BRAIN
ORIGINAL ARTICLE

Transdiagnostic inflexible learning dynamics explain deficits in depression and schizophrenia

Hans Kirschner,[1] Matthew R. Nassar,[2,3] Adrian G. Fischer,[4] Thomas Frodl,[5,6,7,8] Gabriela Meyer-Lotz,[5] Sören Froböse,[5] Stephanie Seidenbecher,[5] Tilmann A. Klein[1,9,†] and Markus Ullsperger[1,7,8,9,†]

[†]These authors contributed equally to this work.

See Ryan and Treadway (https://doi.org/10.1093/brain/awad400) for a scientific commentary on this article.

Deficits in reward learning are core symptoms across many mental disorders. Recent work suggests that such learning impairments arise by a diminished ability to use reward history to guide behaviour, but the neuro-computational mechanisms through which these impairments emerge remain unclear. Moreover, limited work has taken a transdiagnostic approach to investigate whether the psychological and neural mechanisms that give rise to learning deficits are shared across forms of psychopathology.

To provide insight into this issue, we explored probabilistic reward learning in patients diagnosed with major depressive disorder ($n = 33$) or schizophrenia ($n = 24$) and 33 matched healthy controls by combining computational modelling and single-trial EEG regression. In our task, participants had to integrate the reward history of a stimulus to decide whether it is worthwhile to gamble on it. Adaptive learning in this task is achieved through dynamic learning rates that are maximal on the first encounters with a given stimulus and decay with increasing stimulus repetitions. Hence, over the course of learning, choice preferences would ideally stabilize and be less susceptible to misleading information.

We show evidence of reduced learning dynamics, whereby both patient groups demonstrated hypersensitive learning (i.e. less decaying learning rates), rendering their choices more susceptible to misleading feedback. Moreover, there was a schizophrenia-specific approach bias and a depression-specific heightened sensitivity to disconfirmational feedback (factual losses and counterfactual wins). The inflexible learning in both patient groups was accompanied by altered neural processing, including no tracking of expected values in either patient group.

Taken together, our results thus provide evidence that reduced trial-by-trial learning dynamics reflect a convergent deficit across depression and schizophrenia. Moreover, we identified disorder distinct learning deficits.

1 Institute of Psychology, Otto-von-Guericke University, D-39106 Magdeburg, Germany
2 Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, RI 02912-1821, USA
3 Department of Neuroscience, Brown University, Providence, RI 02912-1821, USA
4 Department of Education and Psychology, Freie Universität Berlin, D-14195 Berlin, Germany
5 Department of Psychiatry and Psychotherapy, Otto-von-Guericke University, D-39106 Magdeburg, Germany
6 Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen 52074, Germany
7 German Center for Mental Health (DZPG), D-39106 Magdeburg, Germany
8 Center for Intervention and Research on adaptive and maladaptive brain Circuits underlying mental health (C-I-R-C), Jena-Magdeburg-Halle, D-39106 Magdeburg, Germany
9 Center for Behavioral Brain Sciences, D-39106 Magdeburg, Germany

Received July 13, 2023. Revised September 25, 2023. Accepted October 10, 2023. Advance access publication December 7, 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of the Guarantors of Brain.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Correspondence to: Hans Kirschner
Institute of Psychology, Otto-von-Guericke University
Universitätsplatz 2, D-39106 Magdeburg, Germany
E-mail: hans.kirschner@ovgu.de

Keywords: reward learning; computational psychiatry; EEG; depression; schizophrenia

Introduction

Cognitive dysfunction is at the core of many disorders and heavily affects quality of life.[1] A wealth of studies have demonstrated its importance and attempted to discern the mechanisms that give rise to cognitive dysfunction in both patients diagnosed with schizophrenia[2,3] and depression.[4] Common to both patient groups are deficits in tasks that require learning.[5,6] A better understanding of the unique and distinct psychological and neural mechanisms that give rise to learning impairments is pivotal to targeting them with interventions and necessitates parsing the learning process into a finer level of detail.

Recent work has highlighted the advantage of designing and analysing experiments within a computational framework to better characterize learning dysfunction. Specifically, it has been suggested that these patient groups have difficulties adaptively updating the value of actions based on probabilistic feedback and in using these value representations to guide behaviour.[5-8] Since multiple cognitive systems interact to support reward learning, it is difficult to separate the processes that contribute to reward learning deficits into finer levels of detail. A recent line of research demonstrated that reward learning depends on cooperative contributions of working memory and reinforcement learning.[9-14] Here, the prefrontal cortex is involved in using working memory to represent values of prospective outcomes and test hypotheses, while the striatal dopamine system affords prediction error signalling and integration of this over trials (i.e. reinforcement learning).[15] In schizophrenia, it has been shown that reward learning impairments arise predominantly from working memory deficits[9,11] and a reduced ability to take future consequences into account.[16] Similarly, working memory deficits have been shown in depression,[17-19] and there is some evidence of increased discounting (i.e. forgetting) of previous reward history in this patient group with increasing working memory load when compared to healthy controls.[20,21] This may suggests that reward learning dysfunctions in depression are also linked to the interaction of reinforcement learning and working memory systems.

A separate line of research links reward learning deficits in depression and schizophrenia to maladaptive integration of new information into beliefs about internal and external states.[22] For example, it has been shown that patients with schizophrenia use an 'all-or-none' belief updating strategy, rendering their learning less flexible and precise.[23] Yet, the neural mechanism underlying these abnormalities remains unclear.

In the normative literature, a uniform sequence of event-related brain potentials (ERPs) has been associated with reward processing: an early frontocentral negativity (FRN), succeeded by a frontocentral positivity (P3a), that is followed by a more sustained parietal positivity called P3b.[24] The early ERP components FRN and P3a have been linked to coding of prediction errors and attentional orienting.[25,26] The P3b likely reflects a signal that is interpreted by downstream processes to set the stage for future behaviour and belief updating.[27-29] Interestingly, there are now multiple reports of relatively normal FRN in schizophrenia,[30] while blunted P3b amplitudes are well documented.[31-34] This body of work suggests that reward information is processed initially but fails to fully impact subsequent behaviour updates in schizophrenia. The literature on depression is less consistent but also hints to blunted reward related EEG activity.[35,36] Critically, the literature remains unclear as to how these alterations relate to learning deficits.

Here, we combined computational modelling and EEG to provide insight into this issue. We administered a simplified variant of an established probabilistic learning task[29] to patients with schizophrenia and depression. In the task, participants must integrate the reward history of a stimulus to decide whether to gamble or avoid gambling on the respective stimulus. Adaptive learning in this task is achieved through dynamic learning rates that are maximal on the first encounters with a given stimulus and decay with increasing stimulus repetitions. Hence, over the course of learning, the choice preferences would ideally stabilize and be less susceptible to misleading information. We used computational modelling to quantify variables involved in learning dynamics and single-trial regressions analyses to identify their EEG signatures. First, we confirmed performance deficits in patients compared to matched healthy control participants. Second, we predicted that these deficits arise from reduced trial-to-trial learning dynamics. Finally, we hypothesized that these learning deficits would be reflected in altered neural dynamics. We found transdiagnostic deficits in learning dynamics and blunted neural representations of expected rewards in both patient groups as well as disorder-specific decision biases. Specifically, we show evidence of reduced learning dynamics, whereby both patient groups demonstrated more hypersensitive learning (i.e. less decay in learning rates), rendering their choices more susceptible to misleading feedback. Moreover, there was a schizophrenia-specific approach bias and a depression-specific heightened sensitivity to disconfirmational feedback. Hypersensitive learning in both patient groups was accompanied by altered neural processing, including no tracking of expected values in either patient group. Hence, we provide evidence of a bunted neural representation of reward expectation in patients. In line with previous research, we demonstrate that parietal EEG activity during later stages of feedback processing may reflect an adaptive cortical mechanism of stimulus value updating by integrating outcomes and learning rate. Interestingly, this mechanism seems to be preserved in a subgroup of both patient groups that show decaying learning rates.

Materials and methods

Participants

To determine learning and reward processing biases in depression and schizophrenia, patients diagnosed with major depressive disorder (MDD; $n = 33$), schizophrenia ($n = 23$) or schizoaffective disorder ($n = 2$), collectively referred to as SZ, and 34 matched healthy

Table 1 Demographic, cognitive, and clinical measures of controls and patients included in the final analyses

Measure	HC ($n = 33$), mean (SD)	MDD ($n = 31$), mean (SD)	SZ ($n = 24$), mean (SD)	Inferential statistic
Demographic				
Age	34.81 (10.83)	33.85 (8.68)	35.00 (9.17)	$F = 0.12$
Gender	17 F; 16 M	18 F; 13 M	7 F; 17 M	–
Education[a]	2.52 (0.62)	2.53 (0.51)	2.00 (0.83)[b,c]	$F = 5.70^{*}$
Cognitive				
TAP-Alertness	25.45 (14.34)	22.25 (14.06)	23.45 (17.56)	$F = 0.37$
TAP-Working memory	51.00 (27.63)	47.37 (29.76)[c]	30.50 (23.83)[b,c]	$F = 3.93^{*}$
TAP-Divided attention	47.81 (24.51)	45.41 (25.68)	34.58 (30.80)	$F = 1.85$
Clinical				
GAF	93.76 (3.49)	67.41 (11.38)[b]	63.46 (12.53)[b]	$F = 88.74^{**}$
BDI-II	2.18 (2.87)	24.81 (10.17)[b,c]	16.42 (10.13)[b,c]	$F = 62.59^{**}$
HAMD	–	15.18 (6.77)	–	–
PANSS				
Positive subscale	–	–	10.50 (4.68)	–
Negative subscale	–	–	13.88 (5.08)	–
General subscale	–	–	24.79 (4.71)	–

BDI-II = Beck depressions inventar-revision II[39]; F = female; GAF = Global Assessment of Functioning[38]; HAMD = Hamilton Rating Scale for Depression[40]; HC = healthy controls; M = male; MDD = patients diagnosed with major depression; PANSS = Positive and Negative Syndrome Scale for Schizophrenia[41]; SD = standard deviation; SZ = patients diagnosed with schizophrenia; TAP = computer-supported test series to assess attention.[37]

[a]Ranging from 1 (primary education) to 3 (higher education).

[b]Significant difference to HC.

[c]Significant difference between patients.

$^{*}P < 0.01$, $^{**}P < 0.001$.

controls (HCs; matched for age, gender and education) were recruited into this study, which took place in the Department of Psychiatry and Psychotherapy at the university hospital in Magdeburg. All participants were of central European origin. The majority of MDD patients (27 of 31; Supplementary Table 1) had a recurrent depression diagnosis with at least two distinct depressive episodes according to DSM-IV and were all currently treated for their depressive symptomatology. The mean Beck Depression Inventory (BDI) and Hamilton Rating Scale for Depression (HAMD) scores (Table 1) of the MDD group can be regarded as 'mild to moderate severity' depressive symptomatology. All participants gave informed consent according to the Declaration of Helsinki, and the research was approved by the ethics committee of the medical faculty of the Otto-von-Guericke University Magdeburg. All participants received €15 for their participation in the experiment, independent of their performance in the probabilistic learning task. Using the proportion of missed trials, we tested for non-compliant participants using Grubb's test for outlier detection with a one-sided alpha of 0.01 [cut-off > 10%, found in three subjects (HC $n = 1$; MDD $n = 1$; SZ $n = 1$)]. Moreover, one patient in the MDD group was on lorazepam during the testing session and had to be excluded.

Clinical and cognitive assessment

Patients were clinically and pharmacologically stable (no change in drug or dose for at least 2 weeks). All SZ patients were on antipsychotic medications and almost all MDD patients took antidepressants (Supplementary Table 1). The diagnosis of the respective patient groups and the absence of a lifetime diagnosis of a psychosis or current mental disorder in the healthy control group was established based on the DSM-IV using the Mini-International Neuropsychiatric Interview (MINI[42]; see Supplementary Table 1 for details of diagnoses). Additional exclusion criteria included: (i) DSM-IV diagnosis of substance abuse or dependence in the past 6 months for all participants; (ii) DSM-IV diagnosis of current major depressive disorder for SZ (note that two SZ patients had a schizoaffective disorder diagnoses); (iii) past head injury with documented neurological sequelae and/or loss of consciousness for all participants; and (iv) benzodiazepine medication within the last 7 days for all participants. All groups were assessed with the German version of BDI-II,[39] the Global Assessment of Functioning (GAF)[38] and a computer-supported test series to assess attention (TAP; subtests allertness, working memory and divided attention).[37] MDD patients were additionally assessed with the German version of the Hamilton Depression Scale.[40] SZ patients were further assessed with the Positive and Negative Syndrome Scale (PANSS) for schizophrenia.[41] Demographic, clinical and cognitive measures of the final sample are listed in Table 1.

Probabilistic learning task

Participants performed a variant of an established learning task.[29] On each trial, participants decided either to gamble or avoid gambling on a centrally presented stimulus (Fig. 1A). Gambling on a stimulus resulted in winning or losing points depending on the reward probability of the respective stimuli. If subjects decided not to gamble, they avoided any consequences, but were still able to observe what would have happened if they had gambled by receiving counterfactual feedback.

The main task consisted of eight blocks in which two stimuli were presented in pseudorandomized order with up to five repetitions of the same stimulus. This resulted in a total of 16 stimuli. Each stimulus was presented at least 20 times and not more than 23 times, resulting in a total of 360 trials. Eight stimuli had a high chance of winning (70% or 80%) and eight stimuli had a low change of winning (20% or 30%). The trial structure is depicted in Fig. 1A and additional information on the task structure can be found in Supplementary Fig. 1.

Reinforcement learning model fitting

To parse out factors influencing choices, we fit variants of reinforcement learning models to participants choice behaviour using a constrained search algorithm (fmincon in MATLAB 2021b), which computed a set of parameters that maximized the

total log posterior probability of choice behaviour. Here, the base model was a standard Q-learning model with two parameters: (i) a constant learning rate (α); and (ii) a temperature parameter (β) of the softmax function. Specifically, on each trial (t) the expected value of a stimulus (Q) was calculated as follows:

$$Q_{t+1} = Q_t + \alpha \times \delta_t \text{ with } \delta_t = R_t - Q_t \quad (1)$$

Q-values represent the expected value of an action at trial t; α reflects the learning rate; and δ_t represents the prediction error, with R_t being the reward magnitude of that trial. The stimulus values were converted into action probabilities according to a softmax rule:

$$P_{c,t} = \frac{1}{1 + exp\left(-\frac{0.5 + Q_t}{\beta}\right)} \quad (2)$$

Next, we fit additional reinforcement models by complementing the base model with additional parameters that were informed by descriptive analyses on the choice behaviour (Supplementary material). Five such parameters were added to the base model: (i) a parameter ($\mathcal{G}$) that allowed free scaling of the initial expected value; (ii) a perseveration parameter (λ), that added a bonus onto the option value of the most recently chosen stimuli to account for potential tendencies to repeat choices; (iii) different learning rates for confirmational (factual wins and counterfactual losses) and disconfirmational (factual losses and counterfactual wins) feedback; and (iv) a learning rate decay function, whereby a decay scaling factor (η) allowed scaling of the learning rate decay in the respective block. Here, the learning rate on a given trial was given by:

$$\alpha_t = \frac{\eta}{T_{bl}} + (1 - \eta) \times \alpha_1 \quad (3)$$

where α_t is the learning rate on a given trial, T_{bl} is the block trial number of the stimulus (i.e. how often has the respective stimuli been seen in the block) and α_1 is the starting learning rate on the first encounter of the stimuli.

Model comparison and validation

For model comparison, we first computed the Bayesian information criterion (BIC) for each participant according to:

$$BIC = \widehat{LL} - \frac{k}{2} ln(T) \quad (4)$$

Where $\widehat{LL}$ is the log-likelihood value at the best fitting parameter settings, k is the number of free parameters and T is the number of trials.[43] Next, we computed protected exceedance probabilities, which gives the probability that one model was more likely than any other model of the model space.[44] To test whether we could distinguish between our models, we conducted a model-recovery study (Supplementary material). To validate the best fitting model, we tested whether we could reliably estimate our free parameters within a parameter and model recovery analyses. Moreover, we conducted post-predictive checks to confirm that the best fitting model captured the key aspects of the choice behaviour. These additional results are presented in the Supplementary material.

EEG measurements and analyses

EEG was recorded at 500 Hz from 64 channels with CPz as reference channel arranged in the extended 10–20 system using BrainAmp amplifiers (Brain Products). Pre-processing of the EEG data was carried out with MATLAB 2021b (The MathWorks, Natick, MA) and the EEGlab 13 toolbox[45] with custom routines as described previously.[27] Pre-processing steps included: (i) filtering (0.3 Hz high- and 40 Hz low-pass filters); (ii) removal and interpolation of bad channels, rejected channels [mean (standard deviation, SD)]: 4.47 (3.51); (iii) re-referencing to a common average; (iv) segmentation into stimulus-locked epochs spanning from 1000 ms pre-stimulus to 4000 ms post-stimulus; (v) automatic epoch rejection, as described by Kirschner *et al.*,[46] rejected epochs [mean (SD)]: 19.55 (9.61); and (vi) removal of blink and eye-movement components, rejected components [mean (SD): 7.92 (3.05)], using adaptive mixture independent component analysis (AMICA).[47]

Following baseline correction (−250 to −50 ms relative to feedback onset), epochs spanning −400 to 1000 ms around feedback presentation were then used for multiple robust single-trial regression analyses.[29] Specifically, for each feedback condition (factual versus counterfactual), we designed two EEG regression models and regressed the single-trial model predictions onto the EEG signal. General linear model 1 (GLM1) included the RPE (δ_t); GLM2 included the RPE components[48] outcome (R_t) and expected value (Q). The groups were comparable in terms of the number of trials these EEG analyses were conducted on [factual trials, HC, mean (SD): 189.93 (24.90); MDD: 183.96 (38.68); SZ: 203.37 (44.34); $F(2,85) = 2.02$, $P = 0.139$; counterfactual trials, HC, mean (SD): 138.12 (23.60); MDD: 139.03 (35.51); SZ: 120.95 (46.43); $F(2,85) = 2.19$, $P = 0.118$]. In a third GLM we investigated the interplay between surprise and learning across all trials. Based on previous research we focused on the P3b in this GLM, as this neural correlate is assumed to be related to behavioural adjustments.[27-29] Here, we included the absolute RPE and the learning rate (α_1) in the EEG regression model. All models included trial number as a regressor of no interest, to account for unspecific task effects like fatigue. Statistics were run on averaged regression weights around the peak (time and location, i.e. electrode) of the respective ERP component (FRN: ±20 ms around peak; P3a: ±40 ms around peak; P3b: mean value between 450 and 550 ms). For the extraction of regression weights of GLM3, we averaged over the time window in which the covariation between the regressor and EEG signal was significant after correction for multiple comparison through false discovery rate (FDR).[49] We chose these broader time windows to investigate whether shifts in EEG activity are associated with behaviour.

Associations

To explore the associations between model-parameters, their EEG correlates, cognitive variables and symptoms, we calculated zero-order Pearson correlations.

Results

Behaviour

In general, participants learned the task well. This was reflected in their choice behaviour that largely followed the reward probabilities of the stimuli across the groups (Fig. 1C). Importantly, however, we found a decreased task performance in patients [Fig. 1B; $F(2,85) = 3.8$, $P = 0.026$]. Follow-up analyses revealed that both MDD [HC versus MDD, mean (SD): 386.36 (38.97) versus 259.68 (47.88); $t(62) = 2.06$, $P = 0.04$, $d = 0.52$] and SZ [HC versus SZ, mean (SD): 386.36 (38.97) versus 217.92 (49.88); $t(55) = 2.69$, $P = 0.009$, $d = 0.74$] participants earned fewer points in the task than HCs. There was no difference between the patient groups in terms of overall

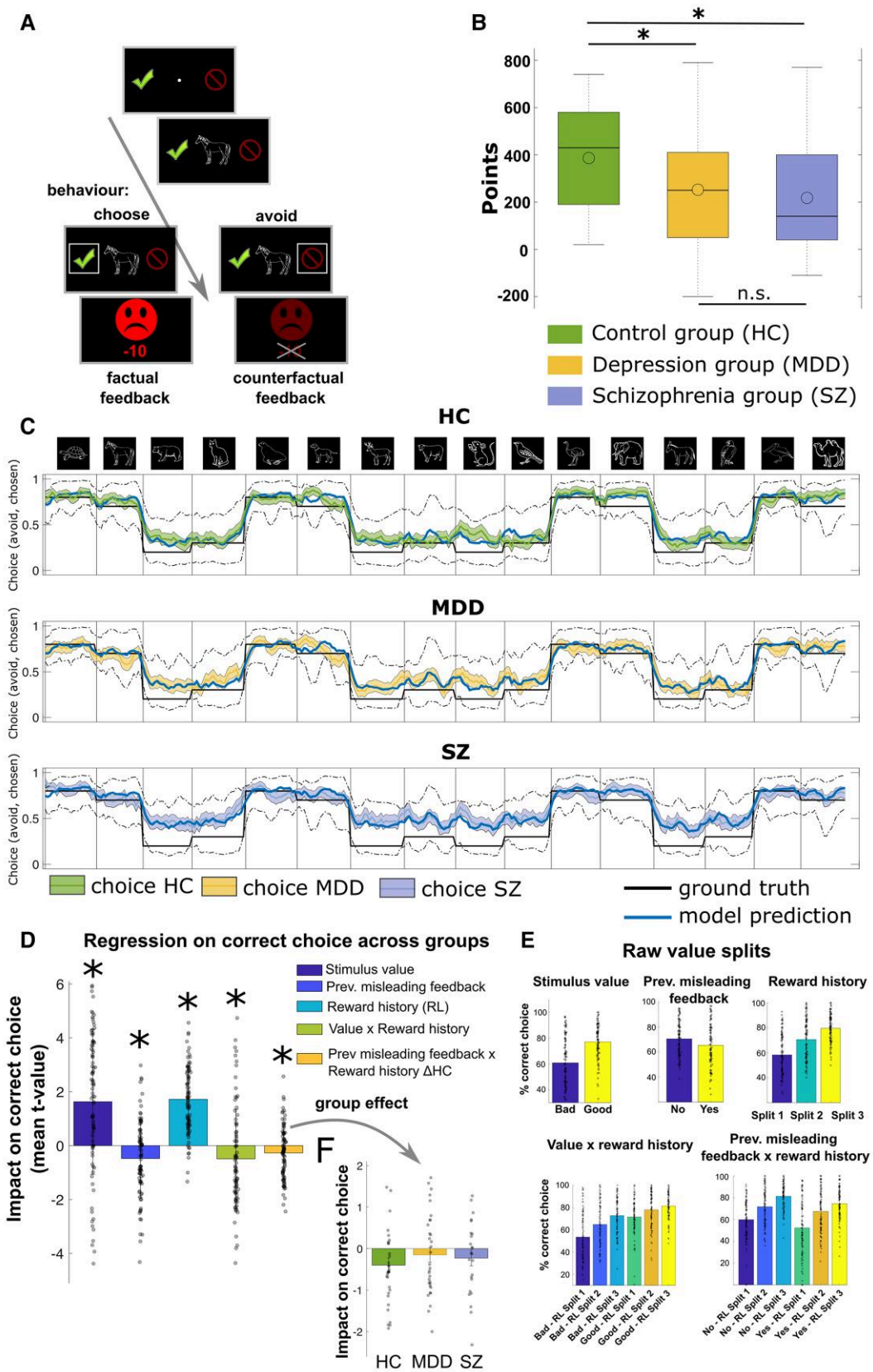

Figure 1 Experimental design, observed behaviour and model predictions. (**A**) Time course of the task. Each trial began with the presentation of a random jitter between 300 and 700 ms. Hereafter, a fixation cross was presented together with two response options (choose, green tick mark; or avoid, red no-parking sign). The response options' sides were counterbalanced across participants and remained in place until the feedback was presented.

(Continued)

performance [mean (SD): 217.92 (49.88) versus 259.68 (47.88); $t(53) = -0.60$, $P = 0.55$, $d = 0.17$]. Moreover, participants committed only a few misses, and the groups were fairly comparable in terms of the number missed trials [HC, mean (SD): 2.66 (3.46); MDD: 2.06 (3.16); SZ: 4.04 (4.51); $F(2,85) = 2.00$, $P = 0.141$].

Learning behaviour

The goal of our study was to understand how factors that govern learning might be altered in depression and schizophrenia and to identify possible neural correlates of these deviations. Behavioural analyses of the choice data (Fig. 1D–F and see the Supplementary material for more details) indicate that participants integrated feedback over time to gamble effectively. Yet, the accuracy of choices was heavily influenced by misleading probabilistic feedback and stimulus values (good versus bad, whereby subjects showed problems in avoiding gambling on bad stimuli) at the beginning of learning in each block. The influence of stimulus value and misleading feedback faded out as learning increased (interaction between reward history and stimulus value/misleading feedback; Fig. 1D and E). Critically, this effect was reduced in both patient groups (indexed by a weaker interaction between reward history and misleading feedback, Fig. 1F). Together, these analyses provide a hint that participants may be changing their learning rate as a function of trials after transitions.

Reduced learning rate decay renders patient choices more susceptible to misleading information

To understand these trial-to-trial learning dynamics, we constructed a nested model set built from reinforcement learning models (Fig. 2A) that included the following features: (i) a temperature parameter of the softmax function used to convert trial expected values to action probabilities (β; higher values indicate increased choice stochasticity); (ii) a learning rate (α) that could either be fixed or change as a function of trials after transitions (scaled by the parameter η, whereby values close to zero indicate fixed learning rates and values close to 1 indicate strong decay in the learning rate); (iii) different learning from confirmational and disconfirmational feedback; (iv) a perseveration parameter to account for potential tendencies to repeat choices irrespective of feedback (λ); and (v) an approach parameter to account for potential tendencies to initially prefer gambling on a stimulus ($\mathcal{G}$, higher values indicate higher initial stimuli values). The winning model (as measured by lowest BIC and achieving protected exceedance probabilities of 92.44%) was one that included all features (Fig. 2B and C and see Supplementary Fig. 9 for more information of the distribution of individual fits across the model space). Sufficiency of the model was evaluated through posterior predictive checks that matched behavioural data on various summary measures (for posterior predictive checks see Supplementary Fig. 2; additional model validation analyses are presented in Supplementary Fig. 5). We did not find evidence for differences in model fit between the groups [$F(2,85) = 0.02$, $P = 0.988$]. The winning model facilitated flexible behavioural adjustments through dynamic learning rates that were maximal on the first trial with a given stimulus and decayed with increasing stimulus repetitions (Fig. 2E). In other words, RPEs influenced behaviour updates more strongly at the beginning of learning when the normative uncertainty is high (in line with the prediction from a Bayesian ideal observer model[50,51]; Supplementary Fig. 5). To validate this interpretation of parameter differences, we conducted an exploratory descriptive analysis and found that participants whose fits indicated less decay in learning rate made more choice switches during the last five trials of each block ($r = -0.46$, $P < 0.001$; Supplementary Fig. 4), when preferences would have ideally stabilized. Critically, learning rate decays were reduced in both patient groups, which was reflected in lower η in both MDD [η MDD: 0.15 (SD: 0.03) versus HC: 0.25 (SD: 0.04); $t(63) = -1.88$, $P = 0.03$, $d = 0.48$] and SZ [η SZ: 0.15 (SD: 0.04); $t(55) = -1.64$, $P = 0.04$, $d = 0.45$]. Consistent with this, both patient groups tended to show more choice switches during the last five trials within a block [switch probability, MDD: 0.36 (SD: 0.02) versus HC: 0.30 (SD: 0.02); $t(63) = 1.51$, $P = 0.0826$, $d = 0.36$; SZ: 0.39 (SD: 0.04); $t(55) = 1.89$, $P = 0.06$, $d = 0.52$]. Moreover, increased switch probability in patients tended to be associated with reduced learning rate decay (MDD: $r = -0.52$, $P = 0.002$, $n = 31$; SZ: $r = -0.33$, $P = 0.06$, $n = 24$; Supplementary Fig. 4). Interestingly, there were no group differences for overall switches [number of overall switches, HC mean (SD): 109.84 (40.27); MDD mean (SD): 110.06 (40.27); SZ mean (SD): 114.45 (54.92); $F(2,85) = 0.08$, $P = 0.925$] or for switches after disconfirmational feedback [i.e. factual loss or counterfactual win; number

Figure 1 Continued
After the fixation cross, the stimulus was shown centrally until the participant responded or for a maximum duration of 2500 ms. If participants failed to respond in time, they were instructed to speed up. Thereafter, participants' choices were confirmed by a white rectangle surrounding the chosen option for 350 ms. Finally, the outcome was presented for 750 ms. If subjects chose to gamble on the presented stimuli, they received either a green smiling face and a reward of 10 points or a red frowning face and a loss of 10 points. When subjects avoided a symbol, they received the same feedback but in a slightly paler colour, and the points that could have been received were crossed out to indicate that the feedback was fictive and had no effect on the total score. (**B**) Total points earned in the task for each group. *Significant difference. Boxes = interquartile range (IQR), horizontal line = median, circle = mean, whiskers = 1.5 IQR. (**C**) Modelled and choice behaviour of the participants in the task stretched out for all stimuli. Note that in the task the different animal stimuli were presented in an intermixed and randomized fashion, but this visualization allows the reader to see that participants' choices followed the reward probabilities of the stimuli. Data plots were smoothed with a running average (±2 trials). Ground truth represents the reward probability of the respective stimuli (good: 70/80%; bad: 20/30%). Dashed black lines represent 95% confidence intervals derived from 1000 simulated agents with parameters that were best fit to participants in each group. Model predictions appear to capture the transitions in choice behaviour well. (**D–F**) Single-trial logistic regression on accuracy (binary variable reflecting whether a participant displayed stimulus-appropriate behaviour on each trial). The regression model included predictors as follows: (i) stimulus value (good versus bad); (ii) previous misleading feedback (i.e. whether the last feedback on the respective stimulus was in line with its reward probability or not); (iii) log-scaled delay (number of trials since the current stimulus was last seen—a simple marker of working memory demand); (iv) reward history (cumulated correct choices for the respective stimulus within each learning block); and (v) trial number (as a regressor of no interest controlling for unspecific task factors like fatigue). Across participants, each of these factors had systematic effects on accuracy, with participants more accurate for high value stimuli and after experiencing cumulative rewards but less accurate after seeing misleading feedback. **E** depicts raw value splits for the predictors allowing visualization of interaction effects. The influence of stimulus value and misleading feedback faded out as learning increased (interaction between reward history and stimulus value/misleading feedback). (**F**) Critically, this effect was reduced in both patient groups. A full description of these results can be found in the Supplementary material. **D** and **F** display average within participant t-values; P-values were derived from t-tests of individual regression weights against zero. *Significant differences after Bonferroni correction. ΔHC = significant group difference; HC = healthy controls; MDD = major depressive disorder; RL = reward learning; SZ = schizophrenia.

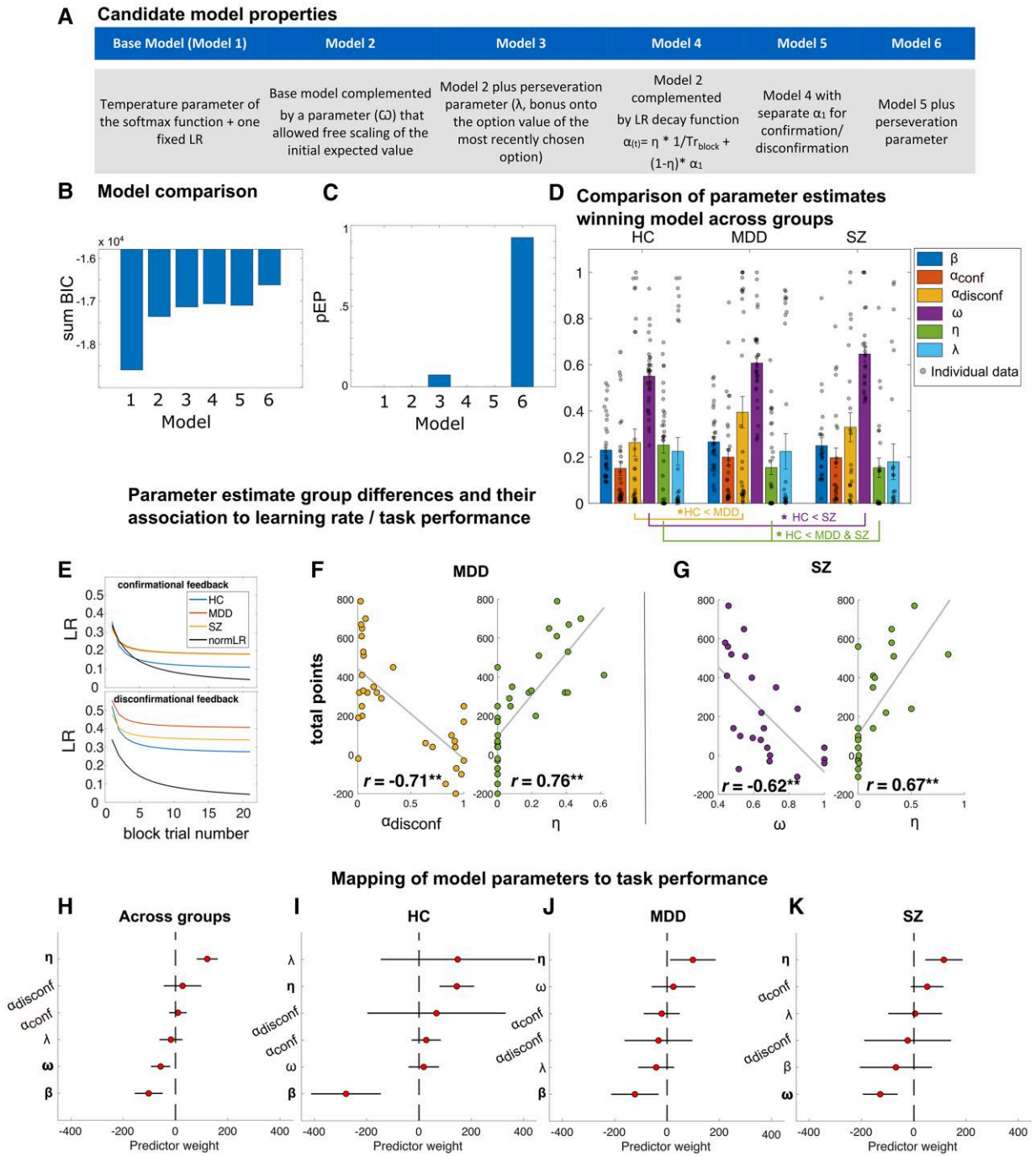

Figure 2 Reinforcement learning modelling, parameter distribution and mapping onto task performance. (**A**) Overview of the candidate model properties. (**B**) Total Bayesian information criterion (BIC) scores over participants of each candidate model. Higher values indicate a better fit of the models to the behavioural data. The results indicate that the full Model 6 provides the best description of participants' choice data. (**C**) Protected exceedance probabilities (pEP) similarly favoured Model 6. (**D**) Comparison of parameter estimates from Model 6 across groups. Asterisks indicate significant group difference for the respective parameter. (**E**) Learning rate (LR) trajectories within each block derived from the learning rate decay function. The data indicate less decay in the learning rate for both patient groups. Moreover, across groups there was a heightened sensitivity to disconfirmational feedback (factual losses/counterfactual wins). While healthy controls (HCs) approached the normative learning rate (black line; derived from an ideal observer model; Supplementary material) during confirmational learning, all groups deviated from normative learning during disconfirmational learning. *Right*: Scatter plots depict correlations between learning biases and task performance for (**F**) major depressive disorder (MDD) and (**G**) schizophrenia (SZ). (**H–K**) Regression coefficients and 95% confidence intervals (points and lines, sorted by value) stipulating the contribution of each model parameter estimate to overall task performance. Here, we regressed model parameters onto task performance. (**H**) Data aggregated over groups. (**I–K**) Data separated by groups. The parameter governing the learning rate decay (η) made a significant positive contribution to overall task performance both across all groups and within each group. Both in HCs and MDD, increased choice stochasticity (β) was associated with worse performance. Across groups and in SZ patients, increased approach bias (ω) made a significant negative contribution to task performance. Note, the overall fit of the general linear models as measured by R^2_{adj} was comparable across groups (HC: $R^2_{adj} = 0.71$; MDD $R^2_{adj} = 0.69$; SZ $R^2_{adj} = 0.74$). λ = perseveration parameter to account for potential tendencies to repeat choices irrespective of feedback; conf = confirmational feedback; disconf = disconfirmational feedback. *$P < 0.05$; **$P < 0.01$.

Table 2 EEG correlates of model estimates

Measure	HC ($n = 33$) mean (SD)	MDD ($n = 31$) mean (SD)	SZ ($n = 24$) mean (SD)	Inferential statistic
Reward prediction error coding				
Factual feedback				
FRN	1.21 (0.29)[a]	0.97 (0.24)[a]	0.74 (0.23)[a]	$F = 0.74$
Outcome	1.23 (0.28)[a]	1.03 (0.26)[a]	0.72 (0.24)[a]	$F = 0.89$
Expected value	−0.66 (0.18)[a]	−0.21 (0.19)	−0.17 (0.2)	$F = 2.21$[†]
P3a	−1.20 (0.36)[a]	−1.19 (0.29)[a]	−1.1 (0.32) (0.36)[a]	$F = 0.03$
P3b	−0.89 (0.27)[a]	−0.73 (0.29)[a]	−0.15 (0.27)[b]	$F = 1.74$[†]
Counterfactual feedback				
FRN	0.01 (0.24)	0.24 (0.22)	0.31 (0.24)	$F = 0.3$
Outcome	0.13 (0.23)	0.29 (0.25)	0.66 (0.25)[a,c]	$F = 1.88$[†]
Expected value	0.36 (0.22)	−0.17 (0.17)	0.54 (0.22)[a,d]	$F = 3.3$[*]
P3a	−0.37 (0.22)	0.43 (0.22)[b]	0.11 (0.28)	$F = 3.29$[*]
P3b	0.65 (0.23)[a]	0.55 (0.25)[a]	0.42 (0.28)	$F = 0.2$
Surprise and LR				
Surprise P3	1.02 (0.24)[a]	0.91 (0.28)[a]	0.35 (0.23)[a,e]	$F = 1.7$[†]
LR P3	0.37 (0.15)[a]	0.56 (0.22)[a]	0.22 (0.25)	$F = 0.45$
Surprise and LR ($\eta > 0.05$)	$n = 20$	$n = 11$	$n = 6$	–
Surprise P3	1.54 (0.30)[a]	2.02 (0.58)[a]	1.52 (0.50)[a]	$F = 0.43$
LR P3	0.93 (0.33)[a]	1.21 (0.41)[a]	0.89 (0.37)[a]	$F = 0.18$

Values represent averaged regression weights around the peak of the respective event-related potential component derived from the respective factor of the single-trial EEG general linear models (GLMs) [feedback related negativity (FRN): 230–270 ms peak at FCz; P3a: 370–450 ms peak at FCz; P3b: 450 and 550 ms]. For the extraction of regression weights of EEG GLM3, we averaged over slightly broader time windows (Surprise P3: 400–650 ms; Learning rate P3: 310–510 ms). HC = healthy controls; LR = learning rate; MDD = patients diagnosed with major depression; RPE = reward prediction error; SD = standard deviation; SZ = patients diagnosed with schizophrenia.

[a]Significantly different from zero.

[b]Significant difference to HC.

[c]Significant difference to HC; trend (P between 0.05 and 0.08).

[d]Significant difference between patients.

[e]Significant difference between patients; trend (P between 0.05 and 0.08).

[†]Trend (P between 0.05 and 0.08).

[*]$P < 0.05$.

of switches, HC mean (SD): 47.45 (23.31); MDD mean (SD): 52.06 (26.18); SZ mean (SD): 53.58 (29.57); $F(2,85) = 0.44$, $P = 0.644$].

The model fits provided some additional results of interest. Across groups, we saw different learning dynamics for confirmational versus disconfirmational feedback, whereby higher learning rates were observed for disconfirmational feedback (Fig. 2D and E). This indicated that participants are more affected by disadvantageous outcomes (i.e. factual losses and counterfactual wins). This effect was particularly pronounced in MDD [$\alpha_{disconf}$ MDD: 0.39 (SD: 0.07) versus HC: 0.26 (SD: 0.06); $t(63) = 1.34$, $P = 0.05$, $d = 0.34$; there was no difference between MDD and SZ, or HC and SZ, all $P > 0.5$]. While the level of perseveration was comparable between the groups (all $P > 0.5$), SZ demonstrated an approach bias that was reflected in higher initial stimulus values [Ω SZ: 0.65 (SD: 0.03) versus HC: 0.55 (SD: 0.02); $t(55) = 2.16$, $P = 0.02$, $d = 0.59$] leading to an increased probability to gamble on stimuli.

Reduced learning rate decay contributes to performance deficits in both patient groups

To investigate how model parameters mapped onto overall task performance, we regressed model parameter estimates onto task performance across and within groups (Fig. 2H–K). The results suggested that the parameter governing the learning rate decay made a significant positive contribution to overall task performance both across all groups and within each group. In other words, stronger decay in the learning rate was associated with better overall task performance. As stated above, there was a general reduction in learning rate decay in both patient groups, and individual differences in both groups correlated strongly with task performance (Fig. 2F and G). Moreover, across groups, increased choice stochasticity was associated with worse performance (Fig. 2H). This may suggest that subjects with reduced transfer of stimulus values to actions (i.e. gamble on stimuli with high values and avoid those with low values) did less well in the task. On the group level, this effect was significant in HCs and MDD (Fig. 2I and J). Across groups and in SZ, increased approach bias made a significant negative contribution to task performance (Fig. 2K). Ideally, the initial stimulus value should be 0.5, as there is no information about the underlying reward probability of a given stimulus during the first encounter. Yet, our data suggested that patients, particularly those diagnosed with schizophrenia, attribute higher initial values to stimuli, leading to an increased probability to gamble on stimuli and decreased task performance (Fig. 2G).

EEG results

For the EEG analyses, we regressed feedback-locked EEG activity across all electrodes onto single-trial model estimates of the RPE and its components (outcome and expected value). This revealed a regression weight time course for single-trial model parameter related activity locked to feedback onset for all electrodes. We then identified the time and location (i.e. electrodes) of maximal activity that mapped onto the ERPs FRN and P3a/P3b and used the average regression weights in a predefined time window around the peak for second level analyses. First level regression weights were scaled by their respective standard errors and thus are comparable across subjects and regressors. The results are shown in Table 2 and Fig. 3 (detailed statistics can be found in the Supplementary material).

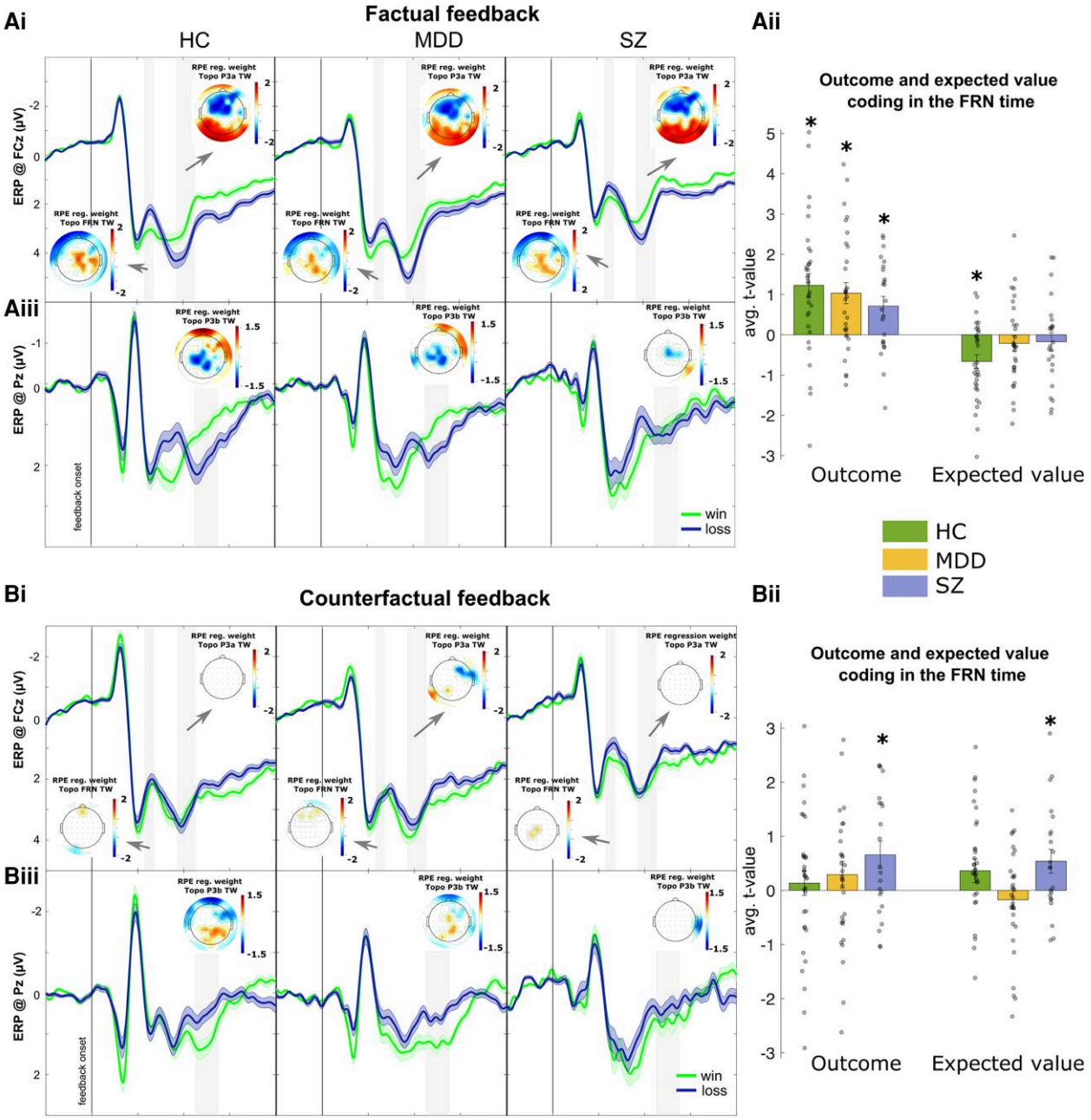

Figure 3 Event-related potentials and regression weight topographies. Event-related potential (ERP) waveforms are depicted separately for factual [**A(ii** and **iii)**] and counterfactual [**B(ii** and **iii)**] wins and losses, together with regression weight topographies. Reward prediction error (RPE) regression weight means [feedback related negativity (FRN): 230–270 ms; frontocentral positivity (P3a): 370–450 ms; sustained parietal positivity (P3b): 450 and 550 ms) thresholded at critical *P*-value from false discovery rate correction]. *Right*: Column depicts the regression weights for the RPE components derived from the RPE component general linear model. *Inset*: Electrodes of maximal effects of early (FRN, P3a at FCz) and late (P3b at Pz) ERP correlates are depicted. Shading indicates 99% confidence intervals. *Significantly different from zero. HC = healthy controls; MDD = major depressive disorder; SZ = schizophrenia; TW = time window.

Factual feedback processing

Significant interrelations between the EEG signal and model derived RPEs were found across all groups contributing to the ERPs FRN and P3a. Interestingly, P3b was present in the HCs and MDD, but nearly absent in SZ. Next, using the RPE component GLM, we tested neural correlates of constituent components of the RPE [Fig. 3A(ii)]. We found that outcome was reflected in the FRN across all groups. Interestingly, the expected value of a given stimulus was only coded in the FRN time window in the HCs but not in the patient groups. Thus, only in HCs does the FRN track the required pattern of an axiomatic RPE signal.[52]

Counterfactual feedback processing

See Fig. 3B for a visualization of the results. Here, in line with previous research, we found that RPEs are not represented in the FRN and P3a during counterfactual feedback processing in HCs and MDD. Interestingly, SZ patients showed a significant central RPE effect in the FRN latency range. When testing RPE subcomponents in

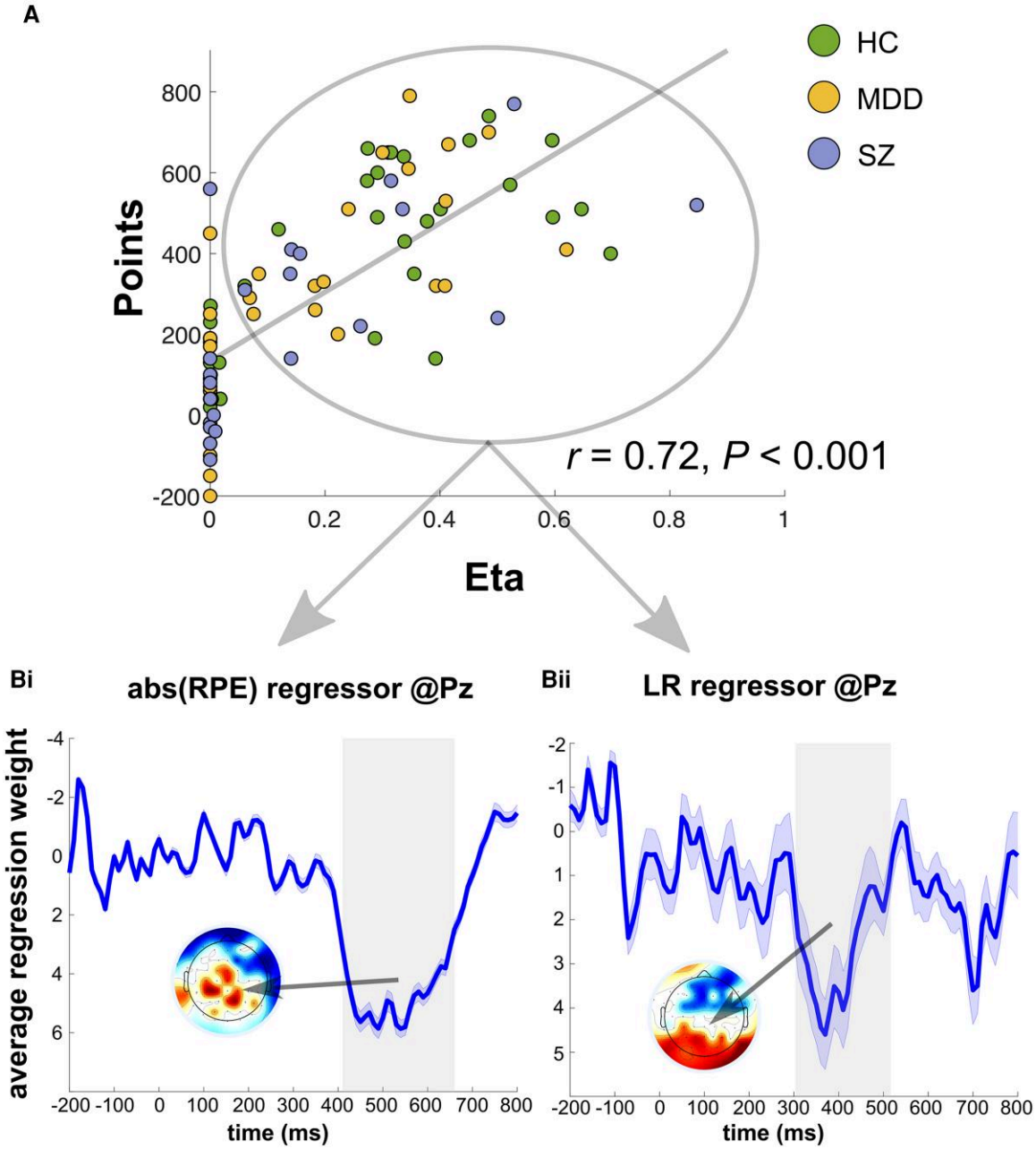

Figure 4 Surprise and learning dynamics. (**A**) Overall task performance (abscissa) was positively related to learning rate decay (ordinate). Across groups, subjects appeared to fall into one of two clusters: a group with a stable learning rate (η around zero) and a group with a substantial decrease in learning rate, whereby the latter cluster was dominated by healthy controls. (**B**) For the learning rate decay cluster, we found a positive central-parietal covariance with the EEG for the absolute reward prediction error (RPE) and the learning rate with a maximal effect at electrode Pz. Grey shaded areas mark the time of significant effects that survived false discovery rate correction and the topographies show average regression weights for the significant time window thresholded at critical *P*-value from false discovery rate correction. Shading indicates 99% confidence intervals. HC = healthy controls; LR = learning rate; MDD = major depressive disorder; SZ = schizophrenia.

the RPE component GLM, we found exclusive outcome coding for SZ, as well es some evidence of expected value coding. As feedback processing continues, RPE are tracked in the P3b. As during factual feedback processing, the late P3b-like effect is blunted in SZ.

Surprise and learning

Next, we investigated the interplay between surprise and learning. In line with predictions from normative learning, subjects with larger decay in their learning rate performed better in the task (Fig. 4A). Here, contradicting information (i.e. unexpected losses or wins) later in a block is downweighed and thus influencing subsequent behaviour less. Across groups, subjects appear to fall into one of two clusters: a group with a stable learning rate [η around zero; this cluster consisted of 25.49% HCs (13 of 33), 39.21% MDD (20 of 31) and 35.29% (18 of 26)] and a group with substantial decrease in learning rate ($\eta > 0.05$), whereby the latter cluster is dominated by HCs [54.05% HCs (20 of 33), 29.72% MDD (11 of 31) and 16.21% SZs (6 of 24)]. To investigate the neural underpinnings of the adaptive value updating described above, we first investigated subjects with a

substantial learning rate decrease and regressed the feedback-locked EEG activity across all electrodes onto the absolute RPE (i.e. surprise) and the learning rate on a given trial. Based on previous research, we focused on late central-parietal EEG activity that coincides with the P3b ERP component in this analysis, as this neural correlate is assumed to be related to behavioural adjustments.[27-29] We found a positive central-parietal covariance with the EEG for the absolute RPE [significant at Pz from 400–650 ms; Fig. 4B(i)] and the learning rate [significant at Pz from 310–510 ms; Fig. 4B(ii)]. In line with previous research,[29] we suggest that the learning rate might modulate the baseline of P3 amplitudes: higher learning rates are associated with a more positive EEG signal and as the learning rate decreases the EEG signal also decreases (but see Nassar *et al.*[28]). Thus, this effect might represent the weighting of the new information, causing less value updates and behavioural adaptation in later trials within each block. Indeed, across groups, future adaptations were best predicted in this time window and topography by multivariate pattern analysis, and the learning rate effect was strongest in people with higher learning rate decay, whereby this relationship appeared to be blunted in SZ (Supplementary material). Interestingly, within the decreasing learning rate cluster, absolute RPEs and trial-by-trial learning rates were similarly tracked in the EEG signal across groups (Table 2). Finally, we extracted mean regression weights (i.e. average regression weights within the significant covariation time window for the absolute RPE and learning rate regressor) for all subjects. Here, absolute RPEs and learning rate were less pronounced in SZ (Table 2).

Associations between model parameters and their EEG correlates to cognition and symptom scores

We also explored whether cognitive and clinical measures are related to model parameters and their EEG correlates (all associations are depicted in Supplementary Fig. 5). Across groups, we found positive correlations between global functioning (indexed by the GAF score[38] and task performance; significant in HCs and SZ). A similar pattern was found for education level (significant in HCs and MDD). In contrast, working memory scores were not related to task performance or any other variables. In HCs and MDD, higher alertness was associated with stronger decay in the learning rate (i.e. higher η) and overall performance. Higher sensitivity to disconfirmational feedback was negatively related to alertness. Learning biases were not related to symptom scores in MDD. In SZ, the approach bias was positively associated with positive and negative symptom scores, which were also negatively related to overall task performed. There was no systematic relationship between EEG correlates and cognition and symptom scores.

Discussion

The goal of this study was to parse out altered learning processes in depression and schizophrenia. To this end, we explored probabilistic learning in MDD and SZ patients by combining computational modelling and single-trial EEG regression. In our task, participants had to integrate the reward history of a stimulus to decide whether it was worthwhile to gamble on it. Advantageous learning in this task was achieved through dynamic learning rates that were maximal on the first encounters with a given stimulus and decay with increasing stimulus repetitions. Consistent with this idea, over the course of the task, choice preferences would ideally stabilize and be less susceptible to misleading information. We found robust evidence for this assumption. Across all groups, decaying learning rates were linked to overall task performance and switching behaviour. Specifically, reduced decay was associated with worse task performance and a higher probability of response switching after misleading feedback. We saw evidence for reduced decay in learning rates for both patient groups. These reductions were linked to performance deficits. The hypersensitive learning in both patient groups was accompanied by altered neural processing, including no representation of expected values in EEG dynamics in either patient group. These results hinted at a transdiagnostic learning deficit that transcends disorders as distinct as depression and schizophrenia. Yet, we also found distinct differences in these patient groups. We discuss these findings in further detail below.

Shared learning deficits in depression and schizophrenia

A key finding of this study is that decaying learning rates are linked to overall task performance. We demonstrated a reduction in learning rate decay in both patients diagnosed with depression and schizophrenia. This reduced learning rate variation suggests that both patient groups show a maladaptive over-responsivity to new information when choice preferences should have ideally stabilized. Here, higher learning rates later in a block lead to over-adjusting, rendering their choices more susceptible to misleading information. In broad strokes, these results are in line with previous research suggesting reinforcement learning deficits in MDD[53-55] and SZ.[7,9,11,16,56-58] Our data suggest that this impairment—at least partly—arises from reduced trial-by-trial learning rate dynamics. It is worth noticing, that this feature would have been missed by averaging learning across trials. Indeed, there are actually no pronounced differences in average learning rates across groups [$F(2,89) = 2.16$, $P = 0.12$]. This highlights the importance of studying trial-by-trial dynamics (see Kirschner and Klein[30] for a discussion on this issue). Moreover, the present findings add to the accumulative evidence of non-normative belief updating in mental disorders.[22] This line of research investigates how individuals adjust their beliefs in light of new information. Here, maladaptive belief updating in clinical populations is characterized by both persistence and hasty changes. While our data do not support the extremes of maladaptive belief updating in depression and schizophrenia [i.e. there was no difference in the perseveration parameter (a proxy of persistence) across groups; moreover, choice behaviour was not better explained by a simple win-stay/lose-shift heuristic (a sign of hasty changes; Supplementary Fig. 8)], they hint at reduced flexibility in belief update in patients diagnosed with depression and schizophrenia. This is in line with recent evidence suggesting less flexible and precise belief updating in schizophrenia.[23]

A different aspect of the reduced learning rate decay findings might be that they reflect increased uncertainty. Here, less decay in learning rates may suggest that patients show higher uncertainty about the underlying value of a stimulus even after they have received enough information to form a judgment on its value. Indeed, previous reports implicating increased uncertainty and intolerance of uncertainty in patients diagnosed with MDD[59,60] and SZ[61] provide at least indirect support for this idea and should motivate future work.

At a neural level, in contrast to HCs, we found no tracking of expected values during early stages of feedback processing in both patient groups. This suggests that outcome expectancy is less represented in SZ and MDD. This is in line with previous research showing deficits in anticipatory mechanisms in patients diagnosed with depression and schizophrenia. For example, Schneider *et al.*[62] found that depressed participants had reduced pupil dilation during reward anticipation. Moreover, a meta-analysis of

functional MRI and EEG studies demonstrated that depressed individuals had reduced striatal activation during reward feedback and anticipation, as well as a blunted FRN.[36] Our results suggest that a blunted FRN partly results from a reduced representation of outcome expectancy and not necessarily blunted outcome processing. In patients diagnosed with schizophrenia, recent data showed that they are less able to mobilize predictive mechanisms to facilitate processing at the earliest stages of accessing the meanings of incoming words.[63]

Distinct learning processes in depression and schizophrenia

While reduced decay in learning rates reflects a convergent deficit across depression and schizophrenia, we also identified disorder distinct deficits that were related to reduced overall task performance. We found a MDD specific heightened sensitivity to disconfirmational feedback, which was reflected in higher initial learning rates for this feedback category. This finding adds to the accumulative evidence suggesting an attentional bias towards negative information in MDD.[64-67] Moreover, there was a SZ- specific approach bias. This was reflected in higher initial expected values for novel stimuli in this patient group. Consistent with this, SZ patients had an increased probability to gamble on stimuli and difficulties avoiding stimuli with low reward probability. This deficit relates to research suggesting less flexible behaviour and increased cognitive rigidity in schizophrenia.[68-70] At a neural level, we found exclusive fictive outcome coding in the FRN in SZ. To the best of our knowledge, this is the first study to demonstrate outcome tracking in the FRN for fictive outcomes. In previous studies,[27,29,71,72] as well as in the HC and MDD groups, there was no representation of the outcome or RPE in the time range of the FRN for fictive outcomes. We speculate that this may reflect a lower sense of agency for fictive outcomes. Thus, the representation of fictive outcomes in the FRN of SZ patients may represent further neural evidence of a disrupted sense of agency that has previously been reported in SZ.[73,74] For factual feedback, we found relatively normal outcome tracking in the FRN, which is in line with multiple previous reports.[30] Yet, as feedback processing continued, SZ patients showed blunted tracking of feedback information in the P3 latency, which is in line with numerous studies.[31-33] Taken together, these findings indicate that feedback information is processed but does not appear to have a full impact on subsequent behaviour.

Several other results deserve discussion. First, in this study we were able to replicate our previous findings, that factual and counterfactual feedback are initially processed differently but converge on a common parietal EEG effect that drives further decisions and learning.[27,29] Strikingly, this adaptive mechanism is preserved in patients that show decay in their learning rate. This finding is in line with research showing intact reversal learning in a subgroup of SZ patients.[69] Second, patient groups did not differ from controls in their noise parameters, a finding that has been reported previously.[10] This suggests that the patients did not simply apply a noisier decision strategy but demonstrated an impairment in learning the appropriate stimulus value. Third, except for a correlation between positive and negative symptoms and the approach bias in SZ patients, we did not see a systematic relationship between our model parameters, EEG activity and symptom profiles across patients. There are several reasons that might explain the lack of symptom associations. One consideration is that our patients were stably medicated, leading to lower symptoms and, together with the moderate sample size, this potentially limited our ability to detect relationships between symptom profiles and task measures. Moreover, our symptom measures did not provide a detailed assessment of symptoms related to reinforcement deficits that might be expected to relate to task behaviour. We encourage future research to use newer generation scales like the Brief Negative Symptom Scale (BNSS) and the Clinical Assessment Interview for Negative Symptoms (CAINS)[75] as the primary measures of negative symptoms, because newer scales have an avolition or anhedonia item, and reduced goal-oriented behaviour was a key symptom of interest examined our study and research into reward learning deficits in general. Yet, the absence of a group difference in the noise parameter and number of missed trials gave at least indirect support that anhedonia and avolition were not driving group differences in our task. Fourth, it should be noted that most patients were on a variety of psychotropic medications, including antipsychotics and antidepressants (see Supplementary Table 1 for further details). While we did not see a systematic association between medication and the identified learning deficits in both patient groups it is still possible that medication interacted with task performance. An important next step will be to examine medication-free patients. Fifth, the lack of working memory findings was surprising, given the established link between working memory and the reinforcement learning system in reward learning[12,76] and seminal work showing that reinforcement learning dysfunction in SZ is linked to the interaction of reinforcement learning and working memory systems.[9,11] Yet, we specifically designed our task to have low demand on working memory, and we did not systematically manipulate working memory load. The fact that we did not see an association between working memory capacity and performance reflects these details of our experimental design. Likely, other systems and mechanisms were predominately engaged in our task. Specifically, hypersensitivity to new information (reflected in less decay in learning rates), approach biases (SZ) and attentional bias to negative information (MDD) seemed to drive performance differences in our task. Finally, it is important to note that the patient groups differed in substantial ways other than diagnostics. For example, the SZ patients were less educated, and there were more females in the MDD group. We did not find a difference between patient groups in the key results of this study (i.e. both groups demonstrated a similar reduced learning rate decay that was linked to overall performance and an absent representation of outcome prediction in the FRN) and demographical differences were not related to disorder specific deficits (Supplementary Fig. 7). Nonetheless, these differences should be considered when interpreting the presented results.

In summary, through the combination of computational modelling and single-trial EEG regression, it appears that learning deficits in depression and schizophrenia arise partly from reduced trail-by-trial learning dynamics. The hypersensitive learning in both patient groups was accompanied by altered neural processing, including no tracking of expected values in either patient group.

Data availability

All code and data that support the findings of this study are available from the corresponding author, upon reasonable request. Researchers who wish to access processed and anonymized data from a reasonable perspective should contact the corresponding author. Data will be released to researchers if it is possible under the terms of the GDPR (General Data Protection Regulation).

Acknowledgements

We thank all our participants who took part in this research for the generosity of their time and commitment. Moreover, we

acknowledge support by the Open Access Publication Fund of Magdeburg University.

Funding

No funding was received towards this work.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Millan MJ, Agid Y, Brune M, *et al*. Cognitive dysfunction in psychiatric disorders: Characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov*. 2012;11:141-168.
2. Dickinson D, Ramsey ME, Gold JM. Overlooking the obvious: A meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. *Arch Gen Psychiatry*. 2007;64:532-542.
3. Schaefer J, Giangrande E, Weinberger DR, Dickinson D. The global cognitive impairment in schizophrenia: Consistent over decades and around the world. *Schizophr Res*. 2013;150:42-50.
4. Snyder HR. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychol Bull*. 2013;139:81-132.
5. Whitton AE, Treadway MT, Pizzagalli DA. Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr Opin Psychiatry*. 2015;28:7-12.
6. Culbreth AJ, Moran EK, Barch DM. Effort-cost decision-making in psychosis and depression: Could a similar behavioral deficit arise from disparate psychological and neural mechanisms? *Psychol Med*. 2018;48:889-904.
7. Gold JM, Waltz JA, Prentice KJ, Morris SE, Heerey EA. Reward processing in schizophrenia: A deficit in the representation of value. *Schizophr Bull*. 2008;34:835-847.
8. Pechtel P, Dutra SJ, Goetz EL, Pizzagalli DA. Blunted reward responsiveness in remitted depression. *J Psychiatr Res*. 2013;47:1864-1869.
9. Collins AG, Brown JK, Gold JM, Waltz JA, Frank MJ. Working memory contributions to reinforcement learning impairments in schizophrenia. *J Neurosci*. 2014;34:13747-13756.
10. Collins AG, Frank MJ. How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur J Neurosci*. 2012;35:1024-1035.
11. Collins AGE, Albrecht MA, Waltz JA, Gold JM, Frank MJ. Interactions among working memory, reinforcement learning, and effort in value-based choice: A new paradigm and selective deficits in schizophrenia. *Biol Psychiatry*. 2017;82:431-439.
12. Collins AGE, Frank MJ. Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory. *Proc Natl Acad Sci U S A*. 2018;115:2502-2507.
13. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011;69:1204-1215.
14. Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A*. 2013;110:20941-20946.
15. Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*. 2004;306:1940-1943.
16. Culbreth AJ, Westbrook A, Daw ND, Botvinick M, Barch DM. Reduced model-based decision-making in schizophrenia. *J Abnorm Psychol*. 2016;125:777-787.
17. Nikolin S, Tan YY, Schwaab A, Moffa A, Loo CK, Martin D. An investigation of working memory deficits in depression using the n-back task: A systematic review and meta-analysis. *J Affect Disord*. 2021;284:1-8.
18. McIntyre RS, Cha DS, Soczynska JK, *et al*. Cognitive deficits and functional outcomes in major depressive disorder: Determinants, substrates, and treatment interventions. *Depress Anxiety*. 2013;30:515-527.
19. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: A systematic review and meta-analysis. *Psychol Med*. 2014;44:2029-2040.
20. Ang YS, Gelda SE, Pizzagalli DA. Cognitive effort-based decision-making in major depressive disorder. *Psychol Med*. 2023;53:4228–44235.
21. Rupprechter S, Stankevicius A, Huys QJM, Steele JD, Seriès P. Major depression impairs the use of reward values for decision-making. *Sci Rep*. 2018;8:13798.
22. Kube T, Rozenkrantz L. When beliefs face reality: An integrative review of belief updating in mental health and illness. *Perspect Psychol Sci*. 2021;16:247-274.
23. Nassar MR, Waltz JA, Albrecht MA, Gold JM, Frank MJ. All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs. *Brain*. 2021;144:1013-1029.
24. Ullsperger M, Fischer AG, Nigbur R, Endrass T. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci*. 2014;18:259-267.
25. Polich J. Updating P300: An integrative theory of P3a and P3b. *Clin Neurophysiol*. 2007;118:2128-2148.
26. Sambrook TD, Goslin J. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychol Bull*. 2015;141:213-235.
27. Kirschner H, Fischer AG, Ullsperger M. Feedback-related EEG dynamics separately reflect decision parameters, biases, and future choices. *Neuroimage*. 2022;259:119437.
28. Nassar MR, Bruckner R, Frank MJ. Statistical context dictates the relationship between feedback-related EEG signals and learning. *Elife*. 2019;8:e46975.
29. Fischer AG, Ullsperger M. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*. 2013;79:1243-1255.
30. Kirschner H, Klein TA. Beyond a blunted ERN—Biobehavioral correlates of performance monitoring in schizophrenia. *Neurosci Biobehav Rev*. 2022;133:104504.
31. Bramon E, McDonald C, Croft RJ, *et al*. Is the P300 wave an endophenotype for schizophrenia? A meta-analysis and a family study. *Neuroimage*. 2005;27:960-968.
32. Qiu YQ, Tang YX, Chan RC, Sun XY, He J. P300 aberration in first-episode schizophrenia patients: A meta-analysis. *PLoS One*. 2014;9:e97794.
33. Ford JM. Schizophrenia: The broken P300 and beyond. *Psychophysiology*. 1999;36:667-682.
34. Frodl T, Meisenzahl EM, Müller D, *et al*. P300 subcomponents and clinical symptoms in schizophrenia. *Int J Psychophysiol*. 2002;43:237-246.
35. Bruder GE, Kayser J, Tenke CE. Event-related brain potentials in depression: Clinical, cognitive, and neurophysiological

implications. In: Kappenman ES, Luck SJ, eds. *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press; 2012:564–592.
36. Keren H, O'Callaghan G, Vidal-Ribas P, *et al*. Reward processing in depression: A conceptual and meta-analytic review across fMRI and EEG studies. *Am J Psychiatry*. 2018;175:1111-1120.
37. Zimmermann P, Fimm B. *Testbatterie zur aufmerksamkeitsprüfung (TAP)*. Psychologische Testsysteme; 2002.
38. Association AP. *Diagnostisches und statistisches manual psychischer störungen—DSM-5®*. Hogrefe; 2015.
39. Hautzinger M, Keller F, Kühner C. *BDI-II. Beck depressions inventar revision—Manual*. Harcourt Test Services; 2012.
40. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62.
41. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13: 261-276.
42. Sheehan DV, Lecrubier Y, Sheehan KH, *et al*. The Mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998; 59(Suppl 20):22-33.
43. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian Model selection for group studies. *Neuroimage*. 2009; 46:1004-1017.
44. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian Model selection for group studies—Revisited. *Neuroimage*. 2014;84: 971-985.
45. Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004;134:9-21.
46. Kirschner H, Humann J, Derrfuss J, Danielmeier C, Ullsperger M. Neural and behavioral traces of error awareness. *Cogn Affect Behav Neurosci*. 2021;21:573-591.
47. Palmer J, Kreutz-Delgado K, Makeig S. *AMICA: An adaptive mixture of independent component analyzers with shared components*: Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech Rep.; 2012.
48. Rutledge RB, Dean M, Caplin A, Glimcher PW. Testing the reward prediction error hypothesis with an axiomatic model. *J Neurosci*. 2010;30:13525-13536.
49. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165–1188.
50. Jang AI, Nassar MR, Dillon DG, Frank MJ. Positive reward prediction errors during decision-making strengthen memory encoding. *Nat Hum Behav*. 2019;3:719-732.
51. Nassar MR, Wilson RC, Heasly B, Gold JI. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci*. 2010;30:12366-12378.
52. Caplin A, Dean M. Axiomatic methods, dopamine and reward prediction error. *Curr Opin Neurobiol*. 2008;18:197-202.
53. Chen C, Takahashi T, Nakagawa S, Inoue T, Kusumi I. Reinforcement learning in depression: A review of computational research. *Neurosci Biobehav Rev*. 2015;55:247-267.
54. Vandendriessche H, Demmou A, Bavard S, *et al*. Contextual influence of reinforcement learning performance of depression: Evidence for a negativity bias? *Psychol Med*. 2023;53:4696–44706.
55. Murphy FC, Michael A, Robbins TW, Sahakian BJ. Neuropsychological impairment in patients with major depressive disorder: The effects of feedback on task performance. *Psychol Med*. 2003;33:455-467.
56. Murray GK, Cheng F, Clark L, *et al*. Reinforcement and reversal learning in first-episode psychosis. *Schizophr Bull*. 2008;34: 848-855.
57. Polli FE, Barton JJ, Thakkar KN, *et al*. Reduced error-related activation in two anterior cingulate circuits is related to impaired performance in schizophrenia. *Brain*. 2008;131(4): 971-986.
58. Somlai Z, Moustafa AA, Keri S, Myers CE, Gluck MA. General functioning predicts reward and punishment learning in schizophrenia. *Schizophr Res*. 2011;127(1–3):131-136.
59. McEvoy PM, Hyett MP, Shihata S, Price JE, Strachan L. The impact of methodological and measurement factors on transdiagnostic associations with intolerance of uncertainty: A meta-analysis. *Clin Psychol Rev*. 2019;73:101778.
60. Carleton RN, Mulvogue MK, Thibodeau MA, McCabe RE, Antony MM, Asmundson GJG. Increasingly certain about uncertainty: Intolerance of uncertainty across anxiety and depression. *J Anxiety Disord*. 2012;26:468-479.
61. Kreis I, Zhang L, Moritz S, Pfuhl G. Spared performance but increased uncertainty in schizophrenia: Evidence from a probabilistic decision-making task. *Schizophr Res*. 2022;243:414-423.
62. Schneider M, Elbau IG, Nantawisarakul T, *et al*. Pupil dilation during reward anticipation is correlated to depressive symptom load in patients with Major depressive disorder. *Brain Sci*. 2020; 10:906.
63. Sharpe V, Weber K, Kuperberg GR. Impairments in probabilistic prediction and Bayesian learning can explain reduced neural semantic priming in schizophrenia. *Schizophr Bull*. 2020;46: 1558-1566.
64. Siegle GJ, Granholm E, Ingram RE, Matt GE. Pupillary and reaction time measures of sustained processing of negative information in depression. *Biol Psychiatry*. 2001;49:624-636.
65. Peckham AD, McHugh RK, Otto MW. A meta-analysis of the magnitude of biased attention in depression. *Depress Anxiety*. 2010;27:1135-1142.
66. Gotlib IH, Krasnoperova E, Yue DN, Joormann J. Attentional biases for negative interpersonal stimuli in clinical depression. *J Abnorm Psychol*. 2004;113:127–1235.
67. Brown VM, Zhu L, Solway A, *et al*. Reinforcement learning disruptions in individuals with depression and sensitivity to symptom change following cognitive behavioral therapy. *JAMA Psychiatry*. 2021;78:1113-1122.
68. Cools R, Brouwer WH, de Jong R, Slooff C. Flexibility, inhibition, and planning: Frontal dysfunctioning in schizophrenia. *Brain Cogn*. 2000;43(1–3):108-112.
69. Reddy LF, Waltz JA, Green MF, Wynn JK, Horan WP. Probabilistic reversal learning in schizophrenia: Stability of deficits and potential causal mechanisms. *Schizophr Bull*. 2016;42:942-951.
70. Baker SC, Konova AB, Daw ND, Horga G. A distinct inferential mechanism for delusions in schizophrenia. *Brain*. 2019;142: 1797-1812.
71. Humann J, Fischer AG, Ullsperger M. The dynamics of feedback-based learning is modulated by working memory capacity. *PsyArXiv*. [Preprint] doi: 10.31234/osf.io/8qzu2
72. Schüller T, Fischer AG, Gruendler TOJ, *et al*. Decreased transfer of value to action in Tourette syndrome. *Cortex*. 2020;126:39-48.
73. Synofzik M, Voss M. Disturbances of the sense of agency in schizophrenia. In: Michela B, ed. *Neuropsychology of the sense of agency: From consciousness to action*: Springer; 2010.
74. Salgado-Pineda P, Fuentes-Claramonte P, Spanlang B, *et al*. Neural correlates of disturbance in the sense of agency in schizophrenia: An fMRI study using the 'enfacement' paradigm. *Schizophr Res*. 2022;243:395-401.
75. Carpenter WT, Blanchard JJ, Kirkpatrick B. New standards for negative symptom assessment. *Schizophr Bull*. 2015;42:sbv160.
76. Yoo AH, Collins AGE. How working memory and reinforcement learning are intertwined: A cognitive, neural, and computational perspective. *J Cogn Neurosci*. 2022;34:551-568.