Reinforcement learning as an intermediate phenotype in psychosis? Deficits sensitive to illness stage but not associated with polygenic risk of schizophrenia in the general population

Authors: Montagnese M[1], Knolle F[1], Haarsma J[1], Griffin J[1], Richards A, Vertes P, Kiddle B, Fletcher PC[1,2], Jones PB[1,6], Owen M, Fonagy P[7], Bullmore ET[1,3,6], Dolan R[4,5], NSPN Consortium, Moutoussis M[4,5], Goodyer I[1,6], Murray GK[1,3].

[1]Department of Psychiatry, University of Cambridge, United Kingdom, [2]Wellcome Trust MRC Institute of Metabolic Science, Cambridge Biomedical Campus, United Kingdom, [3]Behavioural and Clinical Neuroscience Institute, University of Cambridge, United Kingdom, [4] Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London WC1B 5EH, United Kingdom, [5]Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3BG, United Kingdom, [6]Cambridgeshire and Peterborough National Health Service Foundation Trust, Cambridge CB21 5EF,United Kingdom, [7]Research Department of Clinical, Educational and Health Psychology, University College London, London WC1E 6BT, United Kingdom

Correspondence to:

Dr GK Murray, Department of Psychiatry, University of Cambridge, Douglas House, Trumpington Road 18b, Cambridge, CB2 8AH, United Kingdom (gm285@cam.ac.uk)

This work was supported by the Neuroscience in Psychiatry Network (NSPN) Consortium, a strategic award from the Wellcome Trust to the University of Cambridge and University College London (095844/Z/11/Z); by the Cambridge NIHR Biomedical Research Centre.

Abstract

Background: Schizophrenia is a complex disorder in which the causal relations between risk genes and observed clinical symptoms are not well understood and the explanatory gap is too wide to be clarified without considering an intermediary level. Thus, we aimed to test the hypothesis of a pathway from molecular polygenic influence to clinical presentation occurring via deficits in reinforcement learning.

Methods: We administered a reinforcement learning task (Go/NoGo) that measures reinforcement learning and the effect of Pavlovian bias on decision making. We modelled the behavioural data with a hierarchical Bayesian approach (hBayesDM) to decompose task performance into its underlying learning mechanisms. Study 1 included controls (*n*= 29, F|M=0.81), At Risk Mental State for psychosis (ARMS, *n*= 23, F|M=0.35) and FEP (First-episode psychosis, *n*= 26, F|M=0.18). Study 2 included healthy adolescents (*n*= 735, F|M= 1.06), 390 of whom had their polygenic risk scores for schizophrenia (PRSs) calculated.

Results: Patients with FEP (but not ARMS) showed significant impairments in overriding Pavlovian conflict, a lower learning rate and a lower sensitivity to punishment. PRSs did not

significantly predict performance on the task in the general population, which did not strongly correlate with measures of psychopathology.

Conclusions: Reinforcement learning deficits are observed in first episode psychosis but not clinical risk for psychosis, and were not predicted by molecular genetic risk for schizophrenia in healthy individuals. The study does not support the role of reinforcement learning as an intermediate phenotype in psychosis.

Keywords: Psychosis, Schizophrenia, PRS, Bayesian, Reinforcement Learning, Go/NoGo task, Computational Psychiatry

1. Introduction

Cognitive deficits are commonly observed in schizophrenia, including prominent deficits in decision making and in reinforcement learning (trial and error based learning from feedback). Reinforcement learning (RL) is a cognitive domain of interest, not only because impairments in this domain may have a direct impact on educational and occupational outcomes, but also because reinforcement learning deficits may mechanistically contribute to the pathogenesis of positive and/or negative symptoms of schizophrenia and other psychoses (Frank, 2008; Deserno et al., 2013; Murray et al., 2016). Reinforcement learning has been suggested as a candidate process for an intermediate phenotype in schizophrenia, lying on the casual path between identified risk factors and the full clinical expression of the phenotype of illness (Kasanova et al., 2018)

Despite the strong role for genetics in the aetiology of schizophrenia (Tsuang, 2000), there is only indirect evidence that reinforcement learning deficits in schizophrenia are at least partly genetic in origin. Recent evidence indicates shared genetic overlap between the genes underpinning general intellectual function and schizophrenia liability (Toulopoulou et al., 2018), and reinforcement learning correlates significantly with IQ (Chen, 2015). However, much less is known concerning the genetic basis of specific cognitive deficits in schizophrenia. There is evidence that some aspects of reward processing, which is abnormal at different stages of psychosis (Murray et al., 2008; Ermakova et al., 2018), may be an intermediate phenotype in schizophrenia. For example, relatives of people with schizophrenia show altered brain activation during reward anticipation during fMRI scans (Grimm et al., 2014). Furthermore, molecular genetic risk for schizophrenia is associated with reward related brain activation: in the IMAGEN study of about 2000 14-year-olds, Lancaster et al. (2016) found that schizophrenia polygenic risk scores were associated with striatal activation during reward anticipation. If this altered brain activation is manifest in the altered ability to learn about rewards and reward-related decision making, then we expect that reward-based reinforcement learning behaviour should also be related to polygenic risk for schizophrenia.

If reinforcement learning is an intermediate phenotype for schizophrenia, it is also to be expected that individuals who partially express the schizophrenia phenotype genetically, such as people at increased clinical risk of developing psychosis (At Risk Mental States ARMS) ,

should show a degree of deficit in reinforcement learning, but of lesser severity than individuals with the full illness phenotype. Of relevance, here is the study of schizotypal traits in the general population. It is not yet established whether schizotypal traits or clinical risk for psychosis are associated with altered reinforcement learning. Recent evidence has suggested that patients at clinical risk for psychosis show subtle subcortical prediction error abnormalities during reinforcement learning (Ermakova et al., 2018), but whether these neural deficits are associated with the behavioural deficits are not clear, as they may be insufficient to result in altered behaviour or may be compensated for by adaptations in other brain regions (Murray et al., 2010).

There is some suggestion that reinforcement learning abnormalities may not be uniform in schizophrenia but may be particularly prominent in certain patient groups. For example, reward-related reinforcement learning deficits are particularly prominent in patients with negative symptoms, consistent with the possibility that such deficits may causally contribute to the pathogenesis of such symptoms (Gold et al., 2012). Further support for such a link between reinforcement learning deficits and negative symptoms comes from computational modelling studies that tried to tease apart the different learning mechanisms involved. For example, Albrecht et al., (2016) administered a Go/NoGo reinforcement learning task (Guitart-Masip et al., 2012) to a group of chronic schizophrenia patients. Patients showed impaired Pavlovian biases, a tendency to seek a reward with action invigoration and avoid a punishment with action suppression, possibly suggesting a reduction of those mechanisms in the striatal regions that give rise to the Pavlovian biases in the first place, coupled with a disruption in communication between these striatal areas and the prefrontal cortex. The influence of Pavlovian biases on reinforcement learning have not previously been studied in first episode psychosis (FEP) or clinical risk for psychosis, and it is not known whether profiles of reinforcement learning differ across different stages of psychotic illness or in patients who are or who are not taking antipsychotic medication. The effects of Pavlovian biases on learning and decision making are of interest for both theoretical reasons, as in relation to pathogenesis of psychiatric symptoms (Moutoussis et al., 2018) and in decision-making in everyday life (Hunt et al., 2016).

If reinforcement learning is an intermediate phenotype in schizophrenia, we hypothesised to find reinforcement learning deficits in FEP patients, in ARMS individuals, and in members of

the general population with a raised molecular genetic risk for the disorder. Further, we would expect that reinforcement learning performance would relate to trait schizophrenia measures in the population. We thus studied reinforcement learning in a group of FEP patients, ARMS individuals, and healthy individuals. In several hundred healthy individuals we examined whether their performance on a reinforcement learning task related to their molecular genetic risk for schizophrenia and to their psychopathology. We hypothesised that impairments in reinforcement learning would relate to trait level manifestations of subclinical positive and negative symptoms. We combined standard measures of learning with a computational psychiatry analysis approach (Teufel & Fletcher, 2016; Redish & Gordon, 2016), as it offers the possibility of developing rigorous and testable models of behaviour that can contribute to our understanding of how abnormal neurobiological substrates become expressed in clinical phenotypes.

2. Methods and Materials

Participants

Clinical study. We recruited three groups of participants aged 17 to 35 (mean age 22.8 years): *n*= 23 participants for the ARMS group, *n*= 26 FEP patients and *n*= 29 Controls. FEP participants were recruited from the Cambridge First Episode Psychosis service, CAMEO. ARMS participants were recruited through CAMEO, through advertisements at University Counselling Services, and from existing local research databases; ARMS status was confirmed using the CAARMS interview Comprehensive Assessment of At Risk Mental States (CAARMS), as used in the EDIE-II trial (Morrison et al., 2012). Medication details can be found in *Table 6* in the Supplementary Section. Controls were recruited thorough advertisement in Cambridgeshire and through existing University of Cambridge research databases. Exclusion criteria were: current or past history of neurological disorder or trauma, currently or recently participating in a clinical trial of an investigational medical product, learning disability, or not satisfying standard MRI safety exclusion criteria, including pregnancy. The latter requirement was due to the fact that a subset of volunteers had MRI scans, reported elsewhere (Whittaker et al., 2016). Past or current treatment for a mental health problem was an exclusion criterion for controls. The project received ethical approval from the National Research Ethics Service. Written informed consent was signed by all

participants; if they were below 16 years of age, then written parental consent was also required. Further demographic information can be found in *Table 2* in the Supplementary Material.

Healthy adolescent volunteer study. *N*= 785 participants took part (mean age 18.6 years, SD= 2.96; F|M=1.06) and underwent cognitive reinforcement learning testing. Participants were recruited from General Medical Practice lists as a sampling frame as well as by direct advertisement so as to represent the UK population in this age range (Kiddle et al, 2017). Inclusion criteria were age 14 to 24 years old, able to understand written and spoken English, living in Greater London or Cambridgeshire & Peterborough, being willing and able to give informed consent for recruitment into the study cohort and consent to be re-contacted directly. Exclusion criteria were as described above for controls in the clinical study. A detailed analysis of reinforcement performance in these participants is available in Moutoussis et al., (2018), which does not address molecular genetics or schizotypal traits. Further demographic information can be found in *Table 3* in the Supplementary Material.

Psychopathology measures

The participants in the Clinical study were administered: the Comprehensive Assessment of At Risk Mental States (CAARMS) (Yung et al., 2005), providing operational criteria for identification of clinical risk for psychosis; the Mood and Feelings Questionnaire (MFQ) subset of the Young People Questionnaire (YPQ) (Costello & Angold, 1988) to measure depressive symptoms; the Positive and Negative Symptoms Scale (PANSS)(Kay et al., 1987); to measure schizotypy they were administered the 21-items Peters Delusions Inventory (PDI-21) (Peters et al., 2004) and the Schizotypal Personality Questionnaire (SPQ); IQ was measured from combining the scores of two subscales of the Wechsler Abbreviated Scale of Intelligence (WASI), namely the Vocabulary and Matrix subtests. The healthy adolescent participants were administered the following: MFQ; PLIKS (Psychosis-Like Symptoms) to measures unusual experiences, hallucinations and delusions (Zammit et al., 2008). The Schizotypal Personality Questionnaire (SPQ)(Raine, 1991) to measure schizotypy. The SPQ was later scored according to the novel subscales provided by Davies (2017); the Snaith Hamilton Pleasure Scale (SHAPS)(Snaith et al., 1995) to measure some aspects of anhedonia (higher scores reflect higher values of anhedonia); IQ was measured from the WASI, the same way as in the Clinical study.

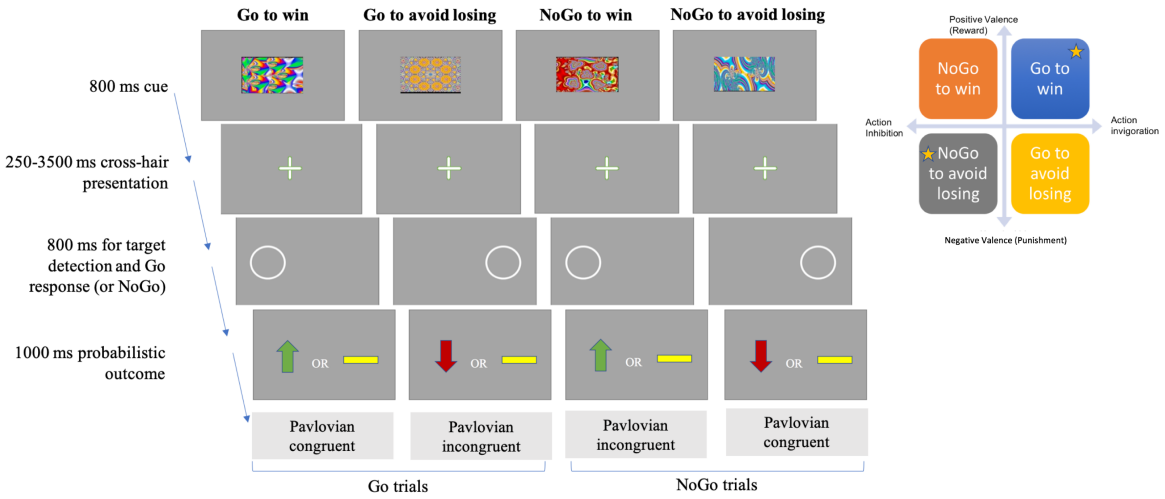

Figure 1 **-** Experimental paradigm schematic. Figure adapted from *Guitart-Masip et al., (2012) and Moutoussis et al 2018.* Top-right figure shows a graphical representation of the four conditions of the modified Go/NoGo task crossing valence (y-axis) and action (x-axis). Yellow stars mark the Pavlovian congruent conditions, while the other two are the Pavlovian incongruent ones.

Reinforcement learning task

All participants were assessed on a modified version of a traditional Go/NoGo reinforcement learning task, developed by Guitart-Masip et al., (2012) that provides several measures of reinforcement learning (*Figure 1)*. The task involved the presentation of four fractal images 36 times each, for a total of 144 trials across the 4 conditions. The order of the stimuli was random and each cue was presented for 800ms, followed by cross-hair in the middle of the screen for 250-3500ms. Then there was a target detection task showing a circle on either side of the screen for a maximum time of 800ms, during which time the participant had to make a button press response (Go) or not (NoGo). The Go response was given via pressing a keyboard button on the side on which the cue was presented (right or left), then the probabilistic outcome was shown. Possible outcomes were: a green arrow upward for wins (£0.5), a red one downwards for losses (-£0.5) and a yellow horizontal bar for neutral outcomes (£0). For the reward conditions, only positive or neutral outcomes were possible, while for the losses conditions participants could experience either a loss or a neutral outcome. Importantly, these outcomes were probabilistic on a 80:20 schedule. Overall, there were four conditions depending on the cue presented at the start of the task: two *Pavlovian congruent* conditions requiring to press the button to get a reward (*Go-to-win*) or to not press

the button to avoid losing (*NoGo-to-avoid-losing*); two *Pavlovian Incongruent* conditions requiring to either not press the button to get a reward (*NoGo-to-win*) or to press the button to avoid losing (*Go-to-avoid-losing*). Further details on the specifics of the task can be found in the Supplementary Material.

Computational modelling: hBayesDM

Behavioural performance on the Go/NoGo task was calculated by summing scores for the task conditions, and by modelling latent task variables using the hBayesDM package (hierarchical Bayesian modeling of Decision Making tasks) for R (version 0.5.0 on MacOS High Sierra version 10.13.1) developed by Ahn et al. (2017). We used this approach to generate posterior distributions of the parameters characterising task performance to improve the balance of within-subject and between-subject random effects, whilst also taking into account within-subject variability and group-level similarities (O'Callaghan et al., 2017). Full information on the details of the modelling parameters and model fitting and comparison can be found in the Supplementary Material. "Model 4" was the best model (lowest LOIC) for both cohorts of participants and included the following parameters: lapse rate (random errors), learning rate, Go bias (tendency to make a response), Pavlovian bias (tendency to make a response to stimuli associated with reward and withhold a response to stimuli associated with punishment), sensitivity to reward, sensitivity to punishment.

Polygenic risk score calculation

Participants in the healthy adolescent study participants were drawn from a larger sample of over 2000 adolescents on whom genetic data were acquired from by saliva sample (Kiddle et al., 2017). Genotyping was carried out by the Cambridge Bioresource on an Affymetrix chip array, yielding genotype at 507,968 SNPs for subjects. Quality control and imputation was performed. The parameters for retaining SNPs were: SNP missingness < 0.01 (before sample removal); SNP Hardy-Weinberg equilibrium ($P > 10^{-6}$) and minor allele frequency MAF > 0.01. Final statistical analyses were carried out on $n = 390$ participants of European ancestry for whom both adequate genotype and reinforcement learning data were available. See *Figure 5* in Supplementary Material for a detailed flowchart of excluded participants. The generation of the PRS was based on the methods described by the *International Schizophrenia Consortium* (2009). Polygenic scores were calculated for each individual using the PLINK (version 1.9) score command. Scores were created by adding up the number of risk alleles for each SNP, i.e. single nucleotide polymorphism, which took the value of 0,1, or

2 and weighted by the logarithm of its odds ratio for schizophrenia from the results reported in Pardinas et al., (2018): the meta-analysis of the CLOZ-UK sample and the Psychiatric Genomics Consortium PGC2 schizophrenia dataset (Jones et al., 2016). The scores used were generated from a list of SNPs with a GWAS training-set P<.05 threshold, as this is the threshold that has been suggested to capture maximal schizophrenia liability (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014; Pardiñas et al. 2018).

Statistical analyses

In the clinical study, group differences on task performance (behavioural and modelled) were examined by one-way analysis of variance (ANOVAs), and Spearman Rank Order correlation coefficients were used investigate the relationships between task performance and clinical measures at each group level. Despite the group differences in IQ in the clinical study, since matching for education and IQ could yield a non-representative sample of patients, and given that both the participants' own level of education and their maternal levels of education were not significantly different from controls, we did not match ARMS and FEP for IQ and, like Albrecht et al., (2016), we did not use IQ as a covariate for the statistical analyses carried out.

In the healthy adolescent study, the relationships between task performance (behavioural and modelled) and clinical measures were examined by Spearman Rank Order correlation coefficients (n= 735). Standard multiple regression analysis was first used to test whether PRS at P-threshold 0.05 predicted learning rate as measured by the computational model (chosen as the main outcome variable given the robust evidence in the literature showing learning deficits in patients with schizophrenia). Covariates included age, sex and the first five primary component analysis factors for ancestry. N= 5 participants were excluded as outliers, with a final sample of n= 390. To test if the PRS scores predicted the other aspects of task performance, standard multiple regression analyses were then run for each of the other cognitive variables of interest. False Discovery Rate (Benjamini-Hochberg) correction was applied to control for the expected proportion of falsely rejected hypotheses and to gain power (Benjamini & Hochberg, 1995). Further, Bayesian linear regressions were also performed in JASP to compare the likelihood of the task performance data under models with, versus without, schizophrenia polygenic risk score.

3. Results

Clinical study

All groups showed the classic pattern of better performance in the Pavlovian congruent conditions. There were significant differences between groups in performance of the task, with FEP overall performing worse than the other groups in several measures of performance.

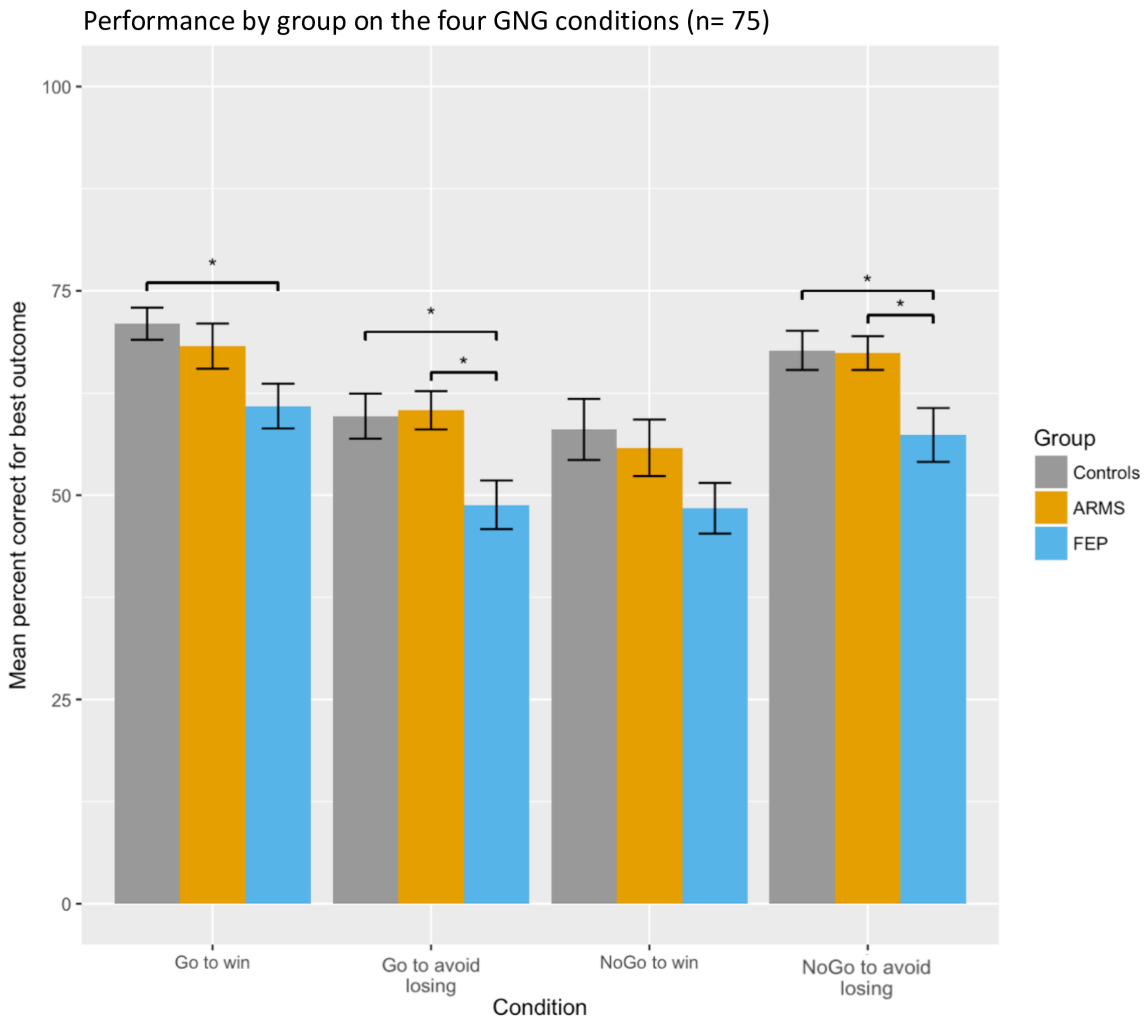

Figure 2 **-** Group differences in overall performance (percent for best outcome) on the four GNG (go/no-go) conditions. Controls n=29, ARMS (At-Risk for Mental Health) n=23 and FEP (First-episode psychosis) n=26. Error bars indicate standard error of the mean. Stars indicate significant t-test group differences at $p<0.05$ after ANOVA testing.

In terms of overall performance (percent for best outcome) on the four GNG conditions, all groups showed better performance in the Pavlovian congruent conditions compared to the Pavlovian incongruent ones. Specifically, when looking at group differences in performance of each condition, FEP performed significantly worse than controls and ARMS in the Punishment conditions (Go-to-avoid losing and NoGo-to-avoid-losing) and also significantly worse than controls only on the easier Go-to-win condition. See *Figure 2* and for the descriptive statistics *Table 4* in the Supplementary Material.

When then looking at the latent variables of performance, we found group differences across the six modelled parameters, largely driven by FEP versus control results. See *Figure 3* below and *Table 5* in the Supplementary Material for the results of group-comparisons and the statistics of each parameter per group.

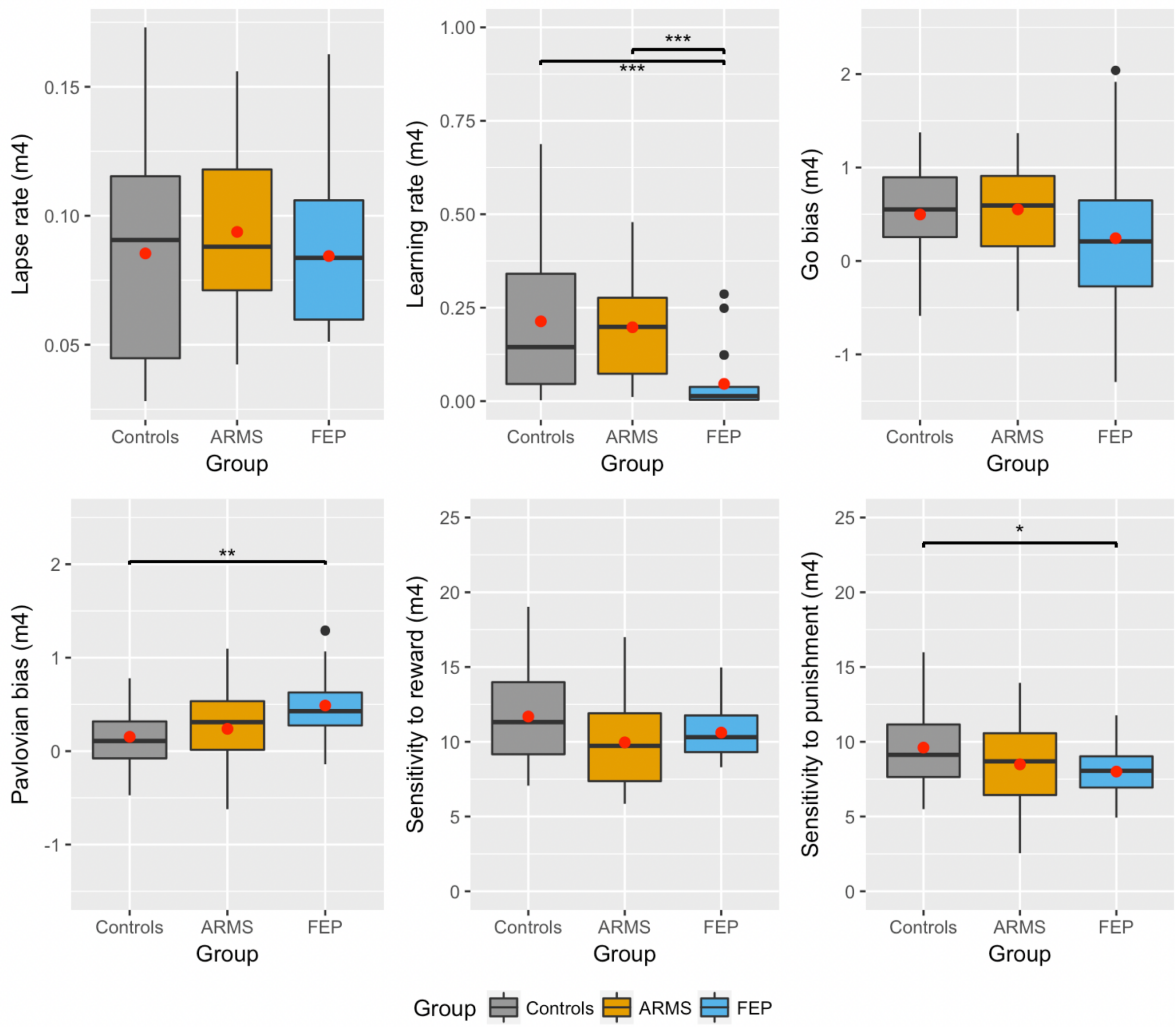

Figure 3 - Group differences in the modelled parameters). For Go Bias and Pavlovian Bias, values >0 indicate the presence of such bias, those <0 indicate the opposite. Horizontal back bar = Median; mean = red circle. Whiskers indicate the interquartile range. (*p<.05, **p<.01, ***<.001).

To further explore the possible effect of antipsychotic medication on our results, we subdivided the FEP group into two different sub-groups: one of FEP individuals who did not take antipsychotics (FEP- *n*= 11) and one with those taking antipsychotics, (FEP+ *n*= 15). The results remained largely the same. See Supplementary Material for more details. FEP+ had a higher sensitivity to punishment compared to the FEP- (1.800, 95% CI [-3.341, .258], p= .019), and this latter group also showed a significantly higher Pavlovian bias compared to Controls (0.363, 95% CI [.0001, .727], p= .05).

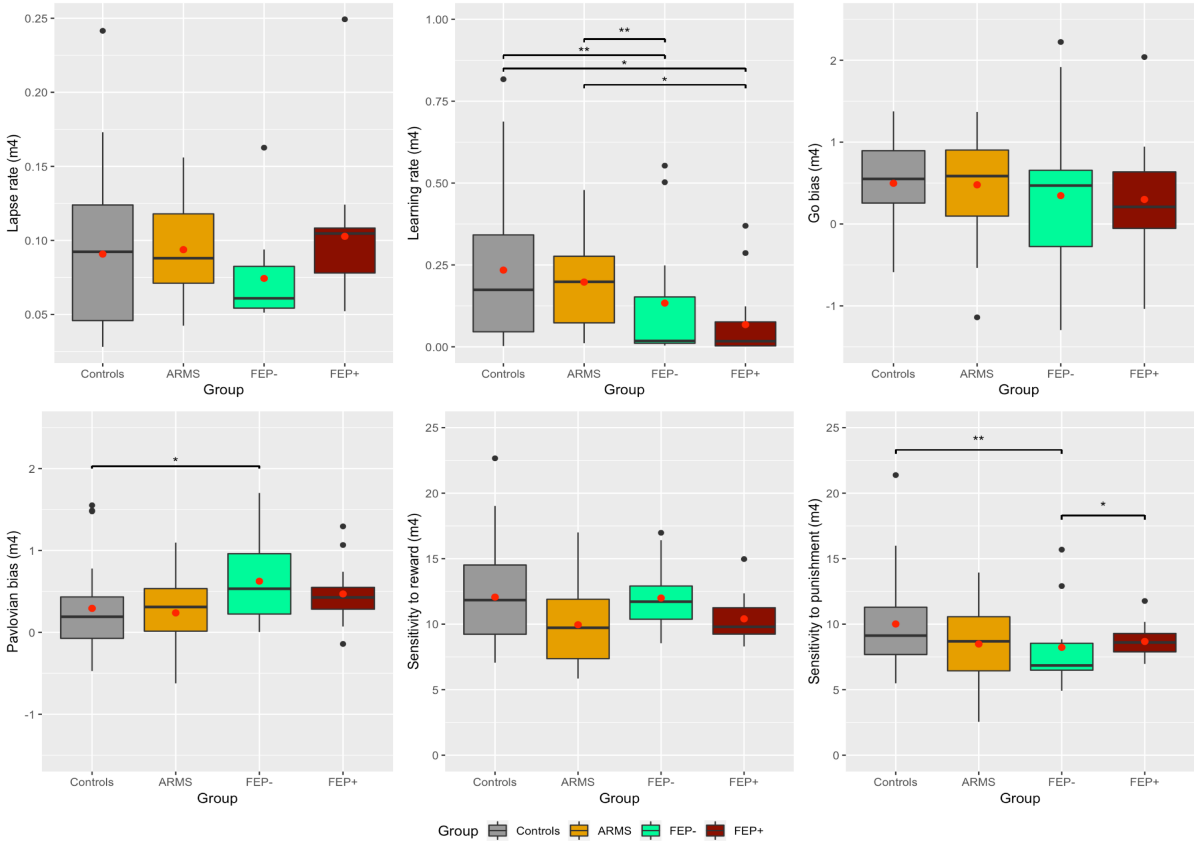

Figure 4 **-** Group differences in the modelled parameters (m4 in the hBayesDM terminology) after subdividing the first episode psychosis (FEP) group into FEP- = those not taking antipsychotics, and FEP+ = taking antipsychotics. For Go Bias and Pavlovian Bias, values >0 indicate the presence of such bias, those <0 indicate the opposite. Horizontal back bar = Median; mean = red circle. Whiskers indicate the interquartile range. (*p<.05, **p<.01, ***<.001).

Results from the Spearman correlational analyses investigating possible relationships between task performance and clinical measures for each group can be found in *Figure 6* in the Supplementary Material.

Healthy adolescent study

The pattern of performance in the healthy adolescent study is reported in detail by Moutoussis et al., (2018). In brief, there were, as expected, significant differences in performance across conditions, with better performance on the Pavlovian congruent conditions compared to the Pavlovian incongruent ones, and similar patterns for the learning curves. The Spearman correlational analyses on the Healthy Adolescent group showed a moderate negative correlation between the modelled parameters of Pavlovian bias and that of learning rate. Moutoussis et al., (2018) reported that there were no significant associations between task indices and mood. Our behavioural results (*Figure 7* in Supplementary Material) indicate weak positive correlations between the Go bias parameter and SPQ tot ($r = .13$, $p = .01$), as well as with two SPQ subscales tapping on social anxiety and eccentricity ($r = .13$, $p = .01$ and $r = 0.10$ $p = .04$). The SPQ subscale reflecting anomalous experiences and beliefs was weakly negatively correlated with the sensitivity to reward in the task ($r = -.11$, $p = .03$). The sensitivity to punishment was weakly negatively associated with the SPQ subscale of paranoid ideation ($r = -.15$, $p < .001$) and with the PLIKS ($r = -.11$, $p = .03$).

The results from the standard multiple regression analysis between PRS at P-threshold 0.05 and the modelled parameter of learning rate (with age, sex, first five primary component analysis factors for ancestry as covariates) was not statistically significant:
$R^2 = .004$, $F(8, 381) = .177$, $p = .994$, adjusted $R^2 = -.017$, Unstandardized B Coefficient = -.001 (Standard error = .008, t-value = -.109, $p = .913$). Standardized Beta coefficient (β)= -.006. See *Figure 8* in the Supplementary Material. Results for the other main cognitive variables of interest are summarised in *Table 1* below in ascending order of adjusted significance p-value. Overall, after corrections, no significant results were found.

Table 1- Summary of the results from the standard multiple regressions carried out with PRS P-threshold of 0.05 before and after False discovery rate (FDR) correction. P-values are shown in order of ascending order of adjusted significance. (*p<.05, **p<.01, ***<.001). SE = Standard error of the unstandardized coefficient.

Cognitive variable of interest (IV)	Unstandardized regression coefficient and SE	Standardized coefficient (β)	t-value	Significance of regression coefficient for the PRS Score and the cognitive variable of interest	Significance after False discovery rate correction
Sensitivity to reward	1.167(.662)	.093	1.761	.079	.348
NoGo-to-avoid-losing %	-1.027(.652)	-.082	-1.575	.116	.348
Sensitivity to punishment	-.755(.473)	-.083	-1.598	.111	.348
Lapse rate	.001(.003)	.027	.525	.600	.927
NoGo-to-win %	.768(1.044)	.038	.735	.463	.927
Go-to-win %	.118(.559)	.011	.211	.833	.927
Go-to-avoid-losing %	-.251(.723)	-.018	-.347	.729	.927
Go bias	-.005(.031)	-.009	-.172	.864	.927
Pavlovian bias	-.002(.023)	-.005	-.092	.927	.927

We also run Bayesian linear regression analyses, comparing a model with PRS to a null model including age, gender and the first five PCA components of ancestry as covariates. Results can be found in *Table 8* in the Supplementary Material. The null model with the covariates out-predicted the model that contained the main predictor of interest for all task-related variables. The only exceptions regarded the two modelled parameters tapping on sensitivity to rewards (BF_{10} = 1.331) and to punishment (BF_{10} = 1.011). Nevertheless, a Bayes Factor between 1 and 3 is considered as providing only weak and inconclusive evidence for the support of H1 over the null model (Lee & Wagenmakers, 2013; Wagenmakers et al., 2017) and the results converge with what found in the standard multiple regression analyses.

4. Discussion and Conclusions

In the Clinical study, overall, all groups showed better performance in the Pavlovian congruent conditions compared to the Pavlovian incongruent ones. We found group differences in behavioural and modelled performance on the task, with FEP performing worse than the other two groups. Both ARMS and FEP showed a decrease in behavioural performance across all conditions of the task compared to controls, but the differences were only significant for the FEP group. Further to this, and contrary to what was expected, FEP performed relatively better on the Pavlovian congruent conditions compared to the Pavlovian incongruent ones, and even more so when having to make an action to get a reward than to avoid a punishment, thus suggesting preserved action (Go) reward-related learning.

There were also significant group differences in the modelled parameters. Learning rate was lower in FEP compared to controls and to ARMS (who showed a similar learning rate to controls) and FEP had a significantly higher Pavlovian bias than controls. Finally, although the sensitivity to punishment was intact for ARMS, it was significantly reduced in FEP compared to the other groups, which is seemingly at odds with what was hypothesised and expected from previous literature in chronic schizophrenia patients (Gold et al., 2008). There were no significant differences in the sensitivity to punishment nor in the other modelled parameters.

The finding of a higher Pavlovian bias in first episode psychosis patients compared to controls is in contrast with the findings from Albrecht et al., (2016) in chronic illness. This might be attributable to the progression of the disease which, alongside an extensive use of antipsychotics (Scherer et al., 2004), is linked to the worsening of deficits in gradual reinforcement learning, the neural substrate of which is thought to involve the basal ganglia, and specifically the striatum, the same areas thought to give rise to the Pavlovian biases. In turn, this might have the effect of weakening the Pavlovian biases and result in the pattern observed in Albrecht's study. To investigate further whether disease stage and antipsychotics can have such effects on Pavlovian biases, longitudinal follow up of FEP patients is necessary.

Our results show clear group effects and deficits in reinforcement learning in patients, although these are different from the deficits found in Albrecht et al., (2016). Such discrepancies could be due to the different stages of schizophrenia among the patients. In fact, in Albrecht et al., (2016), patients were at a chronic stage of the disorder and much older (mean age= 37.7 years) than those in the current study. In the current study, we looked at individuals suffering from early psychotic episodes (FEP) and at those who were at-risk (ARMS) for developing schizophrenia, respectively being 24.6 and 21.2 years old on average. Furthermore, the patients in Albrecht's study presented more severe negative symptoms, thus they were potentially a different subgroup of schizophrenia patients compared to ours.

In the Healthy Adolescent study, the pattern of overall performance on the reinforcement learning task is the same as that of controls from the Clinical study, thus showing that, in the general population, individuals learn the Pavlovian congruent conditions more easily and have more difficulties with the incongruent ones. Although similar results had been shown in previous experiments, the main contribution the current findings is that they confirm the influences of Pavlovian biases on reinforcement learning in a bigger and younger sample compared to those in which this task had been used so far. In doing so, the current study strengthens the confidence that the observed pattern of results is representative of how an average healthy individual performs the task.

When correlating task performance and clinical measures of psychopathology, we found some evidence of weak associations between task performance and schizotypy. The Go bias parameter of performance was positively associated with the total score on the SPQ scale measuring schizotypy, as well as with two of the subscales tapping on social anxiety and on eccentricity; this is in contrast to the clinical results, where Go bias was reduced in patients. The SPQ subscale reflecting anomalous experiences and beliefs was negatively associated with the sensitivity to reward – this domain did not differentiate patients and controls in the clinical study. Sensitivity to punishment was negatively correlated with the SPQ subscale of paranoid ideation, and with the PLIKS, and was reduced in first episode psychosis patients; taken together might suggest a link between impaired punishment-related learning and delusional thinking in clinical psychosis and the healthy population.

Our results further show that PRSs for schizophrenia in the general population do not predict performance on this specific reinforcement learning task. There are multiple possible explanations of this lack of significance, which cannot be disentangled in the current study. The first possible explanation is that the PRS for schizophrenia does not specifically bear on the cognitive domain of reinforcement learning, which could be more associated with illness itself rather than illness risk; this explanation would align with the clinical study findings where we did not find significantly impaired performance in the clinical risk (ARMS) group. The second explanation is that the regression analyses were underpowered to detect any small polygenic risk effect sizes present in this sample and/or the GNG task might not have capture sufficient individual variability in performance (we note performance did not significantly relate to any measured psychological traits in the healthy adolescent study). We did not record fMRI responses during reinforcement learning which were shown to associate with schizophrenia PRS in a recent study (Lancaster et al., 2019). We also conducted post-hoc power calculations in order to inform future studies and examine the power of the current study to detect genetic effects on reinforcement learning. Post-hoc power calculations (Soper, 2018) suggested that a sample of 390 individuals in the PRS analysis, with learning rate as the main predictor, had 0.43 power to detect an association, and therefore the analysis might have been underpowered to find any significant results. For a 0.80 power with the same observed effect size, a minimum sample of 959 individuals would have been needed to demonstrate a significant effect. Nevertheless, for the majority of cognitive outcomes measures, Bayesian analysis indicated the data was slightly more likely under a model without schizophrenia polygenic risk score than one including it. Finally, the sample in the Healthy Adolescent study consisted of individuals who were partly recruited on the basis of their good health; it is possible that this lack in mental health variance might have reduced our ability to detect relationships between task performance and other traits.

The study has several limitations. Firstly, in the Clinical study the groups differed significantly on age, and this could possibly be problematic when looking at group differences, as some studies point at age- related effects on reinforcement learning performance (Samanez-Larkin & Knutson, 2015; Radulescu, Daniel and Niv, 2016); however, the group differences in age were only slight, and the behavioural performance differences remained intact when controlling for age. We did not demonstrate significant differences between ARMS and controls. This could be party linked to the conservative approach we used in the modelling, namely fitting the models to all members of the clinical

study as if they were drawn from a single group (with a single mean and variance), rather than assuming they were drawn from separate populations. Although this approach has the benefit of minimising potential false positives, it might have reduced the between-group differences and therefore impacted the overall sensitivity of the study. However, the results in the analysis of the modelled latent parameters were largely consistent with those in the modelled observed performance measures. Finally, we acknowledge the possible influence of severe traumatic stress experiences, which was linked to increased Pavlovian biases in a previous study (Ousdal et al., 2018). In fact, one cannot exclude the possibility that one of the driving mechanisms leading to the observed increase in Pavlovian biases in FEP might be linked to the stress of having experienced a psychotic episode, thus opening interesting avenues for further research.

Overall, the current work makes some important contributions to the field of reinforcement learning in schizophrenia. Firstly, we show that there are specific reinforcement learning deficits in psychotic illness and that such deficits are sensitive to illness stage, being present in frank psychosis but not in At Risk Mental States. Secondly, we show that there is no clear association between these reinforcement learning domains identified as deficient in psychosis and psychopathology in the general population. Lastly, we found no large effects of either clinical risk for psychosis or molecular polygenic risk for schizophrenia in reinforcement learning, with the power calculations indicating that a bigger sample would be required for definitive results; the results do not support reinforcement learning as an intermediate phenotype for schizophrenia.

Acknowledgements & Funding

The authors are grateful to the volunteers who took part in these studies, as well as all the members of staff involved in the recruitment process. This work was supported by the Neuroscience in Psychiatry Network (NSPN) Consortium, a strategic award from the Wellcome Trust to the University of Cambridge and University College London (095844/Z/11/Z); by the Cambridge NIHR Biomedical Research Centre.

Conflicts of interest

ETB is employed 50% by GSK. All other authors declare no conflicts of interest. PBJ was a member of scientific advisory boards for Janssen, Ricordati and Lundbeck.

Author roles

MMontagnese: conceptualization, methodology, formal analysis, writing (original draft preparation, review and editing); FK: conceptualization, methodology, formal analysis, supervision, writing (original draft preparation, review and editing); JH, JG: conceptualization, investigation, methodology, writing (review and editing); AR, PV, BK,: methodology, resources; writing; PCF conceptualization, methodology, writing (review and editing); PBJ, PF, ETB, RD, IG: conceptualization, project administration, funding acquisition, supervision, writing (review and editing); MO: resources; writing; MMoutoussis: conceptualization, methodology, supervision, writing (review and editing); NSPN Consortium: conceptualization, methodology, project administration. GKM, conceptualization, project administration, methodology, supervision, writing (original draft preparation, review and editing).

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Ahn, W., Haines, N. and Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, 1, pp.24-57.

Ahn, W., Haines, N. and Zhang, L. (2018). **hBayesDM Reference Manual.** [online] Cran.r-project.org. (https://cran.r-project.org/web/packages/hBayesDM/hBayesDM.pdf) [Accessed 1 Mar. 2018].

Albrecht, M., Waltz, J., Cavanagh, J., Frank, M. and Gold, J. (2016). Reduction of Pavlovian Bias in Schizophrenia: Enhanced Effects in Clozapine-Administered Patients. *PLOS ONE*, 11(4), p.e0152781.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), pp.289-300.

Chen, C. (2015). Intelligence moderates reinforcement learning: a mini-review of the neural evidence. *Journal of Neurophysiology*, 113(10), pp.3459-3461.

Collins, A., Brown, J., Gold, J., Waltz, J. and Frank, M. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *The Journal of Neuroscience*, 34(41), pp.13747-13756.

Costello, E. and Angold, A. (1988). Scales to Assess Child and Adolescent Depression: Checklists, Screens, and Nets. *Journal of the American Academy of Child & Adolescent Psychiatry*, 27(6), pp.726-737.

Davies, D. (2017). *Psychotic experiences beyond psychotic disorders in young people: from measurement to computational mechanisms*. PhD. University of Cambridge.

Deserno, L., Boehme, R., Heinz, A. and Schlagenhauf, F. (2013). Reinforcement Learning and Dopamine in Schizophrenia: Dimensions of Symptoms or Specific Features of a Disease Group?. *Frontiers in Psychiatry*, 4.

Ermakova, A., Knolle, F., Justicia, A., Bullmore, E., Jones, P., Robbins, T., Fletcher, P. and Murray, G. (2018). Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis. *Neuropsychopharmacology*, 43(8), pp.1691-1699.

Ermakova, A., Knolle, F., Justicia, A., Bullmore, E., Jones, P., Robbins, T., Fletcher, P. and Murray, G. (2018). Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis. *Neuropsychopharmacology*, 43(8), pp.1691-1699.

Frank, M. (2008). Schizophrenia: A Computational Reinforcement Learning Perspective. *Schizophrenia Bulletin*, 34(6), pp.1008-1011.

Gold, J. (2012). Negative Symptoms and the Failure to Represent the Expected Reward Value of Actions. *Archives of General Psychiatry*, 69(2), pp.129-38.

Gold, J., Waltz, J., Prentice, K., Morris, S. and Heerey, E. (2008). Reward Processing in Schizophrenia: A Deficit in the Representation of Value. *Schizophrenia Bulletin*, 34(5), pp.835-847.

Grimm, O., Heinz, A., Walter, H., Kirsch, P., Erk, S., Haddad, L., Plichta, M., Romanczuk-Seiferth, N., Pöhland, L., Mohnke, S., Mühleisen, T., Mattheisen, M., Witt, S., Schäfer, A., Cichon, S., Nöthen, M., Rietschel, M., Tost, H. and Meyer-Lindenberg, A. (2014). Striatal Response to Reward Anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. *JAMA Psychiatry*, 71(5), pp.531-9.

Guitart-Masip, M., Huys, Q., Fuentemilla, L., Dayan, P., Duzel, E. and Dolan, R. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, 62(1), pp.154-166.

Hunt, L., Rutledge, R., Malalasekera, W., Kennerley, S. and Dolan, R. (2016). Approach-Induced Biases in Human Information Sampling. *PLOS Biology*, 14(11), p.e2000638.

International Schizophrenia Consortium, Purcell, Wray, Stone, Visscher, O'Donovan, Sullivan and Sklar (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460(7256), pp.748-52.

Jones, H., Stergiakouli, E., Tansey, K., Hubbard, L., Heron, J., Cannon, M., Holmans, P., Lewis, G., Linden, D., Jones, P., Davey Smith, G., O'Donovan, M., Owen, M., Walters, J. and Zammit, S. (2016). Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry*, 73(3), p.221.

Kasanova, Z., Ceccarini, J., Frank, M., van Amelsvoort, T., Booij, J., van Duin, E., Steinhart, H., Vaessen, T., Heinzel, A., Mottaghy, F. and Myin-Germeys, I. (2018). Intact striatal dopaminergic modulation of reward learning and daily-life reward-oriented behavior in first-degree relatives of individuals with psychotic disorder. *Psychological Medicine*, 48(11), pp.1909-1914.

Kay, S., Fiszbein, A. and Opler, L. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin*, 13(2), pp.261-276.

Kiddle, B., Inkster, B., Prabhu, G., Moutoussis, M., Whitaker, K., Bullmore, E., Dolan, R., Fonagy, P., Goodyer, I. and Jones, P. (2017). Cohort Profile: The NSPN 2400 Cohort: a developmental sample supporting the Wellcome Trust NeuroScience in Psychiatry Network. *International Journal of Epidemiology*, 47(1), pp.18-19g.

Lancaster, T., Linden, D., Tansey, K., Banaschewski, T., Bokde, A., Bromberg, U., Büchel, C., Cattrell, A., Conrod, P., Flor, H., Frouin, V., Gallinat, J., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Martinot, J., Paillère Martinot, M., Artiges, E., Lemaitre, H., Nees, F., Orfanos, D., Paus, T., Poustka, L., Smolka, M., Vetter, N., Jurk, S., Mennigen, E., Walter, H., Whelan, R. and Schumann, G. (2016). Polygenic Risk of Psychosis and Ventral Striatal Activation During Reward Processing in Healthy Adolescents. *JAMA Psychiatry*, 73(8), pp.852-61.

Lancaster, T., Dimitriadis, S., Tansey, K., Perry, G., Ihssen, N., Jones, D., Singh, K., Holmans, P., Pocklington, A., Davey Smith, G., Zammit, S., Hall, J., O'Donovan, M., Owen, M. and Linden, D. (2019). Structural and Functional Neuroimaging of Polygenic Risk for Schizophrenia: A Recall-by-Genotype–Based Approach. *Schizophrenia Bulletin*, 45(2), pp.405-414.

Lee, M. and Wagenmakers, E. (2013). *Bayesian cognitive modeling*. 1st ed. Cambridge University Press.

Morrison, A., French, P., Stewart, S., Birchwood, M., Fowler, D., Gumley, A., Jones, P., Bentall, R., Lewis, S., Murray, G., Patterson, P., Brunet, K., Conroy, J., Parker, S., Reilly, T., Byrne, R., Davies, L. and Dunn, G. (2012). Early detection and intervention evaluation for people at risk of psychosis: multisite randomised controlled trial. *BMJ*, 344(apr05 1), pp.e2233-e2233.

Moutoussis, M., Bullmore, E., Goodyer, I., Fonagy, P., Jones, P., Dolan, R. and Dayan, P. (2018). Change, stability, and instability in the Pavlovian guidance of behaviour from adolescence to young adulthood. *PLOS Computational Biology*, 14(12), p.e1006679.

Murray, G., Corlett, P. and Fletcher, P. (2010). The Neural Underpinnings of Associative Learning in Health and Psychosis: How Can Performance Be Preserved When Brain Responses Are Abnormal?. *Schizophrenia Bulletin*, 36(3), pp.465-471.

Murray, G., Corlett, P., Clark, L., Pessiglione, M., Blackwell, A., Honey, G., Jones, P., Bullmore, E., Robbins, T. and Fletcher, P. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, 13(3), pp.267-276.

Murray, G., Tudor-Sfetea, C. and Fletcher, P. (2016). Can Models of Reinforcement Learning Help Us to Understand Symptoms of Schizophrenia?. In: J. Dreher and L. Tremblay, ed., *Decision Neuroscience: An Integrative Perspective*, 1st ed. Academic Press, pp.261-275.

O'Callaghan, C., Hall, J., Tomassini, A., Muller, A., Walpola, I., Moustafa, A., Shine, J. and Lewis, S. (2017). Visual Hallucinations Are Characterized by Impaired Sensory Evidence Accumulation: Insights From Hierarchical Drift Diffusion Modeling in Parkinson's Disease. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 2(8), pp.680-688.

Ousdal, O., Huys, Q., Milde, A., Craven, A., Ersland, L., Endestad, T., Melinder, A., Hugdahl, K. and Dolan, R. (2018). The impact of traumatic stress on Pavlovian biases. *Psychological Medicine*, 48(02), pp.327-336.

Pardiñas, A., Holmans, P., Pocklington, A., Escott-Price, V., Ripke, S., Carrera, N., Legge, S., Bishop, S., Cameron, D., Hamshere, M., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J., McCarroll, S., Baune, B., Breen, G., Byrne, E., Dannlowski, U., Eley, T., Hayward, C., Martin, N., McIntosh, A., Plomin, R., Porteous, D., Wray, N., Caballero, A., Geschwind, D., Huckins, L., Ruderfer, D., Santiago, E., Sklar, P., Stahl, E., Won, H., Agerbo, E., Als, T., Andreassen, O., Bækvad-Hansen, M., Mortensen, P., Pedersen, C., Børglum, A., Bybjerg-Grauholm, J., Djurovic, S., Durmishi, N., Pedersen, M., Golimbet, V., Grove, J., Hougaard, D., Mattheisen, M., Molden, E., Mors, O., Nordentoft, M., Pejovic-Milovancevic, M., Sigurdsson, E., Silagadze, T., Hansen, C., Stefansson, K., Stefansson, H., Steinberg, S., Tosato, S., Werge, T., Collier, D., Rujescu, D., Kirov, G., Owen, M., O'Donovan, M. and Walters,

J. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 50(3), pp.381-389.

Peters, E., Joseph, S., Day, S. and Garety, P. (2004). Measuring Delusional Ideation: The 21-Item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin*, 30(4), pp.1005-1022.

Radulescu, A., Daniel, R. and Niv, Y. (2016). The effects of aging on the interaction between reinforcement learning and attention. *Psychology and Aging*, 31(7), pp.747-757.

Raine, A. (1991). The SPQ: A Scale for the Assessment of Schizotypal Personality Based on DSM-III-R Criteria. *Schizophrenia Bulletin*, 17(4), pp.555-564.

Redish, A. and Gordon, J. (2016). *Computational psychiatry*. 1st ed. MIT Press.

Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421–427 (2014).

Samanez-Larkin, G. and Knutson, B. (2015). Decision making in the ageing brain: changes in affective and motivational circuits. *Nature Reviews Neuroscience*, 16(5), pp.278-289.

Snaith, R., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D. and Trigwell, P. (1995). A Scale for the Assessment of Hedonic Tone the Snaith–Hamilton Pleasure Scale. *British Journal of Psychiatry*, 167(01), pp.99-103.

Soper, D. (2018). Free Post-hoc Statistical Power Calculator for Multiple Regression - Free Statistics Calculators.[online] (https://www.danielsoper.com/statcalc/calculator.aspx?id=9) [Accessed 1 Mar. 2018].

Teufel, C. and Fletcher, P. (2016). The promises and pitfalls of applying computational models to neurological and psychiatric disorders. *Brain*, 139(10), pp.2600-2608.

Toulopoulou, T., Zhang, X., Cherny, S., Dickinson, D., Berman, K., Straub, R., Sham, P. and Weinberger, D. (2018). Polygenic risk score increases schizophrenia liability through cognition-relevant pathways. *Brain*, 142(2), pp.471-485.

Tsuang, M. (2000). Schizophrenia: genes and environment. *Biological Psychiatry*, 47(3), pp.210-220.

Wagenmakers, E., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., de Jong, T., van den Bergh, D., Sarafoglou, A., Steingroever, H., Derks, K., Rouder, J. and Morey, R. (2017). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25(1), pp.58-76.

Whitaker, K., Vértes, P., Romero-Garcia, R., Váša, F., Moutoussis, M., Prabhu, G., Weiskopf, N., Callaghan, M., Wagstyl, K., Rittman, T., Tait, R., Ooi, C., Suckling, J., Inkster, B., Fonagy, P., Dolan, R., Jones, P., Goodyer, I. and Bullmore, E. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proceedings of the National Academy of Sciences*, 113(32), pp.9105-9110.

Yung, A. and McGorry, P. (1996). The Prodromal Phase of First-episode Psychosis: Past and Current Conceptualizations. *Schizophrenia Bulletin*, 22(2), pp.353-370.

Yung, A., Yuen, H., Phillips, L., Francey, S. and McGorry, P. (2005). Mapping the onset of psychosis: The comprehensive assessment of at risk mental states (CAARMS). *Schizophrenia Research*, 60(1), pp.30-31.

Zammit, S., Horwood, J., Thompson, A., Thomas, K., Menezes, P., Gunnell, D., Hollis, C., Wolke, D., Lewis, G. and Harrison, G. (2008). Investigating if psychosis-like symptoms (PLIKS) are associated with family history of schizophrenia or paternal age in the ALSPAC birth cohort. *Schizophrenia Research*, 104(1-3), pp.279-286.