Title: 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

Authors: Jiachao Wang, David S. Tait, Verity J. Brown, Eric M. Bowman

PII: S0166-4328(19)30210-4
DOI: https://doi.org/10.1016/j.bbr.2019.112037
Article Number: 112037

Reference: BBR 112037

To appear in: *Behavioural Brain Research*

Received date: 9 February 2019
Revised date: 12 June 2019
Accepted date: 12 June 2019

Please cite this article as: Wang J, Tait DS, Brown VJ, Bowman EM, 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* (2019), https://doi.org/10.1016/j.bbr.2019.112037

Title: 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

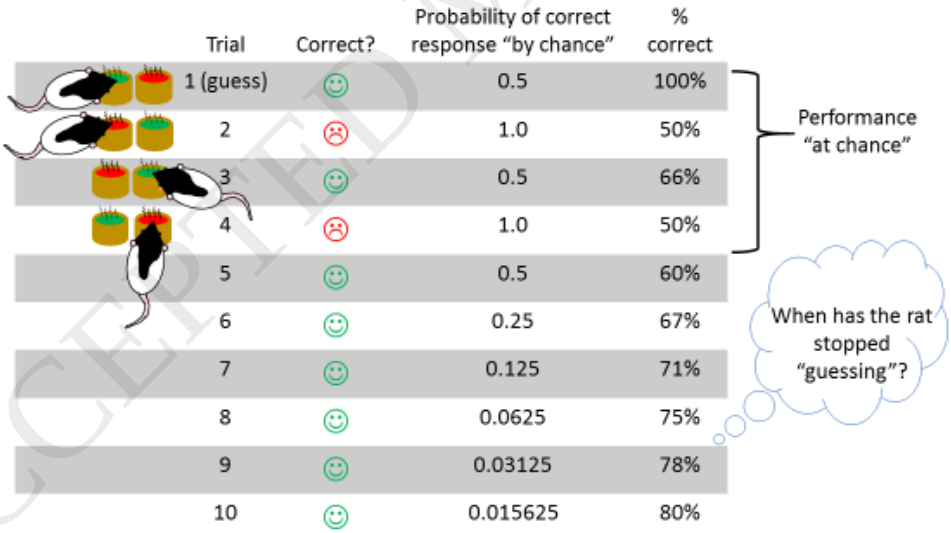
Abbreviated title: Cognitive flexibility in the pre-solution period of learning

Authors: Jiachao Wang, David S Tait, Verity J Brown, Eric M Bowman

Affiliation: School of Psychology & Neuroscience, University of St Andrews St Mary's Quad, South Street, St Andrews, KY16 9JP, U.K.

*__Corresponding Author:__ Verity J Brown (https://orcid.org/0000-0001-5762-1797)
School of Psychology & Neuroscience, University of St Andrews
St Mary's Quad, South Street, St Andrews, KY16 9JP, U.K.
vjb@st-andrews.ac.uk
Tel: +44 (0) 7792 383 103

Graphical Abstract



Abstract:

Our internal models of the world help us to process information rapidly: in general model-based learning is more rapid than model-free learning. However, the cognitive flexibility required to overcome cognitive predispositions can let us down: it is not fully developed until adulthood; predispositions can be unconscious biases; and cognitive flexibility is impaired in many psychiatric and neurological conditions. To understand these limits to flexibility, we need to know how brain generates predispositions and deploys flexibility.

We performed a detailed analysis of the exploratory behavior of rats in the pre-solution period of a two-alternative forced choice discrimination learning task. Rats readily learn in which of two bowls, filled with differentially scented and textured digging materials, there is hidden bait. In a single session, they are presented with a series of discrimination learning and reversal stages. We performed a simple Bayesian analysis on the data from 68 rats, 33 of which had lesions of the medial prefrontal cortex), to examine patterns of responding in the pre-solution period.

Control rats rapidly focussed on the relevant stimulus attributes and showed flexibility when required to learn about a different stimulus attribute. Rats with prefrontal cortex damage had reduced sensitivity to negative feedback. They were able to overcome this deficit and solve the credit assignment problem when there were limited alternatives or when attention was appropriately focused and predispositions matched the required response. However, the learning impairment presents as a problem with shifting attention due to the additional difficulty of solving the credit assignment problem when the attentional set is inconsistent with the required response.

Keywords: prefrontal cortex; set-shifting; executive function; Bayes; learning

1 Introduction

Understanding how cognitive predispositions – also called 'mental set' – shape perception, decision-making and learning is important, not least because predispositions can constrain thoughts and decisions. Indeed, cognitive biases that are unconscious can be particularly pernicious, as unacknowledged biases can drive actions with wide-ranging implications for society (for reviews see [1–3]). Furthermore, we cannot always rely on cognitive flexibility to overcome predispositions: it is not fully developed until adolescence[4,5] and it is impaired in many psychiatric and neurological disorders (e.g., [6]). It is even impaired in healthy, but stressed, adults, for example when under academic pressure[7].

The fact that cognitive flexibility is impaired in so many human clinical conditions inspired inclusion of the intradimensional/extradimensional (ID/ED) shifting task[8] in the Cambridge Neuropsychological Test Automated Battery (CANTAB)[9], which is used in both humans and monkeys, in clinical[10,11] and preclinical research[12,13]. Versions of the ID/ED task for rats and mice have also been developed (e.g., [14–23]), and used extensively for preclinical research testing potential pharmaceutical treatments. In the ID/ED task, the subject learns a two-alternative forced-choice discrimination between complex, multidimensional stimuli. Initial discriminations are based on one stimulus feature (for example, the color) and other features of the stimuli (for example, shape or size) are uncorrelated with feedback and are therefore irrelevant. Because there are multiple differences between the two stimuli, both perceptual and spatial, learning depends on correctly assigning 'credit' to the relevant stimulus element. The credit assignment problem is most acute when feedback is unreliable (e.g., probabilistic) or stimuli are ambiguous and the relevance of different features can change[24]. Assigning credit to particular stimulus feature is a cognitive bias: prior experience of relevancy results in a tendency to assume the same feature will still be relevant. For this reason, subsequent learning of novel ambiguous stimuli is more rapid if the discrimination is based on the same stimulus feature (an intradimensional (ID) shift), compared to when another, previously irrelevant, stimulus feature signals the solution (an extradimensional (ED) shift). The greater rapidity of ID, compared to ED, learning (generally, the number of trials to a criterion level of performance) reflects the need to suppress the cognitive predisposition to attend to previously relevant information and shift attention to newly relevant features. Thus, the difference between ID and ED, or the 'cost' of shifting attention, is used as an operational measure of cognitive flexibility (e.g., see [25]).

Variants of the ID/ED task have been used to explore the nature of the shift cost. For example,

whether it is an inability to suppress an established cognitive set (i.e., perseveration) or an inability to shift attention once disengaged (i.e., a problem overcoming learned irrelevance)[26–28]. These approaches compare the relative rates of learning, but they do not provide direct information about the learning process as it unfolds.

To focus on the learning process, it is necessary to examine responding in the period leading up to the point at which learning is complete: the pre-solution period. The point at which an animal is deemed to have learned a solution in a discrimination task is determined by a statistical analysis of performance based on frequentist (e.g., *x*-correct-out-of-*y* trials) null hypothesis testing. Specifically, the probability of a correct response is greater than chance. However, this does not mean that responding in the pre-solution period is *by* chance (random) even when the probability of a correct choice is *at* chance (the null hypothesis accepted). Behavior in the pre-solution period can be characterized by looking at patterns of errors. For example, in a reversal learning context, failure to find the solution could be due to perseverating with a previous solution or to a problem in finding or maintaining a new solution (e.g., [21]). However, when the stimulus-response or response-reward association is novel, the meaning of an incorrect response is ambiguous.

It has long been acknowledged that animals (including humans) do not respond randomly. Indeed, trying to generate random responses is remarkably difficult and taxes executive function[29]. Even rats make 'guesses' with purpose[30–32]. Rats have cognitive predispositions, leading them to process stimuli differently according to prior experience of the relevance of different features[8,33–35]. It is well established that, when placed in an environment where they can find food, rats "*experiment with many solutions*.. [and these] *attempted solutions… are within the rat's customary range of activity*" (p.135[30]). Furthermore, it has been suggested that the rat may simultaneously be entertaining more than one hypothesis, so that the "*pre-solution period is not sharply delimited into several series of mutually exclusive hypotheses*" (p.108[36]). Thus, at different points in the pre-solution period, an animal may know nothing about the stimuli, may have ruled out one or more potential solutions, or may have increasing confidence in the correct stimulus, while still responding 'at chance' with respect to the solution. Using the number of trials to reach a criterion for learning is neither intended nor able to address the psychological processes taking place in the pre-solution period. To understand cognitive biases and flexibility in model-based learning (see [37]) requires knowledge of behavior in the pre-solution period as the animal is testing various hypotheses in search of a solution.

To our knowledge, there has been no detailed examination of patterns of both correct and incorrect responses across the entire pre-solution period in individual animals, in order to identify cognitive predispositions and how these are overcome to find a solution. Therefore, in the present study, we describe a straightforward analysis of behavior in the pre-solution period of discrimination learning based on Bayes' theorem. A Bayesian approach has been used to model the decision-making processes, for example in solving "bandit" problems[38] or to model reinforcement learning[39,40]. However, we are using it here in a simpler way, not to produce a model of latent cognitive processes, but rather to quantify the consistency of behavioral data with hypothetical response patterns.

The approach was evaluated on archival data from a large number of rats (n=68) performing the ID/ED task using the bowl-digging paradigm (see Figure 1)[23]. This task is particularly well suited to illustrating the value of analysis of pre-solution behavior because there are a number of potential solutions at each learning stage; the solutions change at different points in the task; and the task is completed in a single test session, with each learning stage following immediately after the previous so over-night retention of learning is not an issue. The Bayesian analysis of the data of individual rats provides detailed information about behavioral patterns emerging over trials during learning. We were particularly interested to compare the patterns in the pre-solution period of learning in the context of consistent mental set (the ID stage) and inconsistent set (the ED stage). Nevertheless, although this Bayesian approach is applied here to a particular task in rats, it is sufficiently general to analyse the pre-solution period behavior of any discrimination learning regardless of species and task. It thus offers a powerful approach for the study of learning in intact or impaired animals.

2 Methods

2.1 The data set

2.1.1 Animals

We analysed high-quality archival data from 68 male Lister Hooded rats (Harlan or Charles River, UK) performing the ID/ED task. Thirty-three had received excitotoxic lesions of the medial prefrontal cortex (mPFC) (Figure 2) and the remainder (n = 35) were controls that had undergone sham-mPFC lesion surgery. The use of the term 'mPFC' has been recently challenged, with 'prelimbic' suggested as a preferred term for referring to this area of the rat brain[41]. We do acknowledge that nomenclature is an issue, particularly when comparing different species and perhaps inferring homology or functional equivalence. The medial

prefrontal cortex has anatomical and probably functional subdivisions. However, we have not attempted to subdivide these rats according to lesion extent. They were selected for inclusion because the lesions were centered on Area PL. Referring to the lesions as prelimbic would not be appropriate, however, because it might mislead, implying that they do not extend well beyond prelimbic cortex (Area PL). We are using medial prefrontal cortex (abbreviated to mPFC) in a descriptive sense, to indicate the location of lesion damage. We acknowledge that the primate brain has prefrontal cortex on the medial wall but, with different afferent and efferent connections, it is unlikely to be homologous to the rat's medial wall prefrontal cortex. Therefore, we stress that in referring to mPFC we are referring to location and we do not intend to imply homology or functionally equivalence to medial prefrontal cortex of primates. The lesions were centred on Area PL, but also extended beyond to include Area IL and Cg1 in most rats (Figure 2).

At the time of surgery, the rats weighed from 271g to 588g but there was no difference in the weight of animals with lesions compared to the controls, either before or after surgery.

The work was carried out in accordance with the UK Animals (Scientific Procedures) Act 1986, under UK project licences 60/3138, 60/3837, and 60/4459, and was approved by the University of St Andrews Animal Welfare and Ethics Committee.

The rats came from different cohorts, tested in the years 2006-2009. They were tested using the same protocol[23], with counterbalancing of stimuli and direction of shift within each cohort. At the time of testing, they were naive to behavioral testing and, specifically, to the ID/ED test: the data came only from the first test that each individual rat performed. All rats went on to be used in other studies. In particular, rats with mPFC lesions were used to test the effects of various pharmaceutical compounds, but the data presented here is prior to pharmacological manipulations. As many of these studies were commissioned by pharmaceutical companies, the results have not been and will not be published. Data from 20 of the rats (10 lesion and 10 control) are included in [42]. All intact rats are also included in a larger, but more heterogeneous, dataset of 375 intact rats reported in [23].

The rats were housed in our standard vivarium conditions, in pairs, with a 12-h light/dark cycle (lights on at 7:00 am; testing in the light phase). They were maintained on a diet of 15-20 g of standard lab chow per rat per day, given after testing, and water was available *ad libitum* in the home cage.

2.1.2 Surgery

The rats were anaesthetized with isoflurane and were given single fixed doses of the anti-inflammatory, carprofen (Rimadyl™ Pfizer, UK; 2.5mg in 0.05ml, s.c.) and diazepam (1.25mg in 0.25ml, i.p.). They were placed into a stereotaxic frame (tooth-bar +5.0 mm). Lesions were made by infusing 0.3µl of ibotenic acid (0.06M, in phosphate buffer) over two min at each of two mPFC sites bilaterally using a 0.5 µl Hamilton syringe with a 30-gauge needle. The coordinates of the injection sites were: (1) AP +3.5 mm; ML±0.6 mm; DV −5.2 mm from skull surface, and (2) AP +2.5 mm; ML±0.6 mm; DV −5.0 mm from skull surface[43]. The control rats were treated in the same way, except that the infusions were phosphate buffer vehicle solution only.

2.1.3 The task

The task involved choosing between two complex multidimensional stimuli (namely, bowls filled with various digging materials with different odors added[23] to retrieve food bait (half of a Honey Loop; Kellogg, Manchester, UK) buried in one of the bowls. The rats discovered, by trial and error learning, which bowl had the food. There were seven discrete stages, which were performed in a fixed order (see Figure 1) according to a standardized testing protocol. Each stage ended, and the next began, when the rat made six correct choices consecutively (i.e., a frequentist learning criterion). For most stages of the test, this was typically achieved in <12 trials and the entire task was completed in a single session lasting up to several hours.

A pre-determined pseudo-random trial order was used for each stage and this was the same for all rats (although there were instances in which a rat did not dig: after 10 minutes, the trial was terminated, and the next pair of bowls was presented). The trial order determined which pair of bowls was used (i.e., the odor/digging material combination of each bowl) and the left or right position of each bowl (and therefore the location of the reward) for each trial. We imposed constraints (i.e., counterbalancing) in the trial order to minimize the reinforcement of spurious patterns. For the first four trials of each stage, no trial repeated: each pair of bowls was presented twice, once in each of the two spatial configurations. These four trials thus potentially provide ‘perfect information’ to the rat about the association of the stimulus exemplars with reward. In addition, for these trials, if the rat dug first in the un-baited bowl, the trial was recorded as incorrect but the rat was permitted to subsequently dig in the other bowl, thus finding and consuming the bait. On the fifth and subsequent trials, the rat was permitted to explore both bowls but was only permitted to dig in one of them.

For each individual response in the pre-solution period, it is not possible to know on which stimulus element(s) the choice is based. However, it is possible to estimate the extent to which a choice is consistent with a pattern evident across previous choices. From the observed behavior on each trial we know: which stimulus exemplars (one odor; one medium) characterized the chosen bowl; the left or right location of the bowl; and whether the rat dug in the first bowl or explored both before digging in either. Based on extensive experience observing rats' responding in this task, we have seen evidence to suggest that, at different points in the pre-solution period, responding might resemble a spatial hypothetical response pattern (which could be 'place' hypotheses (e.g., location-perseverative or location-alternating), or a 'spatial response' hypothesis, such as win-stay/loose-shift or the converse) or a perceptual hypothetical response pattern, based on the differing properties of the bowls (i.e., odor or media). We generated eight potential hypothetical response patterns that were plausible possibilities to label the rats' behavior and applied Bayes' rule to all of the previous trials to estimate the posterior probability that the most recent choice was consistent with each of the hypothetical response patterns.

2.1.4 Updating the prior probability

Initially, at the start of the SD stage, when there was only one perceptual dimension, there were six potential hypothetical response patterns, so the priors were initially all set to one sixth (i.e., 0.167). For the CD and subsequent stages, there were eight potential response patterns.

The likelihood associated with each hypothesis was determined by the choice the rat made and whether it dug in the first bowl approached (favoring the spatial hypotheses) or sampled both bowls before selecting where to dig (see Figure 3).

The posterior probabilities were updated with each choice made by the rat, using Bayes' rule. Specifically, at the first trial of a new stage, all the priors are equal. With each subsequent response, the previous trial's posterior probability is the prior for the current trials. The prior is multiplied by the likelihood, and then this value is normalized (i.e., each is divided by the sum of the updated posterior probabilities of all hypotheses). This gives a quantitative indication of the consistency between the observed response pattern and the response pattern expected if hypothesis h_i governed responding over successive trials, with values increasing towards unity indicating consistency and values decreasing to approach zero indicating inconsistency. The accompanying annotated data file provides the raw data (trial by trial choices made by each rat, in every stage), as well as the likelihood, given the actual

response, and the resulting normalised posterior probabilities for each trial.

At the start of the CD stage, the priors were the posterior probabilities from the final trial of the SD stage, but because two additional priors were introduced at one eighth each, with the values of the SD priors were adjusted proportionately. All priors were reset to one eighth when the new stimuli were introduced at the ID and the ED stages. The calculations were performed using MATLAB, running in a Windows 10 environment. We used SPSS v24 for statistical analyses.

2.1.5 Estimating the likelihood of hypothetical response patterns

If a hypothetical rat chose a bowl based on a spatial rule (whether location-based or a response strategy, such as win-shift), the sensory properties of the bowl would be unknown when the location decision was made. Neither would the subsequently-discovered properties of the bowl influence the decision to dig, as the choice of location was already made. If the rat digs in the first bowl it encounters, it could be because the decision to dig was based on bowl location or because the first-encountered bowl had the perceptual properties the rat was seeking. However, if the rat does not dig in the first bowl it encounters, we have additional information: namely, the rat is acquiring perceptual information and is likely not making a choice based entirely on spatial location. Therefore, different decision trees were used to compute the likelihood of spatial and perceptual hypothetical response patterns (Figure 3) to decrease the weight given to spatial hypothetical response patterns if the rat explored both bowls before making a choice.

3 Statistical treatment of the data

All the findings reported below were obtained with a specific set of pre-fixed parameter values in the Bayesian model. Some of the values, for example those assigned to differentially weight spatial hypotheses based on exploration pattern, were selected based on reasonable assumptions about the behavior. Nevertheless, they are estimations and re-analysis of the data with different sets of parameters (particularly in the likelihood function) did not reveal qualitative differences in the outcome, suggesting that these parameters are as robust as others.

3.1 Confirmation of validity of assumptions

The eight hypothetical response patterns were proposed as plausible possibilities based on our experience and observations of many hundreds of rats performing this task as well as

published accounts of rat behavior in similar behavioral tasks. Nevertheless, it is necessary to confirm that each pattern is seen, in at least some of the rats, in the data. It is worth restating at this point that we are labelling hypothetical patterns to describe the characteristics of a response pattern: we do not intend to imply or infer a mental state (and, in particular, not 'having a hypothesis') to the rat.

We shall refer to the Bayesian posterior probability associated with each hypothetical pattern as the 'b-value'. We set a minimum for the b-value: this was done to acknowledge that there can never be certainty with respect to the absence of a pattern. In addition, all the b-values sum to one at each trial, which means that there is also a maximum for any one hypothetical response pattern. When the b-value for any one of the hypothetical patterns is above 0.5, this indicates that there is more evidence for this pattern than for other patterns. However, because it is possible that another hypothesis could also have a high value (>0.4), we chose a more conservative level (b-value >0.6) to accept as evidence, on a trial-by-trial basis for each individual rat, that the response choice was consistent with one of the hypothetical response patterns. With one value >0.6, all others must be <0.4. As judged by this criterion, all rats, at some point in the test, generated response patterns consistent with at least one of the spatial patterns. In particular, the majority of rats showed response patterns that increased the b-values of the perseveration (62/68 rats; 91%) and win-shift (56/68 rats; 82%) patterns to >0.6 and about half the rats, for one or more trials, responded in a manner that resulted in b-values >0.6 for the alternation and/or win-stay hypothetical patterns.

Having ascertained that the hypothetical patterns were reasonable assumptions and the consistent patterns were observed in the collected data, we next sought to determine the relationship between the frequentist criterion used in the task (six correct trials in a row, with a p-value on the final trial of 0.0156) and the final Bayesian posterior probability associated with the reward-related hypothetical pattern. For each stage, for each of the 68 rats, we computed the posterior probability associated with the rewarded hypothetical pattern on the final trial of each stage, when the learning criterion of 6-consecutive correct trials had been met. When a b-value is higher than 0.5, this indicates that one pattern is dominating. Higher b-values generally indicate increased confidence that a pattern is evident in the data.

It is important to note that the b-values are relative, rather than absolute, indicators of confidence in a pattern. Furthermore, some differences arise from the way the b-values are computed and therefore do not indicate anything meaningful about behavior. For example, the SD and CD stages had higher mean b-values on the final trials (SD stage = 0.87 (SEM = 0.03); CD stage = 0.92 (SEM = 0.02)) compared to the five following stages (range of mean b-values:

0.78 - 0.80 (SEM = 0.02 - 0.03) (Main effect of Stage: corrected-$F_{(5.5, 359.4)}$ = 3.9, p = 0.001, η_p^2 = 0.06). Planned comparisons confirmed the SD and CD did not differ, but both differed from each of the other five stages (all $p < 0.05$), none of which differed from each other. However, this effect arises from the way we computed the b-values: the b-values in the SD stage are higher because there are only six hypotheses so for the final six correct trials, each b-value increment is greater in this, compared to the other stages. At the start of the CD (when the same stimulus is still associated with reward), the high b-values for the relevant SD stimulus carry forward (lowering slightly, as b-values are assigned to irrelevant dimension stimuli), meaning that correct responses in the CD stage increment the b-value from an initial high base. By contrast, at the ID and ED stages, all b-values reset to 0.125 (one eighth) at the start of the stage. At the reversal stages, the lower values of the now-rewarded/previously-unrewarded stimulus are the ones that carry forward. Therefore, it is to be expected that the final b-values will tend to be slightly higher for the SD and CD stages. We did not seek to 'correct' this computationally, to avoid over-complicating the analysis, but we note that this difference between these stages does not have implications for cognition.

4 Results

4.1 Performance according to the frequentist criterion

It is well established (see [23]) that the different stages of the ID/ED task require a different number of trials to reach the learning criterion, and this was true for each of the subsets of the data included here. Thus, unsurprisingly, the effect of stage was seen in the larger group overall with a large effect size ($F_{(4.33, 277.7)}$ = 22.7, $p < 0.001$, η_p^2 = 0.26) (Figure 4, upper panel). Intact rats completed the CD stage in 9.54 (SEM = 0.55) trials and required 5.57 (SEM = 1.28) additional trials to complete the first reversal. Compared to the ID acquisition stage, intact rats required an additional 3.11 (SEM = 0.97) trials at the ED stage. Similarly, consistent with the statistical significance of the pattern observed in each subset of data, rats with mPFC lesions required a greater number of additional trials at the ED stage (Stage x Group: corrected-$F_{(4.34, 277.7)}$ = 4.1, p = 0.002, η_p^2 = 0.06; planned comparison of Group at ED: corrected-$F_{(1, 277.7)}$ = 21.5, $p < 0.001$, η_p^2 = 0.25). There was also a difference between the groups in the final reversal stage following the ED (planned comparison at Reversal 3: corrected-$F_{(1, 277.7)}$ = 5.74, $p < 0.02$, η_p^2 = 0.08). This is a weaker effect that is statistically reliable only when data are pooled across studies. Discrimination of odors or media was not statistically different (main effect of Dimension: $F_{(1, 64)}$ = 2.95, p = 0.09, η_p^2 = 0.04; Stage x Dimension: $F_{(4.34, 277.7)}$ = 2.05, p = 0.08, η_p^2 = 0.03).

It is impossible to know in the pre-solution period whether a correct response of an individual rat was merely a lucky guess ("at chance" responding). Nevertheless, with this large sample of rats, it is possible to determine that, as a group, there is strong evidence of learning within the initial six trials (Figure 4, lower panel), which is before the frequentist criterion of six consecutively correct responses could be met. Over 80% of the rats made a correct response on the fourth trial in the ID stage. In all three reversal stages, at the group level, there was no evidence of a period of perseverative responding: rather, an increasing proportion of rats made correct responses with each trial. Furthermore, the pattern is very similar in both groups, even in the ED stage.

4.2 Characterizing ED shift performance in the pre-solution period

4.2.1 Responding to stimuli in the irrelevant dimension

By the ED stage in this fixed-order test, the rats had already attained the frequentist criterion in five previous stages. For all stages prior to the ED, one of the perceptual dimensions (either odor or medium) had been relevant for finding the food bait. A principal motivation behind performing the current analysis of the pre-solution period response patterns was to identify the extent to which the rats initially respond to the stimuli in this perceptual dimension, perhaps trying first one and then the other exemplar, before ruling them out and 'shifting' responding to the other dimension. Furthermore, rats with mPFC lesions require significantly more trials to complete the ED stage. The b-values can indicate whether there is a yet greater propensity for rats with mPFC lesions to respond to the previously-relevant (now irrelevant) stimulus dimension.

We computed the number of trials per rat when the b-values were >0.6 for either of the stimuli in the irrelevant dimension in the ED stage. A surprisingly small number of animals responded in a consistent manner to stimuli in the irrelevant dimension (15/68; 22%). Furthermore, such responding was not more likely in the rats with mPFC lesions: 9/35 (27%) of controls and 6/33 (18%) of the rats with mPFC lesions had one or more trials where the b-values for stimuli in the irrelevant dimension were >0.6. For the nine control rats, the median number of trials where the b-value >0.6 was three (~25%) trials, compared to a median of four (~19%) trials for the six rats with mPFC lesions. Thus, there was no evidence to support the supposition that the additional trials to learn at the ED compared to the ID stage are due predominantly to persistent responding to stimuli in the newly-irrelevant dimension. Furthermore, there was no evidence to suggest that the deficit at the ED in mPFC-lesioned rats is due to an exaggeration of such persistent responding.

4.2.2 Spatial patterns of responding

As the lesioned rats were not responding to stimuli in the irrelevant dimension, we next looked to see whether there was any other consistent pattern in their responding. When the b-value associated with any of the hypothetical response patterns on any given trial in the ED or Reversal 3 stages was >0.6, the trial was classified according to whether responding was consistent with either a perceptual or a spatial hypothetical response pattern. Otherwise (i.e., when all b-values <0.6), the trial was classified as supporting 'no dominant pattern'.

In the ED stage, half of intact rats (18/35; 51%) had b-values >0.6 for spatial hypothetical patterns on one or more trials, accounting for 20% of the trials of the whole group. By contrast, a greater proportion of the rats with mPFC lesions (29/33; 88%) had b-values >0.6 for a spatial hypothetical pattern, accounting for over 40% of the trials in that group. A similar effect, albeit somewhat less pronounced, was seen in Reversal 3: 17/35 (49%) control and 24/33 (73%) of rats with lesions had spatial hypothetical pattern b-values >0.6 on one or more trials, accounting for 16% and 32% of trials, respectively. Trials with a b-value >0.6 for any one of the spatial hypothetical patterns accounted for *all* of the additional trials required by rats with mPFC lesions at both the ED and Reversal 3 stages (Figure 5; interaction of Group x Hypothetical pattern (Perceptual; Spatial; None): corrected-$F_{(1.73,113.96)} = 10.93$, $p < 0.001$; $\eta_p^2 = 0.14$; planned comparison for Group at Spatial, $p < 0.05$; interactions with Stage (i.e., ED; Reversal 3), n.s.). There was no difference between the groups in the number of trials classified as no pattern or perceptual. The additional trials by the rats with mPFC lesions at both the ED and Reversal 3 stages were accounted for by an increase in a pattern of responding matching a spatial, and not a perceptual, hypothetical pattern.

4.2.3 Emergence of the ED deficit

To see how quickly the lesion and control groups diverge in their responding to the new stimuli at the ED stage, we looked at the b-values in the initial six trials. Before the first ED trial, the priors are all set to equal (0.125) and, as trial one must be a 'guess' because the stimuli are new, any divergence of the priors at the first trial cannot be meaningful, even though the rat is collecting information and potentially learning from errors to guide subsequent choices. By the fourth trial, however, the rat has been exposed to full information about the correct stimulus: it has encountered both pairs of bowls twice, once in each of the two spatial configurations and has been allowed to recover the bait even when the initial dig was an incorrect choice.

As early as trial four of the ED, half of the rats (49%) of each group (lesion: 17/35; control: 16/33) had a b-value >0.6 for one of the hypothetical patterns. However, there was a group difference in the likelihood of the different hypothetical patterns. For the 16 control rats with a consistent pattern, an equal number were spatial (n = 8) or perceptual (n = 8) patterns and, of those with a perceptual pattern dominating, four rats were already responding consistently to the rewarded stimulus, whilst the other four rats were responding to one of the stimuli in the non-rewarded dimension. For the rats with mPFC lesions, most (15/17) response patterns supported a spatial hypothetical pattern and only two rats were responding with a perceptual pattern, both to stimuli in the irrelevant dimension. By trial six, 29% (10/35) of the control rats had b-values of >0.6 for the reward-related stimulus, and two of these rats had met the frequentist criterion (p = 0.015625). By contrast, for none of the lesioned rats was the b-value of the rewarded stimulus >0.6 by trial six.

Figure 6 shows the b-values, over the initial six trials, for the correct and the incorrect stimuli in the relevant dimension. It is clear that, in these early trials of the ED stage, the lesioned animals do not show the same increase in the b-value for the correct stimulus as controls. This indicates that the performance of the two groups diverges within the initial few trials, even though this is not because either group is persisting in responding to the previously-rewarded dimension.

There is another striking pattern in Figure 6, which is of particular interest: the rats with mPFC lesions also have lower b-values associated with the correct stimulus at stages other than the ED and Reversal 3. The divergence in the b-values of the lesioned and control animals over the initial six trials was statistically reliable (interaction of Group and Trial: $F_{(2.1,138.7)} = 5.54$, p = 0.004, $\eta_p^2 = 0.08$), but did not differ by stage (Stage X Group interaction: $F_{(4.22, 278.58)} = 2.15$, p = 0.07, $\eta_p^2 = 0.03$; Stage X Trial X Group: $F_{(12.75, 689.34)} = 1.32$, p = 0.196, $\eta_p^2 = 0.02$). This indicates that the groups are different very early in the pre-solution period and that this difference is not limited to the ED stage but is present throughout.

4.2.4 Final trial b-values are lower for rats with mPFC lesions

Figure 7 shows the b-values for the correct and the incorrect stimuli in the final six trials of each stage. Except for Reversal 1, where the curves for lesion and control groups precisely overlap, the b-values associated with the rewarded stimulus tend to be lower in the lesion group in every stage. On the final trial, the b-values for the rewarded hypothetical pattern were lower overall for the rats with lesions (mean = 0.80, SEM = 0.02) compared to controls (mean = 0.85, SEM = 0.02) (main effect of Group: $F_{(1,66)} = 4.6$, $p = 0.04$, $\eta_p^2 = 0.07$), but, as was

seen for the initial trials, there was no significant difference as a function of stage (Stage x Group: corrected-$F_{(5.4,\ 359.4)}$ = 1.4, p = 0.24, η_p^2 = 0.02). There are two important points to be made about this observation. The first concerns confidence in learning within each group. While the p-value on the final trial is the operational definition of learning (because it is the criterion), the b-value on the final trial is a measure of confidence in the presence of a behavioral pattern. Notwithstanding both the ED and Reversal 3 required more trials to reach the p-value frequentist criterion, for the rats with lesions the final b-values indicate, for all stages, an equivalent level of confidence in the presence of a pattern by the final trial. From this, we can have equal confidence in the strength of learning in the ED and Reversal 3 stages as for the other stages. The second point is more surprising: the main effect of group indicates that there is a generalized behavioral difference between the groups in the pre-solution period at *other* stages of the test. Previous analyses (e.g., trials to criterion or analysis of errors) had suggested there was no a behavioral effect of the lesion at other stages of this test. This Bayesian analysis picks it up both in the final b-values as well as in the initial trials of the pre-solution period (Section 4.2.3).

When using a frequentist criterion of n-consecutively-correct, the p-value resets to 1.0 with every error and thus reflects only the number of correct trials since the last error. The b-value, by contrast, reflects the entire response history from the point at which the priors are reset at the start of a new acquisition. If the response patterns in the pre-solution period are consistent with unrewarded hypothetical patterns, this will force the posterior probability of the reward-associated hypothetical pattern to be low immediately prior to the last six, correct, trials. Evidence *against* these previously supported alternative hypothetical patterns of responding, as well as evidence in support of the rewarded hypothetical pattern, must be accumulated to drive the reward-associated b-value higher. Therefore, the lower b-values of the rats with mPFC lesions suggest that the pattern of behavior in the pre-solution period is different compared to the controls, even in those stages where the rats achieve the frequentist learning criterion without additional trials to solve the discriminations.

4.3 Performance deficits in other stages of the test

As noted above (Sections 4.2.3 and 4.2.4), for the rats with lesions, the b-values of the rewarded hypothetical patterns were lower than controls in both the initial six trials and the final six trials, and this effect was not restricted to the ED or Reversal 3 stages. We noted that, at every stage, the rats with mPFC lesions were less likely to explore both bowls and more frequently dug in the first bowl approached (main effect of Group: $F_{(1,66)}$ = 11.07, p = 0.001, η_p^2

= 0.14; Group x Stage: $F_{(6,396)}$ = 1.13, ns, η_p^2 = 0.02). Spatial hypothetical response patterns, relative to the perceptual patterns, receive greater support if the rat does not explore both bowls prior to selecting the one in which to dig. Therefore, not exploring both of the bowls will have the effect of suppressing the b-values of perceptual patterns. It will only increase the b-values of spatial patterns if there is a consistent spatial choice, which was only seen in the ED and Reversal 3 stages.

4.3.1 Acquisition stages

At the SD and ID stages, there is no overall statistically significant group difference in the trials to the frequentist criterion. Nevertheless, there is an obvious and statistically significant suppression of the b-value, particularly at the SD and ED stages (Figure 7). Therefore, we examined the acquisition behavior of the rats in greater detail.

On the first trial, when the rats are encountering stimulus exemplars for the first time, they can only "guess" which bowl is baited. However, if the rat digs in the un-baited bowl on the third trial, it would have a chance to dig in the other bowl to recover the bait, which means that as early as trial three, the rat could have full information, having dug in each of the four bowls (there being two bowls in each pair) at least once. This is not true if the rat responded correctly on the third trial: it would not have dug in the un-baited bowl of the second pair. Given that patients with schizophrenia have impaired processing of negative feedback[44], which is detectable in the very early trials of the Wisconsin Card Sort Test[45], we looked at the responses of the rats to positive and negative feedback in the acquisition stages (SD, ID and ED).

For intact rats, after positive feedback (i.e., a correct response) on trial three, 71% of trial four responses were correct, compared with 91% after negative feedback (an incorrect response on trial three). This benefit following an error was expected, given that we were using a 'correction' procedure and the rat would gain additional information from digging first in the un-baited bowl and then in the baited bowl. However, the rats with lesions showed no additional benefit from trial three errors: 61% of trial four responses were correct following positive feedback but only 59% of responses were correct after negative feedback (Chi-squared = 13.85, df = 3, $p < 0.01$).

As noted above (Section 2.1.5), we cannot know whether a rat that correctly digs in the first bowl it encounters made a lucky guess or had intended to seek out that bowl, but a rat that digs incorrectly, without seeking out another stimulus, may have made an unlucky guess. However, if a rat explored both bowls before digging, it is plausible to suggest that this is

indicating that the rat is 'seeking' a particular stimulus. On the fourth, fifth and sixth trials of the SD, no rat made an error if they explored both bowls: rats who moved to the other bowl before digging were always correct. This applied to 12/35 (34%) controls and 6/33 (18%) lesions on the fourth trial; 15/35 (43%) controls and 3/33 (9%) lesions on the fifth; and 12/35 (34%) controls and 13/33 (40%) lesions on the sixth trial. Fischer Exact tests were used to compare the proportion of rats in each group that made a correct dig having explored both bowls and those who made an incorrect dig in the first bowl encountered. In the SD stage, a greater proportion of control rats correctly dug in the second bowl on trials four ($p = 0.022$) and five ($p = 0.015$), but there was no group difference for trial six ($p = 0.73$). At the ED stage, but not the ID stage, there was also a difference in the proportion of rats from each group correctly digging in the second bowl on trials four ($p = 0.037$), five ($p = 0.005$) and six ($p = 0.032$).

Together, these observations are evidence for learning within the initial six trials, before the frequentist criterion could be met (see also Section 4.1; Figure 4, lower panel). It also suggests that the rats with mPFC lesions have a mild acquisition impairment, evident in the initial trials of the pre-solution period but quickly overcome. The nature of this impairment appears to be due to the rats deriving less benefit from negative (i.e., following an error) feedback.

4.3.2 Reversal learning

4.3.2.1 Spontaneous shifting during reversal learning

Very few rats (3/35 controls; 2/33 with lesions) had trials with b-values >0.6 for stimuli in the irrelevant dimension in Reversal 3. This suggests that, after the ED stage, neither control nor rats with mPFC lesions show a response tendency indicative of 'shifting back', despite more experience responding to the other dimension. It also rules out 'shifting back' as a possible explanation for the rats with mPFC lesions requiring more trials to reach the learning criterion of Reversal 3.

In Reversals 1 and/or 2, many more rats (30/68; 44%) had one or more trials in which the b-value for the irrelevant dimension was >0.6, suggesting that there was a tendency to 'spontaneously shift' responding to the dimension which was not associated with reward. However, there was no difference in the proportion of each group who showed evidence of responding to the other (irrelevant) dimension (15/35 controls; 15/33 with lesions).

4.3.2.2 Dimensional selectivity

In the reversal stages, when a rat stops responding to the previously-correct exemplar, it could immediately start to respond to the other, reversed, stimulus. Alternatively, there may be intervening trials when performance is 'at chance' (i.e., responding with no pattern, or responding with an inappropriate pattern, such as to a stimulus in the irrelevant dimension or matching one of the spatial hypotheses). Figure 8 shows b-values of two 'representative' rats across the seven stages of the test. In Reversal 1, both of these rats show a simultaneous cessation of responding to the previously correct stimulus as they begin responding to the alternative (now correct) exemplar. Clearly, immediately reversing in such a manner results in a high inverse correlation between the b-values for the correct and incorrect responses. Thus, the magnitude of this inverse correlation (i.e., converting the inverse correlation into a positive number) can be used as a dimensional selectivity index (DSI) during reversal learning, with values approaching 1.0 indicating efficient reversal without 'at chance' responding. Obviously, there could never be a positive correlation between these two b-values (the animal can only respond to either the correct or the incorrect exemplar), so the DSI will not fall to zero. However, if there are intervening trials when responding is at chance (either because responding favors another hypothetical response pattern or there is no pattern), the b-values for both the correct and incorrect responses will converge, diminishing the inverse correlation and so reducing the DSI.

Figure 9 shows a histogram of the distributions of DSIs for each group and each reversal. For Reversal 1, the median DSI was 0.74 (mean = 0.64; SEM = 0.03; range 0.996 to 0.1), but, for both groups, over half of the rats have high dimensional selectivity (DSIs >0.7, indicating that the correlation accounts for >50% of the variance) while the rest have DSIs <0.5, accounting for <25% of the variance.

It was expected that responding in Reversal 2 might have greater dimensional selectivity than Reversal 1, because the rats had more experience of attending to one dimension. However, this was not the case: if anything, there was slightly less dimensional selectivity. A slightly larger proportion of rats in both groups had DSIs of <0.5 at Reversal 2, decreasing the median DSI to 0.45 (mean = 0.52, SEM = 0.03; range 0.986 to 0.13). Nevertheless, the difference between Reversals 1 and 2 was not statistically significant for either group, nor did the groups differ from each other at either reversal (Fischer Exact Tests, all $p > 0.10$).

4.3.2.3 Relationship between DSI and trials to criterion

For all three reversal stages and for both groups, the trials to criterion and the DSIs were significantly and highly inversely correlated (Control (n = 35) and Lesion (n = 33), respectively, for Reversal 1: r = -0.67 and -0.77; Reversal 2 r = -0.73 and -0.57; and Reversal 3: r = -0.63 and -0.52; all $p < 0.01$, all accounting for >25% of the variance, which is a large effect size). This indicates that when reversing was less efficient (i.e., required more trials), it was not due to perseverative responding (which would maintain dimensional selectivity, keeping the DSI high, while the trials to criterion also increased). Rather, when rats (lesion or control) required more trials in the reversal, it was due to responding 'at chance' with respect to the correct stimulus.

Reflecting the increase in trials to criterion for Reversal 3, the rats with mPFC lesions had lower DSIs in Reversal 3 (median = 0.47, mean = 0.48) compared to the control group (median = 0.74, mean = 0.68) and, for the lesion group, the distribution of DSIs for Reversal 3 was significantly different to Reversal 1 (Fischer Exact Test: $p < 0.001$) but not Reversal 2 ($p = 0.08$).

4.3.2.4 Negative feedback

During the acquisition stages, rats with lesions showed less benefit of negative feedback (Section 4.3.1). However, in the reversal stages, the previously correct exemplar is still available and therefore, in the early trials of all three reversals, most of the responses of rats in both groups were errors and most of the feedback was negative. Interestingly, however, the rate of learning was not different between the groups in these early trials of reversals.

4.3.3 Consistency across different stages of the test

For the control rats, there was a relationship between performance the different stages of the task that was not seen in the rats with mPFC lesions. Intact rats who acquired the SD and CD in fewer trials also reversed more rapidly (positive correlation between n trials to criterion for the SD+CD and Reversal 1: $r = 0.47$, $p < 0.01$) and with higher DSIs (Inverse correlation with the DSI: $r = -0.41$, $p < 0.015$). This suggests that an early focus on the rewarded stimulus enhanced dimensional selectivity and reversal learning. Similarly, rapid and efficient reversal learning (indicated by high DSIs and low trials to criterion at Reversal 1) was also associated with rapid ID (inverse correlation with Reversal 1 DSI: $r = -0.396$, $p = 0.02$), ED performance ($r = -0.349$, $p = 0.04$) and more rapid reversal learning in Reversal 3 ($r = -0.375$, $p = 0.03$).

For the rats with mPFC lesions, DSIs at Reversal 1 were not associated with performance at any other stage: the largest (non-significant) correlation was with SD+CD trials to criterion (r =

0.25, n = 33, p = 0.17).

5 Conclusion

We have presented an analysis of behavior in the pre-solution period, comparing rats with lesions of the mPFC and controls performing a bowl-digging task involving intradimensional and extradimensional shifts and reversal learning. Although numerous studies (for review see [46]) using frequentist criteria ("x-correct-out-of-y" or "n-consecutive-correct" trials) have repeatedly confirmed that there is a deficit in this task following mPFC lesions that is specific to the ED stage of the test, the present Bayesian analysis, looking for patterns in responding in the period before a frequentist criterion for learning has been reached, has revealed differences resulting from the lesion in the pre-solution period in all of the stages of the task.

During data collection in this task, our standard protocol [46] is to use a 'correction' procedure in the initial four trials of each stage, meaning that the rats are permitted to dig and retrieve the bait from the other bowl if they initially dig incorrectly. This is a well-established method in animal learning and it is known to be associated with more rapid acquisition and faster reversal learning[47]. In the bowl-digging context, it also ensures that all rats have equivalent exposure to the correct stimulus, even if they make an incorrect response in the initial trials. Having a correction procedure is less of an issue in discrimination tasks where visual stimuli are presented adjacent to each other and both exemplars can be seen before one is selected, or in spatial discrimination task, where alternative spatial responses are equally salient and available. By contrast, exemplars of haptic or olfactory stimuli need to be spatially separated, meaning that they cannot be simultaneous sampled. This notwithstanding, the rats with lesions appeared to have learned less about the perceptual exemplars in the four correction trials than control intact rats had learned in the stages where new stimuli were presented (i.e, the SD, ID and ED stages). Specifically, on the fifth trial (the first trial that was not 'corrected') of both SD and ED stages, rats with lesions were less likely to sample both bowls, digging in the first bowl encountered, and thus more likely to make an incorrect response.

After the SD stage, both groups benefit from having an attentional set that serves to 'narrow' the search for the solutions at subsequent stages. An early focus on the correct exemplar in the SD/CD stages was associated with more rapid reversals in both groups. For intact rats, it was also associated with more efficient learning at both ID and ED stages. However, the rats with mPFC lesions learn inefficiently and this impairment became particularly marked at the ED stage when the attentional set is no longer relevant and the 'credit assignment problem'

once again requires to be solved.

The Bayesian analysis detected patterns in responding in the pre-solution period that revealed that rats with mPFC lesions behaved differently throughout the task, not just in the ED stage. For example, they learned the discriminations differently, even the very first, simple discrimination (SD). Although the frequentist criterion was met in an equivalent number of trials, there was evidence that rats with mPFC lesions derived less benefit from negative feedback than controls in the very early trials of learning. Similar effects have been reported in patients with schizophrenia: specifically, there was a learning impairment in the initial SD stage, even while the ID stage did not appear to be affected[48] and patients also show decreased sensitivity to negative feedback on the initial trials of the Wisconsin Card Sort Test, even before responding was established and they had been required to shift sorting categories[45]. In humans and monkeys, the dorsal anterior cingulate cortex (dACC) is thought to facilitate learning by signaling mismatches between the expected compared to experienced outcomes[49,50]. An apparent improvement in serial reversal learning with inactivation of Area PL[51] has been interpreted as resulting from reduced sensitivity to negative feedback. Serial reversal learning is different to the task used here in many respects: serial reversals are rehearsed and anticipated and appear to involve learning 'that reversals can occur' [52] and performance depends on sensitivity to the change in reinforcement contingency (i.e., changing behavior on the receipt of negative feedback). In our task, the animals have limited opportunity to learn 'reversals can occur': indeed, Reversal 1 is the first reversal the rats have ever encountered. We also see a change in sensitivity to negative feedback, but we detect it during acquisition rather than in these 'novel', as opposed to serial, reversals. There is ample evidence [46,53–55], including the present data, that inactivation of Area PL typically does not impair reversal learning with assured outcomes (i.e., traditional, rather than probabilistic, reversals). This is likely due to the fact that credit assignment is easier when feedback is unambiguous[24]. However, by using this Bayesian approach, we have been able to detect the same reduced sensitivity to negative feedback in acquisition stages that results in impaired performance in probabilistic reversal learning. The fact that the manifestation of a cognitive deficit as a behavioral impairment is task dependent is surely unsurprising. Nevertheless, it highlights the importance of understanding the nature of the differences between different tasks designed to tease apart cognitive functions.

When we embarked on this approach to the analysis of the data, we assumed we might see evidence of patterns in responding that would indicate something about attentional engagement. However, we did not see a tendency for the rats, whether intact or lesioned, to

respond persistently to stimuli in the previously relevant dimension – indeed, when the rats were unsure, the default response was to a spatial pattern of responding. Nevertheless, our intention was not to use this analysis to 'guess what the rat was thinking', but rather to deduce from systematic response patterns what the rat was *doing* in the pre-solution period. We were able to ascertain from this analysis that the rats with lesions responded differently to negative feedback in acquisition stages compared to intact rats, even though this difference was not detected using standard analyses of errors and correct responses. This suggests that the mPFC-lesion impairment that manifests as cognitive inflexibility when a shift of attention is required is better characterized as a general impairment in credit assignment that is most marked when ambiguity is increased.

Acknowledgements: JW received a PhD scholarship from the University of St Andrews. We would like to thank Drs Alexander Chase and Alonzo Whyte for their contributions to the development of this work. We would also like to thank Trevor Hawkey and the husbandry staff of the St Mary's Animal Unit for their expert care of the animals.

References

[1] A.G. Greenwald, L.H. Krieger, Implicit Bias : Scientific Foundations, Calif. Law Rev. 94 (2017) 945–967. https://www.jstor.org/stable/20439056.

[2] C.A. Moss-Racusin, J.F. Dovidio, V.L. Brescoll, M.J. Graham, J. Handelsman, Science faculty's subtle gender biases favor male students, Proc. Natl. Acad. Sci. 109 (2012) 16474–16479. doi:10.1073/pnas.1211286109.

[3] I.M. Handley, E.R. Brown, C.A. Moss-Racusin, J.L. Smith, Quality of evidence revealing subtle gender biases in science is in the eye of the beholder, Proc. Natl. Acad. Sci. 112 (2015) 13201–13206. doi:10.1073/pnas.1510649112.

[4] P. Anderson, Assessment and Development of Executive Function (EF) During Childhood, Child Neuropsychol. 8 (2002) 71–82. doi:10.1076/chin.8.2.71.8724.

[5] S. Doebel, P.D. Zelazo, A meta-analysis of the Dimensional Change Card Sort: Implications for developmental theories and the measurement of executive function in children, Dev. Rev. 38 (2015) 241–268. doi:10.1016/j.dr.2015.09.001.

[6] T.W. Robbins, A.F.T. Arnsten, The Neuropsychopharmacology of Fronto-Executive Function: Monoaminergic Modulation, Annu. Rev. Neurosci. 32 (2009) 267–287. doi:10.1146/annurev.neuro.051508.135535.

[7] C. Liston, B.S. McEwen, B.J. Casey, Psychosocial stress reversibly disrupts prefrontal processing and attentional control, Proc. Natl. Acad. Sci. 106 (2009) 912–917. doi:10.1073/pnas.0807041106.

[8] B.E. Shepp, P.D. Eimas, Intradimensional and extradimensional shifts in the rat, J. Comp. Physiol. Psychol. 57 (1964) 357–361. doi:10.1037/h0043967.

[9] A.M. Owen, A.C. Roberts, C.E. Polkey, B.J. Sahakian, T.W. Robbins, Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man, Neuropsychologia. 29 (1991) 993–1006. doi:10.1016/0028-3932(91)90063-E.

[10] A.M. Owen, J.J. Downes, B.J. Sahakian, C.E. Polkey, T.W. Robbins, Planning and spatial working memory following frontal lobe lesions in man, Neuropsychologia. 28 (1990) 1021–1034. doi:10.1016/0028-3932(90)90137-D.

[11] J.J. Downes, A.C. Roberts, B.J. Sahakian, J.L. Evenden, R.G. Morris, T.W. Robbins, Impaired extra-dimensional shift performance in medicated and unmedicated Parkinson's disease: evidence for a specific attentional dysfunction, Neuropsychologia. 27 (1989) 1329–1343.

[12] R. Dias, T.W. Robbins, A.C. Roberts, Dissociation in prefrontal cortex of affective and attentional shifts, Nature. 380 (1996) 69. http://dx.doi.org/10.1038/380069a0.

[13] A.C. Roberts, T.W. Robbins, B.J. Everitt, The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates, Q. J. Exp. Psychol. Sect. B. 40 (1988) 321–341. doi:10.1080/14640748808402328.

[14] J.M. Birrell, V.J. Brown, Medial frontal cortex mediates perceptual attentional set shifting in the rat, J. Neurosci. 20 (2000) 4320–4324.

[15] G.B. Bissonette, G.J. Martins, T.M. Franz, E.S. Harper, G. Schoenbaum, E.M. Powell, Double Dissociation of the Effects of Medial and Orbital Prefrontal Cortical Lesions on Attentional and Affective Shifts in Mice, J. Neurosci. 28 (2008) 11124–11130. doi:10.1523/JNEUROSCI.2820-08.2008.

[16] J.L. Brigman, T.J. Bussey, L.M. Saksida, L.A. Rothblat, Discrimination of

multidimensional visual stimuli by mice: Intra- and extradimensional shifts, Behav. Neurosci. 119 (2005) 839–842. doi:10.1037/0735-7044.119.3.839.

[17] C. Graybeal, M. Feyder, E. Schulman, L.M. Saksida, J. Bussey, Timothy, J.L. Brigman, A. Holmes, Paradoxical reversal learning enhancement by stress or prefrontal cortical damage : rescue with BDNF, Nat. Neurosci. 14 (2012) 1507–1509. doi:10.1038/nn.2954.

[18] D. Scheggia, A. Bebensee, D.R. Weinberger, F. Papaleo, The ultimate intra-/extra-dimensional attentional set-shifting task for mice, Biol. Psychiatry. 75 (2014) 660–670. doi:10.1016/j.biopsych.2013.05.021.

[19] A.M. Brady, S.B. Floresco, Operant Procedures for Assessing Behavioral Flexibility in Rats, J. Vis. Exp. (2015) 1–13. doi:10.3791/52387.

[20] R.C. Trueman, S.B. Dunnett, S.P. Brooks, Operant-based instrumental learning for analysis of genetically modified models of Huntington's disease, Brain Res. Bull. 88 (2012) 261–275. doi:10.1016/j.brainresbull.2011.03.015.

[21] S.B. Floresco, Dissociable Roles for the Nucleus Accumbens Core and Shell in Regulating Set Shifting, J. Neurosci. 26 (2006) 2449–2457. doi:10.1523/JNEUROSCI.4431-05.2006.

[22] M.E. Ragozzino, The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility, Ann. N. Y. Acad. Sci. 1121 (2007) 355–375. doi:10.1196/annals.1401.013.

[23] D.S. Tait, E.M. Bowman, L.S. Neuwirth, V.J. Brown, Assessment of Intradimensional/Extradimensional Attentional Set-Shifting in Rats, Neurosci. Biobehav. Rev. (2018). doi:10.1016/j.neubiorev.2018.02.013.

[24] W.T. Fu, J.R. Anderson, Solving the credit assignment problem: Explicit and implicit learning of action sequences with probabilistic outcomes, Psychol. Res. 72 (2008) 321–330. doi:10.1007/s00426-007-0113-7.

[25] D.R. Dajani, L.Q. Uddin, Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience, Trends Neurosci. 38 (2015). doi:10.1016/j.tins.2015.07.003.

[26] A.M. Owen, A.C. Roberts, J.R. Hodges, B.A. Summers, C.E. Polkey, N. Birbaumer, Contrasting mechanisms of impaired attentional-shift in patients with frontal lobe damage or Parkinson's disease, Brain. 116 (1993) 1159–1175.

[27] J. Gauntlett-Gilbert, R.C. Roberts, V.J. Brown, Mechanisms underlying attentional set-

shifting in Parkinson's disease, Neuropsychologia. 37 (1999) 605–616. doi:10.1016/S0028-3932(98)00049-9.

[28] J.H.R. Maes, M.D.C. Damen, P.A.T.M. Eling, More learned irrelevance than perseveration errors in rule shifting in healthy subjects, Brain Cogn. 54 (2004) 201–211. doi:10.1016/j.bandc.2004.01.003.

[29] A. Baddeley, H. Emslie, J. Kolodny, J. Duncan, Random Generation and the Executive Control of Working Memory, Q. J. Exp. Psychol. 51A (1998) 819–852.

[30] K.S. Lashley, Brain mechanisms and intelligence: A quantitative study of injuries to the brain., University of Chicago Press, 1929.

[31] I. Krechevsky, Brain mechanisms and "hypotheses.," J. Comp. Psychol. 19 (1935) 425–468. doi:10.1037/h0056416.

[32] K.W. Spence, The nature of discrimination learning in animals, Psychol. Rev. 43 (1936) 427–449. doi:10.1037/h0056975.

[33] D.H. Lawrence, Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus, J. Exp. Psychol. 39 (1949) 770–784. doi:10.1037/h0058097.

[34] D.H. Lawrence, Acquired distinctiveness of cues: II. Selective association in a constant stimulus situation, J. Exp. Psychol. 40 (1950) 175–188. doi:10.1037/h0063217.

[35] D.H. Lawrence, The transfer of a discrimination along a continuum., J. Comp. Physiol. Psychol. 45 (1952) 511–516.

[36] I. Krechevsky, A study of the continuity of the problem-solving process, Psychol. Rev. 45 (1938) 107–133. doi:10.1037/h0056641.

[37] B.B. Doll, D.A. Simon, N.D. Daw, The ubiquity of model-based reinforcement learning, Curr. Opin. Neurobiol. 22 (2012) 1075–1081. doi:10.1016/j.conb.2012.08.003.

[38] M. Steyvers, M.D. Lee, E.J. Wagenmakers, A Bayesian analysis of human decision-making on bandit problems, J. Math. Psychol. 53 (2009) 168–179. doi:10.1016/j.jmp.2008.11.002.

[39] Y. Niv, Reinforcement learning in the brain, J. Math. Psychol. 53 (2009) 139–154. doi:10.1016/j.jmp.2008.12.005.

[40] J.R. Busemeyer, T.J. Pleskac, Theoretical tools for understanding and aiding dynamic decision making, J. Math. Psychol. 53 (2009) 126–138. doi:10.1016/j.jmp.2008.12.007.

[41] M. Laubach, L.M. Amarante, K. Swanson, S.R. White, Cognition and Behavior What, If Anything, Is Rodent Prefrontal Cortex?, ENeuro. 5 (2018) 315–333. doi:10.1523/ENEURO.0315-18.2018.

[42] D.S. Tait, H.M. Marston, M. Shahid, V.J. Brown, Asenapine restores cognitive flexibility in rats with medial prefrontal cortex lesions, Psychopharmacology (Berl). 202 (2009) 295–306. doi:10.1007/s00213-008-1364-8.

[43] G. Paxinos, C. Watson, The Rat Brain in Stereotaxic Coordinates, 6th ed., Academic Press, 2007.

[44] E.A. Martin, A. McCleery, M.M. Moore, J.K. Wynn, M.F. Green, W.P. Horan, ERP indices of performance monitoring and feedback processing in psychosis: A meta-analysis, Int. J. Psychophysiol. 132 (2018) 365–378. doi:10.1016/j.ijpsycho.2018.08.004.

[45] K.J. Prentice, J.M. Gold, R.W. Buchanan, The Wisconsin Card Sorting impairment in schizophrenia is evident in the first four trials, Schizophr. Res. 106 (2008) 81–87. doi:10.1016/j.schres.2007.07.015.

[46] D.S. Tait, E.M. Bowman, L.S. Neuwirth, V.J. Brown, Assessment of intradimensional/extradimensional attentional set-shifting in rats, Neurosci. Biobehav. Rev. 89 (2018) 72–84. doi:10.1016/j.neubiorev.2018.02.013.

[47] C.L. Hull, K.W. Spence, “Correction” vs “non-correction” method of trial-and-error learning in rats, J. Comp. Psychol. 25 (1938) 127–145.

[48] V.C. Leeson, T.W. Robbins, E. Matheson, S.B. Hutton, M.A. Ron, T.R.E. Barnes, E.M. Joyce, 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 (2009) 586–593. doi:10.1016/j.biopsych.2009.05.016.

[49] B.Y. Hayden, S.R. Heilbronner, J.M. Pearson, M.L. Platt, Surprise signals in anterior cingulate cortex: Neuronal encoding of unsigned reward prediction errors driving adjustment in behavior, J. Neurosci. 31 (2011) 4178–4187. doi:10.1523/JNEUROSCI.4652-10.2011.

[50] A.R. Weiss, M.J. Gillies, M.G. Philiastides, M.A. Apps, M.A. Whittington, J.J. FitzGerald, S.G. Boccard, T.Z. Aziz, A.L. Green, Dorsal Anterior Cingulate Cortices Differentially Lateralize Prediction Errors and Outcome Valence in a Decision-Making Task, Front. Hum. Neurosci. 12 (2018) 1–15. doi:10.3389/fnhum.2018.00203.

[51] G.L. Dalton, N.Y. Wang, A.G. Phillips, S.B. Floresco, Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning, J. Neurosci. 36 (2016) 1996–2006. doi:10.1523/JNEUROSCI.3366-15.2016.

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

[53] M.E. Ragozzino, S. Detrick, R.P. Kesner, Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning., J. Neurosci. 19 (1999) 4585–94. http://www.ncbi.nlm.nih.gov/pubmed/10341256.

[54] S.B. Floresco, A.E. Block, M.T.L. Tse, 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 (2008) 85–96. doi:10.1016/j.bbr.2008.02.008.

[55] J.M. Birrell, V.J. Brown, Medial Frontal Cortex Mediates Perceptual Attentional Set Shifting in the Rat, J. Neurosci. 20 (2000) 4320–4324. doi:10.1523/JNEUROSCI.20-11-04320.2000.

Figure legends

Figure 1. The schematic shows, from top to bottom, the order of the seven stages of the ID/ED task. The first is a simple discrimination (SD) between two media with different textures (e.g., sand and gravel; media are represented in the illustration, but for half the rats, odor was the initially relevant dimension). One of the bowls is baited with half of a Honey Loop cereal piece (shown at the lower left of the bowl on the left, but hidden from the rat). Distinct odors (represented by lines above the bowls) are added to these media to create a compound stimulus for the compound discrimination (CD), but the same medium indicates which bowl is baited. The first reversal (Rev1) uses the same two pairs of bowls, but the bait is now in the alternative medium. New bowls are used for the intradimensional (ID) shift and its reversal (Rev2) but the relevant dimension is still medium. Finally, new bowls are introduced for the extradimensional (ED) shift and its reversal (Rev3), but the relevant dimension changes to odor.

Figure 2. Coronal sections of rat brain (from[43]) showing the maximum extent of any of the medial prefrontal cortical (mPFC) lesions (pale grey), the extent of a median size lesions (darker) and the area of common damage (darkest grey). The damage is represented on one hemisphere, but the lesions were bilateral. The damage was centred on prelimbic (PrL) and infralimbic (IL) cortex, as intended.

Figure 3. We nominated eight potential ('hypotheses') to describe potential behaviors. Four were associated with responding to perceptual properties of the stimuli and four were descriptions in spatial terms (left table). A value – the likelihood – was assigned to each of the eight hypotheses (right table) according to whether the rat dug in the first bowl it encountered or selected the bowl having explored both and whether the choice was consistent or not with that responding according to particular hypothesis. These values were used to update the posterior probabilities, on a trial by trial basis, using Bayes formula. The result was a b-value for each hypothesis for each trial and each rat.

Figure 4. The top panel shows a bar graph of mean (± SEM) trials to the frequentist criterion of six consecutively correct ($p = 0.015625$) for each of the seven stages of the task, shown from left to right in order of testing. Dark bars are the group with mPFC lesions; white bars are the intact group. *indicates different from control ($p < 0.05$). The lower panel shows a subset (the first six trials) of the same data to illustrate the percentage of each group who make a correct

response on each of the initial trials. Both groups show evidence of learning within the first six trials.

Figure 5. Stacked bar graph showing the mean total number of trials at the ED and Reversal 3 stages, for the control (left of each pair) and lesioned (right of each pair) rats. The total number of trials has been subdivided to show the mean number of trials in which no pattern dominated (black), or a perceptual (grey) or spatial (white) pattern dominated (b-value > 0.6 for any of the relevant hypotheses). Error bars: ± SEM for each type of trial. All of the additional trials at both the ED and Reversal 3 were accounted for by an increase in the number of trials in which spatial patterns dominated (* $p < 0.05$).

Figure 6. The b-values for the initial six trials, for the correct (green lines) and the incorrect (red lines) stimulus in the relevant dimension, are shown for each group (open symbols, control; closed symbols, lesion). For the acquisition stages (SD, ID and ED), the prior probabilities are all equal before the first trial and then they diverge; prior probability is not reset at the CD or reversal stages. There is a statistically significant difference between the groups as a function of trial, but no interaction with stage. Nevertheless, the difference is particularly marked at the ED and SD stages.

Figure 7. The b-values for the final six trials, for the correct (green lines) and the incorrect (red lines) stimulus in the relevant dimension, are shown for each group (open symbols, control; closed symbols, lesion). It should be noted that for those rats who completed a stage in fewer than 12 trials, there is overlap between the data in Figure 6 (the initial six trials) and the data shown here. As for the initial six trials, there was a general tendency for the rats with lesions to have lower b-values associated with the rewarded stimulus. The effect of group interacted with trial but not with stage.

Figure 8. Trial by trial b-values in each of the seven stages for the rewarded stimulus (green); non-rewarded stimulus (red); either of the stimuli in the irrelevant dimension (blue); any one of the four spatial hypotheses (grey). The position of the grey dot indicates whether the rat dug in the first bowl (dot at the bottom) or explored both before digging (dot at the top). Upper panel is 06/182 (Control) and lower panel is 09/134 (lesioned). Patterns at the SD, CD, Rev1, ID and Rev2 are very similar for the two rats: the lesioned rat 'guesses' the first trial of both the SD and ID incorrectly and makes an additional incorrect response at Rev1 and Rev2. Having 'guessed' incorrectly at the SD, the lesioned rat responses initially match a spatial "win-shift/loose stay" pattern, before six correct responses complete the stage. In Rev2, both rats show spatial patterns (alternation for 06/182 and perseveration for 09/134) and the irrelevant

dimension hypothesis is also supported in both rats' responding. The stark difference is seen at the ED stage, where the performance of 06/182 is like the ID but 09/134 shows patterns that support spatial hypotheses (the distinct peaks are due to different hypotheses: first perseveration, then win-stay, followed by alternation and finally win-stay again). Most striking is the fact that the rat first explores the second bowl on the 21st trial after a correct dig on the 20th which began the run of six consecutively correct to criterion. While 06/182 performs Rev3 as rapidly as Rev1, 09/134 perseverates for a few trials and then adopts a pattern supporting the spatial alternation hypothesis.

Figure 9. Histograms of the distributions of dimensional selectivity index (DSI) for each group and each reversal. DSIs greater than 0.7 indicate that the responding to the correct and incorrect stimulus had a high inverse correlation, suggesting that as responding to one declined, responding to the other increased. In this regard, there was no difference between the groups for the first two reversals, which also did not differ from each other. For the third reversal, the rats with lesions had significantly lower DSIs, reflecting the fact that there was an increase in trials in which spatial response patterns dominated.

Fig 1

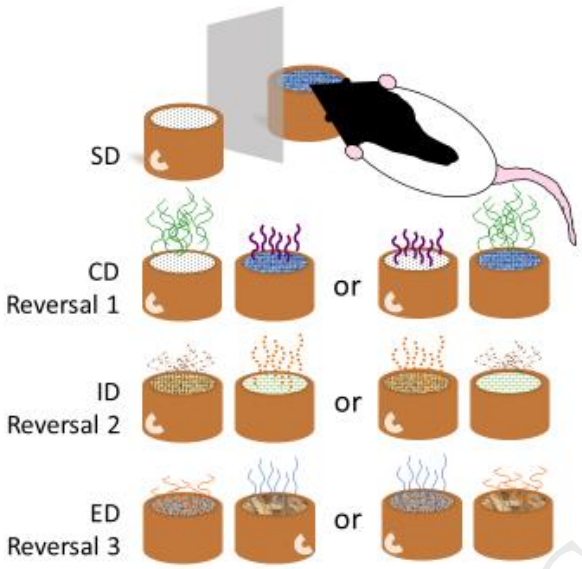


Fig 2

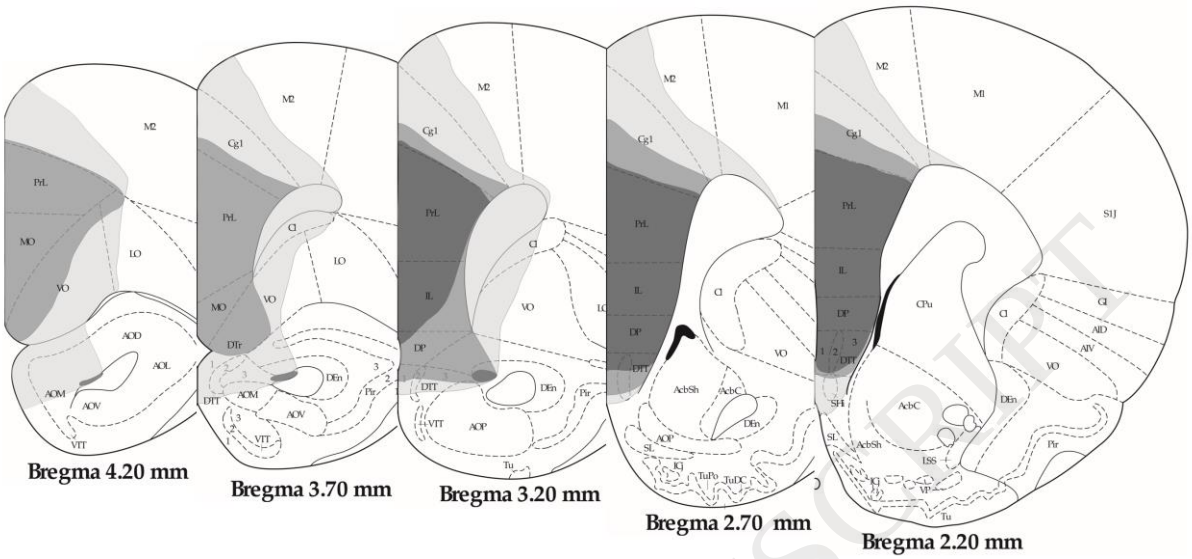


Fig 3

Behavior	Hypothesis	Sampled 2nd bowl?	Consistent with Hi?	Likelihood
Returns to same location	H_1 **Spatial** (location): Perseverative	No (0.9)	Yes	0.8
Alternates sides	H_2 **Spatial** (location): Alternate		No	0.1
Returns to location if rewarded	H_3 **Spatial** (reward): Win-Stay	Yes (0.1)	Yes	0.05
Shifts location if rewarded	H_4 **Spatial** (reward): Win-Shift		No	0.05
Digs in Medium A	H_5 **Perceptual** – Medium A	No (0.5)	Yes	0.45
Digs in Medium B	H_6 **Perceptual** – Medium B		No	0.05
Digs in Odor A	H_7 **Perceptual** – Odor A	Yes (0.5)	Yes	0.45
Digs in Odor B	H_8 **Perceptual** – Odor B		No	0.05

Fig 4

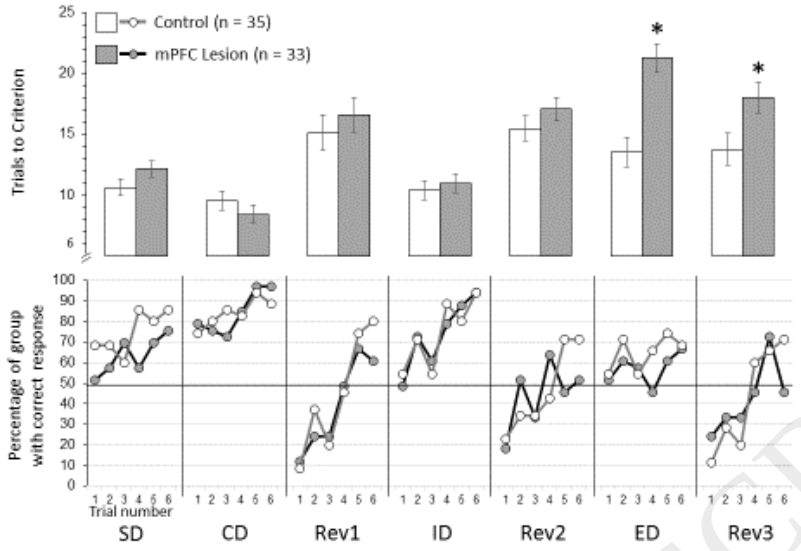


Fig 5

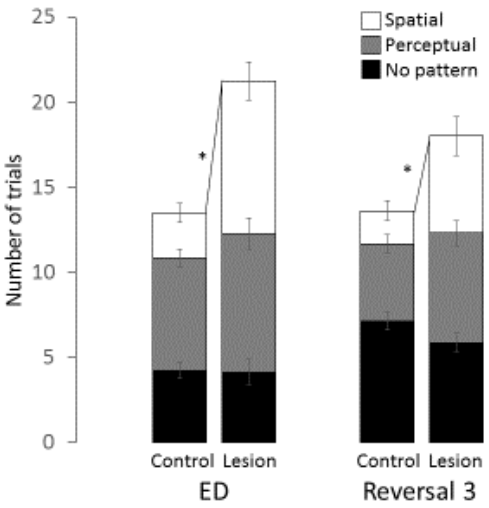


Fig 6

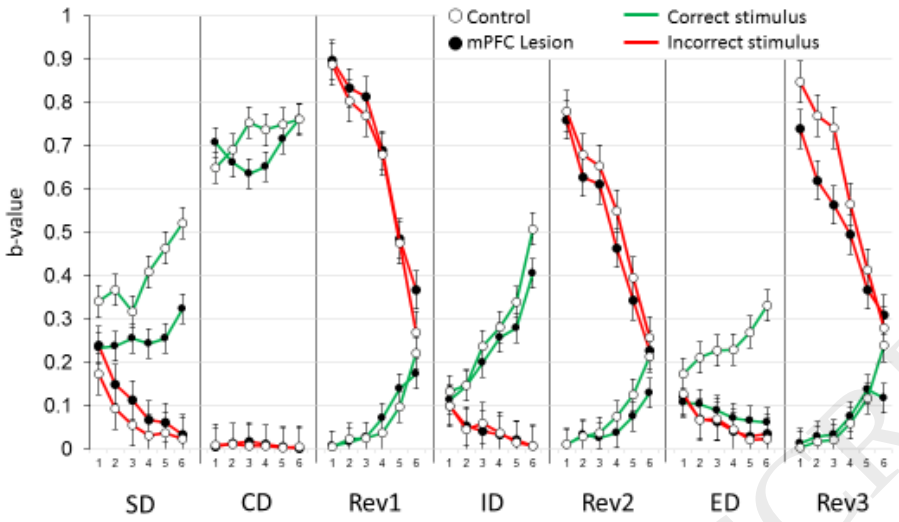


Fig 7

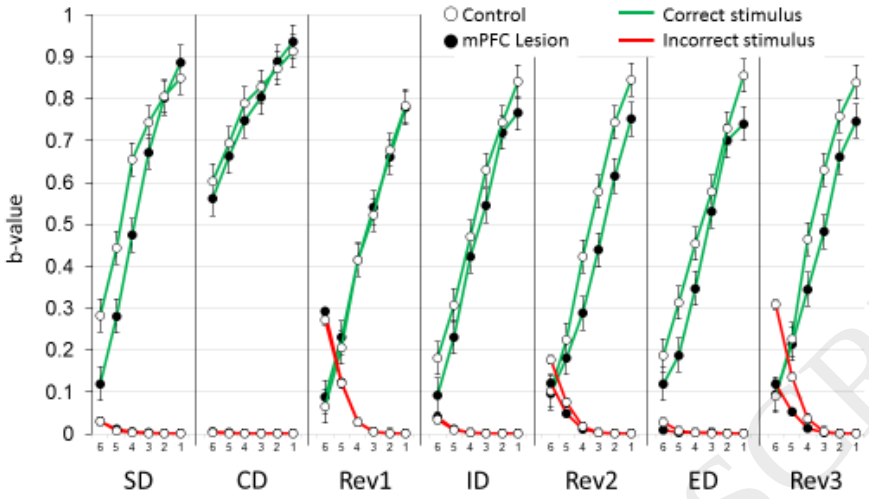


Fig 8

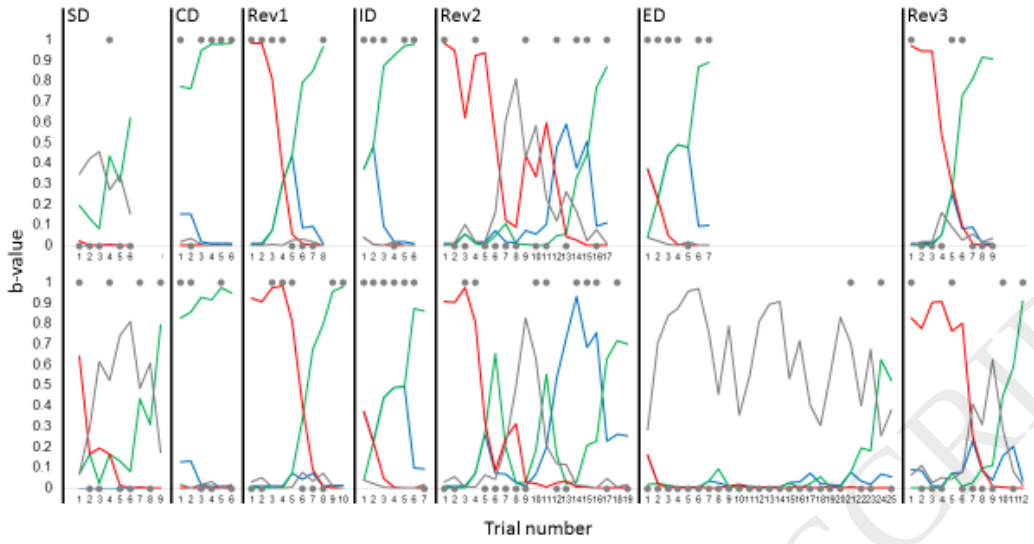


Fig 9