Robbins, T.W., 2007. Shifting and stopping: Fronto-striatal substrates, neurochemical modulation and clinical implications. Philos. Trans. R. Soc. B 362, 917–932.
Robbins, T.W., 2017. Cross-species studies of cognition relevant to drug discovery: a translational approach. Br. J. Pharmacol. 174, 3191–3199.
Robinson, S., Granata, L., Hienz, R.D., Davis, C.M., 2019. Temporary inactivation of the medial prefrontal cortex impairs the formation, but not the retrieval of social odor recognition memory in rats. Neurobiol. Learn. Mem. 161, 115–121.
Roiser, J.P., Stephan, K.E., Den Ouden, H.E.M., Barnes, T.R.E., Friston, K.J., Joyce, E.M., 2009. Do patients with schizophrenia exhibit aberrant salience? Psychol. Med. 39, 199–209.
Roth, B.L., 2016. DREADDs for Neuroscientists. Neuron 89, 683–694. https://doi.org/10.1016/j.neuron.2016.01.040.
Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annu. Rev. Neurosci. 23:473–500. Available at: https://www.annualreviews.org/content/journals/10.1146/annurev.neuro.23.1.473 [Accessed March 20, 2024].
Sharpe, M.J., Killcross, S., 2014. The prelimbic cortex contributes to the down-regulation of attention toward redundant cues. Cereb. Cortex 24, 1066–1074.
Sharpe, M.J., Killcross, S., 2018. Modulation of attention and action in the medial prefrontal cortex of rats. Psychol. Rev. 125, 822–843.
Slabosz, A., Lewis, S.J.G., Smigasiewicz, K., Szymura, B., Barker, R.A., Owen, A.M., 2006. The role of learned irrelevance in attentional set-shifting impairments in Parkinson's disease. Neuropsychology 20, 578–588.
Smith, K.S., Bucci, D.J., Luikart, B.W., Mahler, S.V., 2016. DREADDs: Use and Application in Behavioral Neuroscience. Behav. Neurosci. 130, 1–19.
Snigdha, S., Neill, J.C., McLean, S.L., Shemar, G.K., Cruise, L., Shahid, M., Henry, B., 2011. Phencyclidine (PCP)-induced disruption in cognitive performance is gender-specific and associated with a reduction in Brain-Derived Neurotrophic Factor (BDNF) in specific regions of the female rat brain. J. Mol. Neurosci. 43, 337–345.
Tait, D.S., Marston, H.M., Shahid, M., Brown, V.J., 2009. Asenapine restores cognitive flexibility in rats with medial prefrontal cortex lesions. Psychopharmacology (Berl) 202, 295–306.
Tait, D.S., Bowman, E.M., Neuwirth, L.S., Brown, V.J., 2018. Assessment of intradimensional/extradimensional attentional set-shifting in rats. Available at: Neurosci. Biobehav. Rev. 89, 72–84 http://linkinghub.elsevier.com/retrieve/pii/S0149763417307510.
Vertes, R.P., 2004. Differential Projections of the Infralimbic and Prelimbic Cortex in the Rat. Synapse 51, 32–58.
Waltz, J.A., Kasanova, Z., Ross, T.J., Salmeron, B.J., McMahon, R.P., Gold, J.M., Stein, E.A., 2013. The Roles of Reward, Default, and Executive Control Networks in Set-Shifting Impairments in Schizophrenia. PLoS One 8.
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. Behav. Brain Res. 372, 112037. https://doi.org/10.1016/j.bbr.2019.112037.
Weiner, I., 1990. Neural substrates of latent inhibition: the switching model. Psychol. Bull. 108, 442–461.
Weiner, I., Feldon, J., 1997. The switching model of latent inhibition: an update of neural substrates.
Winefield, A.H., 1978. The effect of prior random reinforcement on brightness discrimination learning in rats. Q. J. Exp. Psychol. 30, 113–119.