REINFORCEMENT LEARNING IN MDD

Spared internal but impaired external reward prediction error signals in Major Depressive Disorder during reinforcement learning

Jasmina Bakic* and Gilles Pourtois* ,Department of Experimental Clinical & Health Psychology, Ghent University, Ghent, Belgium

Marieke Jepma, Institute of Psychology, Leiden University, Leiden Institute for Brain and Cognition, Leiden, The Netherlands

Romain Duprat, Department of Psychiatry and Medical Psychology, Ghent University, Universitair Ziekenhuis Gent, Ghent, Belgium

Rudi De Raedt, Department of Experimental Clinical & Health Psychology, Ghent University, Ghent, Belgium

Chris Baeken, Department of Psychiatry and Medical Psychology, Ghent University, Universitair Ziekenhuis Gent, Ghent, Belgium

Corresponding author: Jasmina Bakic, Brain Stimulation and Cognition group Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Oxfordlaan 55, 6229 EV Maastricht, The Netherlands, E-mail: jasmina.bakic@maastrichtuniversity.nl

* These two authors contributed to the manuscript equally

ABSTRACT

Background. Major depressive disorder (MDD) creates debilitating effects on a wide range of cognitive functions, including reinforcement learning (RL). In this study, we sought to assess whether reward processing as such, or alternatively the complex interplay between motivation and reward might potentially account for the abnormal reward-based learning in MDD.

Methods. A total of 35 treatment resistant MDD patients and 44 age matched healthy controls (HCs) performed a standard probabilistic learning task. RL was titrated using behavioral, computational modeling and event-related brain potentials (ERPs) data.

Results. MDD patients showed comparable learning rate compared to HCs. However, they showed decreased lose-shift responses as well as blunted subjective evaluations of the reinforcers used during the task, relative to HCs. Moreover, MDD patients showed normal internal (at the level of error related negativity, ERN) but abnormal external (at the level of feedback related negativity, FRN) reward prediction error (RPE) signals during RL, selectively when additional efforts had to be made to establish learning.

Conclusions. Collectively, these results lend support to the assumption that MDD does not impair reward processing per se during RL. Instead, it seems to alter the processing of the emotional value of (external) reinforcers during RL, when additional intrinsic motivational processes have to be engaged.

Keywords: depression, EEG/ Evoked potentials, cognition

INTRODUCTION

In an attempt to shed light on the defining emotional deficit characterizing MDD, many bets in the state of the art research are currently placed on anhedonia, one of the cardinal symptoms of this mental illness. Defined as a "loss of pleasure or lack of reactivity to pleasurable stimuli"[1], anhedonia is hypothesized to account for learning deficits visible in MDD when reward processing and utilization is crucial, such as in reinforcement learning (RL). Using this framework, two studies previously showed the reduced development of an implicit positivity bias (or active pursuit of rewarding outcomes) across time in MDD patients with high anhedonia [2,3]. However, in these earlier studies, monetary/secondary reward was used [4]. Unlike monetary reward for which a fixed value is usually provided to the participant, goal attainment relates to the (subject-specific) hedonic experience encountered (or anticipated) when a cue signals that the task at hand has been fulfilled, and self-efficacy is in turn transiently reinforced [5,6].

Because reward-related cues informing about self-efficacy (e.g. feedback on task performance) necessarily provide potent motivational signals to the organism, their swift use to guide learning might be compromised by MDD. The goal of this study was to test this prediction, using a multi-methods approach. RL is paradigmatic example of a situation where internal and external cues have to be used timely to guide the course of learning. At the electrophysiological level, this process has been associated with the generation of the ERN (response-locked) and FRN (feedback-locked) event related potential (ERP) component, respectively [7]. The ERN and FRN are thought to reflect phasic reward prediction error (RPE) signals (either based on an internal/motor or external cue)

In this study, we tested a well-defined cohort of treatment resistant MDD patients (with high level of anhedonia) and compared their learning performance and RPE signals

(using conventional EEG/ERP methods) during a probabilistic learning task [8,9] to a group of age and education-level matched healthy controls (HCs). We assessed if MDD could impair internal (ERN) and/or external (FRN) RPE signals, and whether it would be associated with decreased RL (at the behavioral level) compared to HCs in this task [2]. Given that we used motivationally significant (self-efficacy related) reward and punishment cues as learning signals [10, 11], we surmised that MDD might very well influence it in a way that directly depends on reward probability and effort investment to achieve learning [12]. More specifically, when extra efforts are required to establish learning, abnormal reward prediction error signals (and hence abnormal RL) should be observed in this condition (see [13] for evidence with non-human data).

METHODS

Participants

Sixty non-depressed HCs (35 females, 25 males, mean age: 37.90, *SD* = 12.82) and forty-two individuals meeting the *Diagnostic and Statistical manual of Mental Disor*ders 4 criteria [14] for MDD (30 females, 12 males, mean age: 41.40, *SD* = 12.04) participated in the current study. The two groups were matched for age, sex and education. All participants had normal or corrected to normal vision.

The patients were all diagnosed with MDD by using the Mini-International Neuropsychiatric Interview [15]. Depression severity was assessed with the 17-item Hamilton Rating Scale for Depression (HRSD) [16], and the 21-item Beck Depression Inventory (BDI) [17] by a certified psychiatrist. They filled in the Snaith-Hamilton Pleasure Scale [18], and the Temporal Experience of Pleasure Scale [19]. These patients were classified as at least Stage I treatment resistant [20]. All patients were free from any antidepressant (AD), neuroleptic and mood stabilizer for at least two weeks. Exclusion criteria were (a) bipolarity, (b) a history of

neurological disorders, including epilepsy, head injury, and a loss of consciousness, (c) a history of electroconvulsive therapy, (d) a past or present substance abuse, (e) past or present experience of psychotic episodes. Finally, some of those admitted to the study were excluded a posteriori for the following reasons: (f) balancing average age between the two samples (n = 4 HCs), (g) insufficient or no learning during the RL task, (i.e. below chance level). The data of 16 participants (11 in the HC and 5 in the MDD group) were excluded accordingly, and (j) additional 3 (1 in HC and 2 in MDD group) due to excessively noisy EEG signal. Based on these criteria men were excluded significantly more than women ($\chi^2(3) = 9.44$, $p = .024$). The two groups did not differ significantly for the number of participants excluded ($p = .172$). Importantly, inclusion of these participants did not change the results of the analyses reported below, however it was decided not to include them in these analyses to reduce the noise in the data. The final sample consisted of 44 HCs and 35 MDD patients. Demographic and clinical data are presented in Table 1. The study was approved by the ethics committee of the Ghent University Hospital.

Probabilistic learning task

We used a probabilistic learning task previously devised by Eppinger[8] and used by Bakic[21], as well as by Unger[9]. After a fixation cross of 250 ms duration, and a blank screen (250 ms), a visual stimulus (S) was presented for 500 ms on each trial against a white homogenous background on a 17-inch computer screen. Its mean size was 7 cm width x 5 cm height, corresponding to 5 x 3,6 degrees of visual angle at 80 cm viewing distance. Participants performed a two-alternative forced choice task and decide (with a 800 ms response deadline) whether the stimulus was associated with response (R) 1 or 2. After a 500 ms blank, they received (visual) feedback (500 ms), informing about the accuracy of their action. The inter-trial interval was 500 ms. Unbeknownst to the participant, three stimulus conditions (corresponding to three different reward probabilities) were used in random order: the S-R

association was deterministic, probabilistic or random (see supplementary materials). Each participant completed two blocks of 240 trials. Each block had six different stimuli (there were each time 2 different stimuli used per condition), each repeated forty times. Trial order within a block, as well as the order of the two blocks was alternated across participants.

Procedure

Prior to the actual testing session, participants were asked not to consume any caffeine or nicotine. After the EEG preparation, they first performed a practice of 20 trials, after which the experimental session began. After each block, participants were asked to indicate, for each of the 6 stimuli, the clarity and certainty of each of the six S-R associations, by means of a horizontal 10-cm visual analogue scale (VAS). Furthermore, they were asked to rate the amount of positive vs. negative feedback they thought they received during this last block (using a 10 cm VAS going from "exclusively negative" to "exclusively positive"), as well as how much they liked or disliked this positive vs. negative feedback when receiving them (using a Likert scale spanning from 0 to 100).

EEG recording

EEG was recorded continuously using 64-channels by means of a Biosemi Active Two system (www. Biosemi.com). The EEG was sampled at 512 Hz, with CMS-DRL serving as the reference-ground. The EEG signal was filtered off line, using a 0.016 to 70 Hz filter (12db/oct), with a 50 Hz notch and re-referenced using the linked (average) mastoids. For response-locked ERPs (ERN), individual epochs were segmented using a -/+ 500 ms interval around the response (see ref [22-24]). For feedback-locked ERPs (FRN), epoching was made 200 prior to until 800 ms following feedback onset. Eye blinks were removed automatically via vertical ocular correction [25], using two electrodes, placed above and below the right eye. Individual epochs were baseline corrected using the first 200 ms of the pre-response time-

interval for the ERN (i.e. from -500 to -300 ms prior to response onset) and the entire pre-stimulus time interval for the FRN (i.e. 200 ms).

Artifact rejection was based on a ± 100 μV amplitude cutoff. For response-locked segments, it led to 84.64% of the individual segments being kept and eventually included in the individual averages. No significant group difference [HCs: *M* = 84.46, *SEM* = 0.84; MDD patients: *M* = 84.39, *SEM* = 1.08; *t* (84) = 0.51, *p* = .96] was found for this metric. For feedback–locked segments, 84.86% of the individual epochs were kept. No group difference was found for this metric either [HCs: M= 85.25, SEM= 0.97; MDD: M= 84.42, SEM= 1.22, *t* (75) = 0.54, p=.59]. Finally, individual epochs were averaged separately for the different conditions and subjects, and an additional low pass filter set to 30 Hz was applied on the individual averages before grand-averaging.

Data analysis

Behavioral data (accuracy and switch after negative feedback) were analyzed by means of a mixed model ANOVA with group as a between subjects factor, and condition (n=3) and bin (n=4, where trials were grouped in four parts of 60 trials, 20 per condition) as a within subject factor. Switch after negative feedback captures the sensitivity to negative feedback and has been described as a change of lose-shift strategy (see ref [26,27]). Where necessary, Greenhouse-Geisser correction for sphericity was performed, and corrected p-values were reported, together with the effect size and the 95% confidence interval (CI) around this value. Description of the reinforcement learning model can be found in supplementary materials. The resulting learning rate (α), calculated separately for positive and negative feedback, was analyzed using an ANOVA, followed up by an independent sample t-test. Possible changes in the concurrent exploration parameter (β) between the two groups were assessed by an independent sample t-test.

For the ERN, the mean amplitude was calculated in an interval spanning 100 ms after response onset at electrode FCz. For the FRN, we used a similar 100 ms time interval (centered around the peak; 50 ms prior and 50 ms after it) and calculated the mean amplitude of this component at the same fronto-central electrode (see ref [8]). The FRN peak was defined as the most negative deflection arising at electrode FCz in the 230-350 ms time window following feedback onset. A mixed-model ANOVA was performed on the average mean amplitudes with group as between subjects and condition and response accuracy as within subject factors. In a second step, we computed difference waveforms by subtracting the ERP activity of incorrect from correct trials, separately for the ERN and FRN components, following standard practice [8]. The FCz electrode was selected based on previous work [8,10] showing the strongest expression of these two ERP components at this fronto-central location.

RESULTS

Behavioral results

The number of too late responses was modest (M = 3.45, SD = 1.83) and significantly higher for the MDD group than for the HC group (F (1, 77) = 9.51, p =.003, η_p^2 = .11, 95% CI [.02, .22]).

The analysis of the proportions of correct responses (Figure 1a) showed a significant Condition x Bin interaction (F(4.72, 363.30) = 31.92, p<.001, η_p^2 = .29, 95% CI [.22, .34]), as well as significant main effects of condition (F(2, 154) = 295.14, p<.001, η_p^2= .79, 95% CI [.75, .82]) and bin (F(2.74, 210.86) = 73.86, p<.001, η_p^2 = .49, 95% CI [.33, .48]). These effects translated a steep learning across time in the deterministic condition, lower and intermediate in the probabilistic condition, and with no such learning in the random condition. Groups did not differ significantly with respect to these gross accuracy scores, F (1, 77) = 1.68, p=.20, η_p^2 = .02, 95% CI [.00, .09]).

The analysis performed on the mean number of switches after negative feedback showed a significant Group x Bin interaction ($F(3, 231) = 3.47$, $p = .015$, $\eta_p^2 = .04$, 95% CI [.00, .08]; see Figure 1 b). Independent t-tests showed that in the first half of the task the difference between the two groups was not significant (t (77) = 0.25, $p = .804$, d = -0.082), while during the second half of the experimental session the MDD group ($M = 0.24$, $SD = 0.10$) had a lower number of switches after negative feedback compared to the HCs ($M = 0.30$, $SD = 0.10$), (t (77) = 2.88, $p = .013$, d = -0.6). There was a significant main effect of condition (F (2, 154) = 8.13, $p = .002$, $\eta_p^2 = .10$, 95% CI [.03, .17]), and bin ($F(3, 231) = 2.89$, $p = .034$, $\eta_p^2 = .04$, 95% CI [.00, .07]). Main effect of group was not significant (F (1, 77) = 1.82, $p = .181$, $\eta_p^2 = .023$, 95% CI [.00, .10]).

Clarity ratings (Figure 1c) showed a significant Group x Condition interaction (F (2, 154) = 3.04, $p = .051$, $\eta_p^2 = .04$, 95% CI [.00, .09]) and a main effect of condition (F (2, 154) = 311.70, $p < .001$, $\eta_p^2 = .80$, 95% CI [.76, .83]). Independent t-tests showed that in the deterministic condition, the HC group ($M = 77.09$, $SD = 11.33$) rated the S-R associations to be clearer than the MDD group ($M = 70.78$, $SD = 13.93$), (t (77) = 2.22, $p = .029$, d = 0.50). There was no significant group difference for the two other conditions (all p's > .05). Certainty ratings (Figure 1d) revealed a significant main effect of group (F (1, 77) = 5.23, $p = .025$, $\eta_p^2 = .06$, 95% CI [.00, .17]). Additionally, the HC group ($M = 40.73$, $SD = 10.67$) rated that they had received overall significantly more positive feedback than the MDD group ($M = 25.74$, $SD = 9.84$), (t (77) = 4.68, $p < .001$, d = 1, 47). The HC group ($M = 52.74$, $SD = 9.84$) also reported liking the positive feedback significantly more than the MDD group ($M = 44.39$, $SD = 23.73$), (t (77) = 2.12, $p = .037$, d = -0.48). The two groups did not differ significantly with respect to how much they disliked receiving negative feedback (t (77) = -1.27, $p = .208$, d = -0.29).

Computational modeling

For the learning rate, there was a significant main effect of feedback valence (F (1, 77) = 145.93, p<.001, η_p^2 = .66, 95% CI [.55, .72]) showing higher values following positive feedback (M = 0.32, SD = 0.23) than negative feedback (M = 0.04, SD = 0.08), replicating previous results [21]. The interaction with group was non-significant (F (1, 77) = 0.78, p=.380, η_p^2 = .01, 95% CI [.00, .07]), nor the main effect of group (F (1, 77) = 0.23, p=.631, η_p^2 = .003, 95% CI [.00, .09]). The group comparison performed on the inverse-gain parameter/exploration (β) revealed no significant effect (t (77) = 0.63, p=.532, d = 0.14).

ERP results

The analysis carried out on the ERN mean amplitudes showed a significant Condition x Accuracy interaction (F(1.84, 139.98) = 34.59, p<.001, η_p^2 = .31, 95% CI [.21, .40]), and main effects of condition (F(2,152) = 9.32, p <.001, η_p^2 = .11, 95% CI [.03, .18]) and accuracy (F(1,76) = 49.25, p<.001, η_p^2 = .39, 95% CI [.25, .50]). The main effect of group was not significant (F (1,76) = 0.90, p=.347, η_p^2 = .01, 95% CI [.00, .08]), (see Figure 2). As can be seen from the Table 2, the ERN was large and significant in the deterministic condition, intermediate in the probabilistic condition and merely absent in the random condition, with this (internal) reward probability effect being balanced between the two groups.

By comparison, for the FRN, the analysis revealed a significant Group x Accuracy x Condition interaction (F(2,138) = 3.84, p=.025, η_p^2 = .05, 95% CI [.06, .11]), as well as significant main effects of condition (F(2,138) = 22.45, p<.001, η_p^2 = .25, 95% CI [.10, .28]) and accuracy (F(1,69) = 10.32, p<.001, η_p^2 = .213, 95% CI [.09, .34]). The main effect of group was not significant (F (1,69) = 0.13, p=.718, η_p^2 = .00, 95% CI [.00, .06]). As can be seen from the Table 2, while reward probability yielded opposite effects on the ERN and FRN components for HCs (with the FRN effect being the highest for the random and probabilistic condition), MDD patients did not show the normal amplitude variation of the FRN depending on reward probability. When computing difference waves (i.e. negative – positive feedback),

we found that reward probability did influence the amplitude of the FRN in the HC group in the expected direction (F (2, 78) = 3.18, p=. 047, η_p^2 = .075, 95% CI [.00, .17]), while it did not in the MDD group (F (2, 52) = 1.37, p=. 26, η_p^2 = .050, 95% CI [.00, .15]). Strikingly, when the S-R association was probabilistic or random (and hence RL was more difficult to achieve), no reliable FRN effect was detected in this latter group (see Table 2). Importantly, this lack of normal (external) reward probability effect in MDD patients could not be imputed simply to noisy feedback-locked ERP waveforms in this group, as can be seen from Figure 3.

Relation to Anhedonia

We assessed whether these abnormal RL effects seen in MDD (i.e., switches after negative feedback and FRN) might be related to anhedonia severity in this sample. To this aim, we recalculated the ANOVAs presented here above using the SHAPS, TEPS, or the subscale of the BDI as covariate in separate analyses. None of these analyses showed significant results, however.

DISCUSSION

The MDD patients had more too late responses than the HCs, which is often reported in the literature [1, 2]. Yet, their learning slope and accuracy were similar to the HCs. Moreover, neither learning rate, nor exploration differed between the two groups. Noteworthy, an important difference between our study and previous ones is that monetary (or secondary) reward was often used [2, 34], while we did not do so in the present case. Our reward vs. punishment incentives were primarily related to the perceived task-success/failure (i.e., self-efficacy[28]), as opposed to secondary rewards or punishments, the former of which presumably activates more abstract motivational processes [5], and more dorsal prefrontal cortical areas than the latter [4,29].

Notwithstanding the lack of clear group differences for RL when it was assessed using standard quantitative measures, we found that MDD patients had a lower number of switches after negative feedback than HCs, during the second phase of the experimental session, selectively. This difference might stem from a different updating of trial history based on negative feedback in these two groups. MDD patients became more conservative than HCs, as demonstrated by their lower exploration of the alternative response option towards the end of the experiment. Remarkably, despite a learning performance that was matched with the HCs, these patients judged that they had received less often positive feedback (and they liked them less) throughout the experimental session than HCs (which was not the case obviously), unambiguously translating blunted positive affect at the subjective level. They also evaluated the clarity of the S-R associations in the deterministic condition to be lower than the HCs, and they felt overall less certain about the accuracy of their responses than the HCs.

Our ERP results show that while internal reward prediction error signals (at the ERN level) were overall spared in MDD patients relative to HCs, at the external, FRN level, when it was based on the processing of external evaluative feedback it was abnormal. For the probabilistic and random conditions, for which extra efforts needed to be exerted by the agent to learn the complex rule linking the actual R to the preceding S, the FRN was blunted, irrespective of anhedonia's severity . Previous studies [30,31] reported an overactive ERN for negative affect (MDD or anxiety), an effect that we failed to observe here. This discrepancy might be explained by the fact that interference tasks (such as Stroop or Flanker) were primarily used in these earlier studies, as opposed to RL in the present case, where error making acquires a different meaning (errors provide potent learning signals, as opposed to mere lapses of attention or concentration).

Lastly, we have to point out that these results were obtained in a cohort of MDD patients that were qualified as treatment resistant (because they were enrolled in a treatment

study using intermittent theta burst stimulation (iTBS) and treatment resistance was an inclusion criterion therein, [see 33]). This feature makes our results not immediately comparable to earlier studies where no such criterion was met. We also had to exclude some participants and patients because they failed to show normal RL at the behavioral level.

CONCLUSION

Our new results are compatible with recent theoretical accounts [12, 28], as well as older animal models [13], stating that MDD (and anhedonia) does not dampen reward processing per se, but instead it likely alters a core motivational component which in turn decreases or blunts the processing of the hedonic value of external reinforcers during RL. Abnormal RL as a function of MDD is confined to externally-based learning in the present case (switches after negative feedback and FRN), but not visible for internal error monitoring (ERN). Our findings suggest that ERN and FRN are dissociable since they are differentially sensitive to emotional disturbances accompanying MDD. We failed however to find evidence for an association with anhedonia severity.

In this context, clinical interventions meant to improve the timely processing of external evaluative feedback (self-efficacy related) might ultimately provide a valuable approach to reduce the burden of negative affect and distress in MDD.

Acknowledgements

The authors want to thank the MDD patients as well as HCs included in this study for their participation.

Financial support

This work is supported by the Belgian Science Policy, Interuniversity Attraction Poles program (P7/11). RDR and GP are funded by a Concerted Research Action Grant from Ghent

University (#BOF10/GOA/014). GP is supported in part by a 2015 NARSAD Independent Investigator Grant from the Brain & Behavior Research Foundation. This work was also supported by the Ghent University Multidisciplinary Research Partnership "The integrative neuroscience of behavioral control".

Conflict of interest

The authors have no conflict of interest to declare.

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

1. Pizzagalli DA. Depression, stress, and anhedonia: toward a synthesis and integrated model. Annu Rev Clin Psychol 2014;. 10(1): 393–423.
2. Pizzagalli DA, Iosifescu D, Hallett LA. et al. Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. Journal of Psychiatric Research 2009; 43(1): 76–87.
3. Vrieze E, Pizzagalli DA, Demyttenaere K, et al. Reduced reward learning predicts outcome in major depressive disorder. Biol Psychiatry 2013; 73(7): 639–645.
4. Sescousse G, Caldú X, Segura B, Dreher JC. Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. Neurosci. Biobehav Rev 2013; 37(4): 681–696.
5. Bandura A. Self-Efficacy: the exercise of control. Worth Publishers.1997
6. Locke EA, Latham GP. Building a practically useful theory of goal setting and task

motivation. A 35-year odyssey. Am Psychol 2002; 57(9): 705–717.

7. Holroyd C, Coles M. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol Rev 2002; 109(4): 679–709.
8. Eppinger B, Kray J, Mock B, Mecklinger A. Better or worse than expected? Aging, learning, and the ern. Neuropsychologia 2008; 46(2): 521–539.
9. Unger K, Heintz S, Kray J. Punishment sensitivity modulates the processing of negative feedback but not error-induced learning. Front Hum Neurosci 2012; 6:186.
10. Frank MJ, Woroch BS, Curran T. Error-related negativity predicts reinforcement learning and conflict biases. Neuron 2005; 47(4): 495–501.
11. Gründler TOJ, Cavanagh JF, Figueroa CM, et al. Task-related dissociation in ERN amplitude as a function of obsessive-compulsive symptoms. Neuropsychologia 2009; 47(8-9): 1978–1987.
12. Thomsen KR. Measuring anhedonia : impaired ability to pursue , experience, and learn about reward. Front Psychol 2015; 6(September): 1–11.
13. Salamone JD, Correa M, Farrar A, Mingote SM. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology 2007; 191(3): 461–482.
14. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Arlington. APPI 2013
15. 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. Jof Clin Psychiatry 1998; 59(SUPPL. 20): 22–33.
16. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc

Psych 1967; 6(4): 278–296.

17. Beck AT, Steer RA, Brown GK. Manual for the Beck depression inventory-II. San Antonio, TX: Psychological Corporation, 1996.
18. Snaith RP, Hamilton M, Morley S, et al. A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. Br J Psychiatry 1995; 167(1): 99–103.
19. Gard DE, Gard MG, Kring AM, John OP. Anticipatory and consummatory components of the experience of pleasure: A scale development study. J Res Pers 2006; 40(6): 1086–1102.
20. Rush AJ, Thase ME, Dubé S. Research issues in the study of difficult-to-treat depression. Biol Psychiatry 2003; 53(8): 743–753.
21. Bakic J, Jepma M, De Raedt R, Pourtois G. Effects of positive mood on probabilistic learning: behavioral and electrophysiological correlates. Biol Psychol 2014; 103: 223–232.
22. Aarts K, Pourtois G. Anxiety not only increases, but also alters early error-monitoring functions. Cogn Affect Behav Neurosci, 2010; 10(4): 479–92.
23. Pourtois G. Early error detection predicted by reduced pre-response control process: An ERP topographic mapping study. Brain Topog 2011; 23(4): 403–422.
24. Aarts K, Vanderhasselt M, Otte G, et al. Electrical brain imaging reveals the expression and timing of altered error monitoring functions in major depression. J Abnorm Psychol 2013; 122(4): 939–50.
25. Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Clin Neurophysiol 1983; 55(4): 468–484.
26. Bellebaum C, Kobza S, Ferrea S, et al. Strategies in probabilistic feedback learning in Parkinson off medication. Neuroscience 2016; 320: 8–18.
27. Harlé KM, Zhang S, Schiff M, Mackey S, et al. Altered statistical learning and

decision-making in methamphetamine dependence: evidence from a two-armed bandit Task. Front Psychol 2015; 6: 1910.

28. Treadway MT. The neurobiology of motivational deficits in depression - an update on candidate pathomechanisms. Curr Top Behav Neurosci 2015; 1–19.

29. Badre D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci 2008; 12(5): 193–200.

30. Endrass, T., & Ullsperger, M. (2014). Specificity of performance monitoring changes in obsessive-compulsive disorder. Neuroscience and Biobehavioral Reviews, 46 Pt 1, 124–38. http://doi.org/10.1016/j.neubiorev.2014.03.024

31. Weinberg, A., Riesel, A., & Hajcak, G. (2011). Integrating multiple perspectives on error-related brain activity: The ERN as a neural indicator of trait defensive reactivity. Motivation and Emotion, 36(1), 84–100. http://doi.org/10.1007/s11031-011-9269-y

32. Duprat, R., Desmyter, S., Rudi, D. R., Van Heeringen, K., Van Den Abbeele, D., Tandt, H., Bakic, J., et al. (2016). Accelerated intermittent theta burst stimulation treatment in medication-resistant major depression: A fast road to remission? Journal of Affective Disorders, 200, 6-14.

33. Treadway MT, Bossaller NA, Shelton RC, Zald DH.Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. J Abnorm Psychol 2012; 121(3): 553–558.

Table 1. Demographic and Clinical Data. Means (standard deviations) are provided. Independent samples t-test differences are provided for HRSD (df = 77), BDI II (df = 72), Anhedonia subscale of BDI II (df = 77), TEPS with the corresponding subscales (df = 74), and SHAPS (df = 77).

	HC	MDD	t-test	d
N	44	35		
Age	37.89 (12.23)	43.00 (11.67)	-1.88	-0.43
Sex	28F/16M	27F/8M		
	*χ^2 = 1.68, p=.23			
Age at onset		24.6 (11.03)		
Lenght of episode (months)		20.81 (32.05)		
Number of episodes		3.14 (2.61)		
HRSD	1.42 (2.37)	21.83 (5.63)	-21.79**	-4.93
BDI_II	5.98 (6.75)	30.21 (10.27)	-12.16**	-2.86
Anhedonia	0.98 (1.37)	4.66 (2.25)	-8.97**	-2.03
TEPS	75.02 (19.22)	58.97 (17.04)	3.81**	0.88
Consumatory	36.05 (9.57)	28.76 (9.02)	3.39**	0.78
Inhibitory	38.89 (10.94)	30.21 (8.95)	3.76**	0.89
SHAPS	0.55 (2.16)	7.31 (4.09)	-9.45**	-2.14

*p<.05, **p<.01

Table 2. Mean ERP activity (1 standard deviation) for each condition and accuracy level, separately for each component and group. Results of the direct pairwise comparisons (degrees of freedom: 43) between the two accuracy levels (correct vs. incorrect), using post-hoc t-tests. * indicates that p-values were Bonferroni corrected for multiple testing (p = .008).

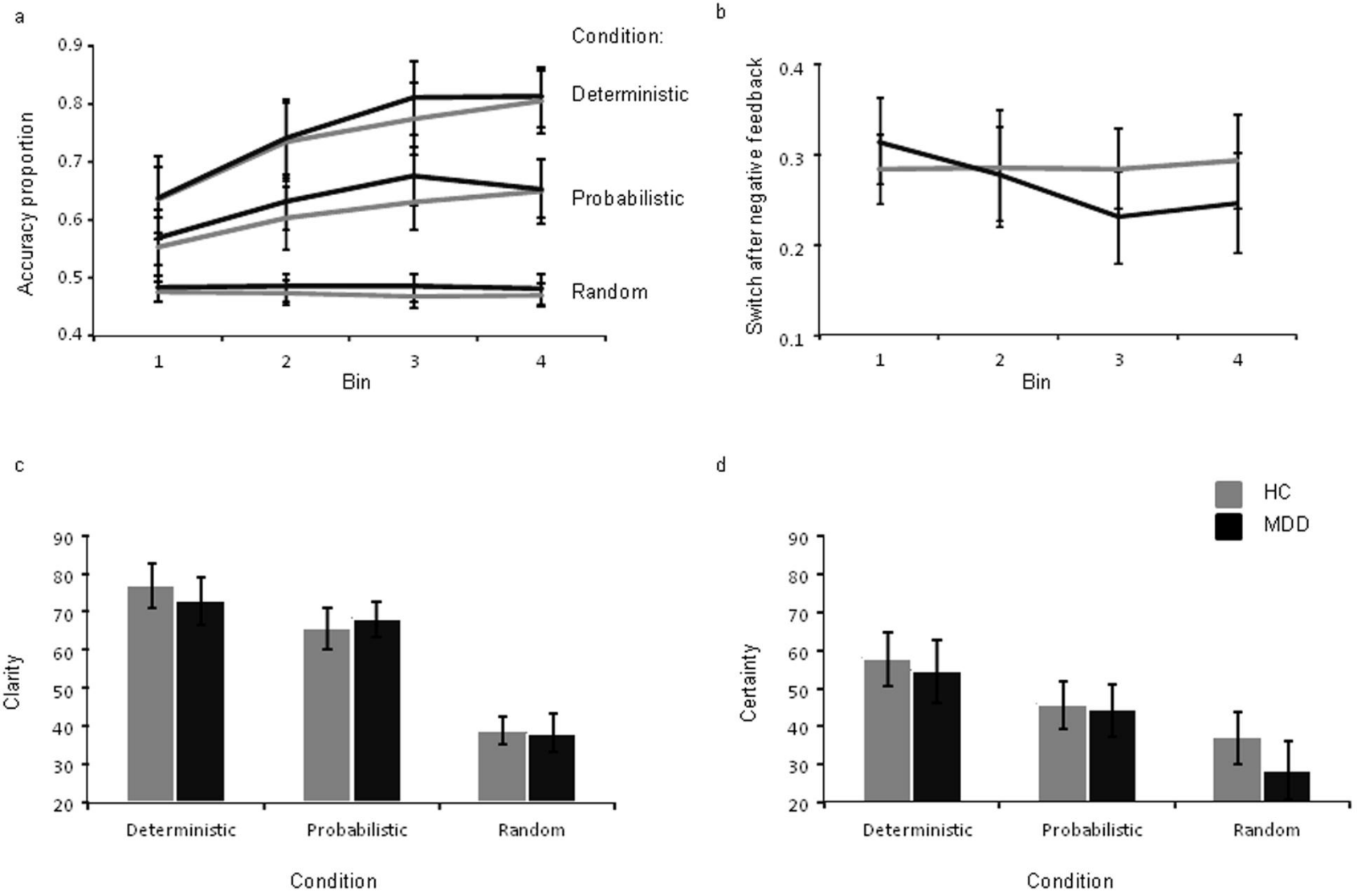
ERP component	Condition	Group					
		HC			MDD		
ERN		Correct	Incorrect	t-test	Correct	Incorrect	t-test
	Deterministic	-1.73	-3.89	5.97*	-1.39	-3.62	5.71*
		(4.33)	(4.79)		(3.84)	(4.64)	
	Probabilistic	-2.25	-2.52	1.18	-1.62	-2.00	1.57
		(4.37)	(4.58)		(3.98)	(3.70)	
	Random	-2.95	-2.68	-1.31	-1.95	-2.03	0.43
		(4.41)	(4.27)		(3.48)	(3.12)	
FRN							
	Deterministic	0.47	0.35	-0.65	0.90	0.29	1.76
		(2.10)	(1.97)		(2.11)	(2.68)	
	Probabilistic	1.11	0.29	2.84*	1.24	1.02	0.71
		(2.34)	(3.28)		(2.58)	(2.90)	
	Random	1.60	1.03	2.91*	1.60	1.59	0.77
		(2.10)	(2.09)		(2.74)	(2.98)	

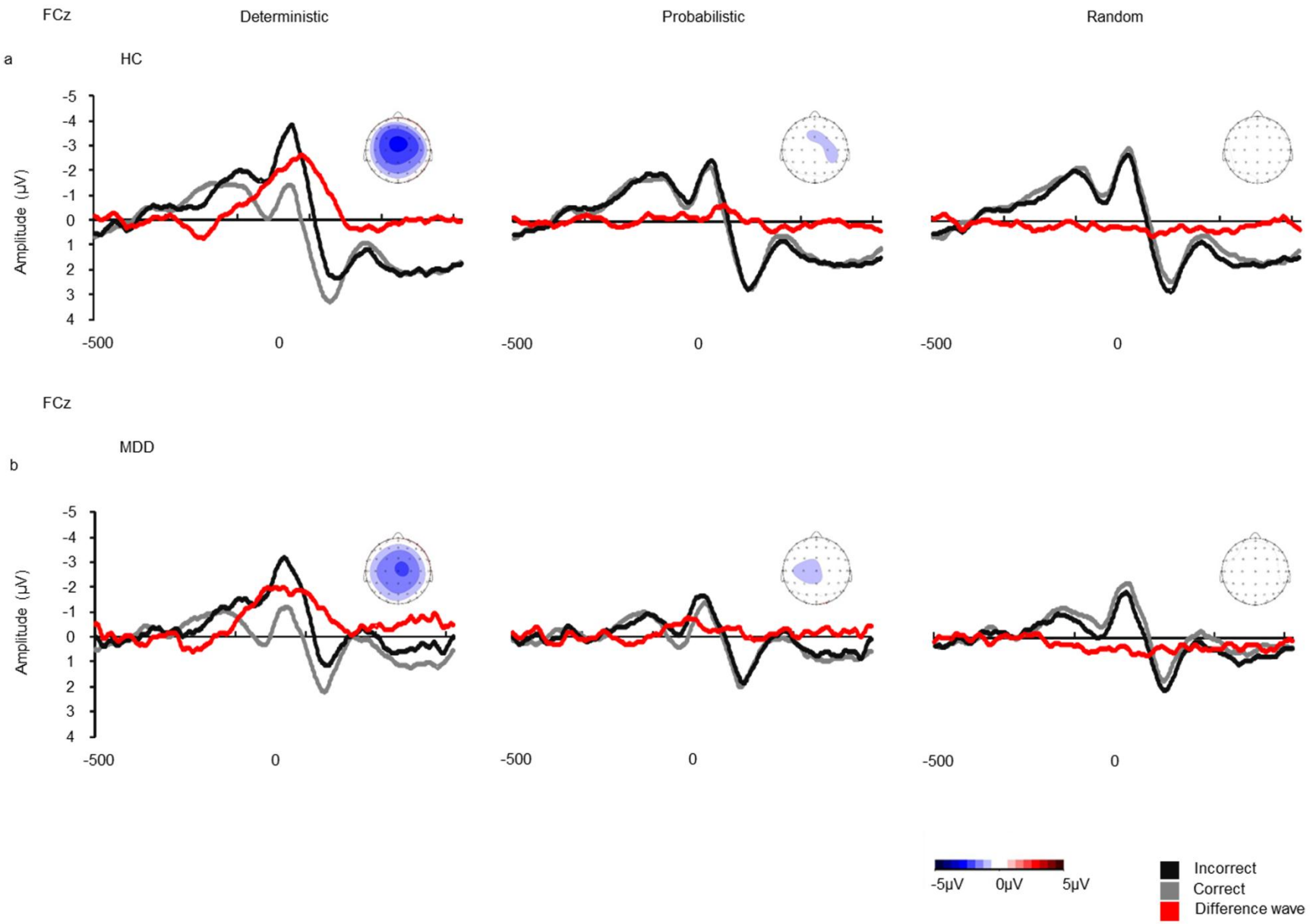
FIGURES LEGEND

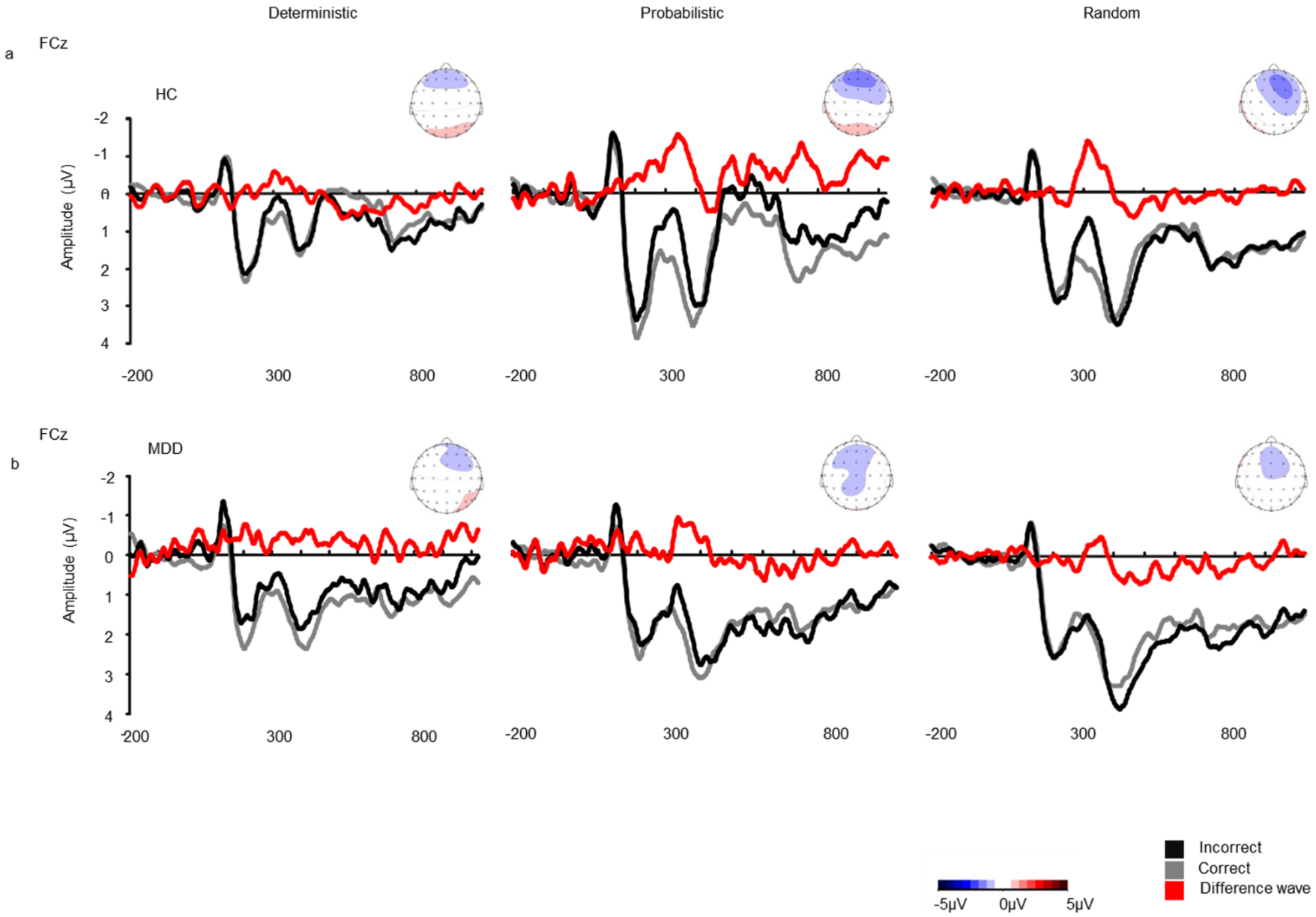
Figure 1. a) Accuracy data (i.e. proportion of correct responses) decomposed as a function of bin, condition and group. b) Mean number of switches after negative feedback (expressed here in proportion) decomposed as a function of bin and group. c) Clarity and d) Certainty ratings decomposed as a function of condition and group.

Figure 2. Grand average ERP waveforms and topographical maps (top view) for the response-locked ERP data (electrode FCz), separately for each condition and accuracy level, for a) HCs b) MDD patients

Figure 3. Grand average ERP waveforms and topographical maps (top view) for the feedback-locked ERP data (electrode FCz), separately for each condition and accuracy level, for a) HCs b) MDD patients.


a
Condition:
Deterministic
Probabilistic
Random
Accuracy proportion
0.9
0.8
0.7
0.6
0.5
0.4
1
2
3
4
Bin
b
Switch after negative feedback
0.4
0.3
0.2
0.1
1
2
3
4
Bin
c
Clarity
90
80
70
60
50
40
30
20
Deterministic
Probabilistic
Random
Condition
d
HC
MDD
Certainty
90
80
70
60
50
40
30
20
Deterministic
Probabilistic
Random
Condition


FCz
Deterministic
Probabilistic
Random
a
HC
Amplitude (μV)
-5
-4
-3
-2
-1
0
1
2
3
4
-500
0
FCz
MDD
b
-5μV
0μV
5μV
Incorrect
Correct
Difference wave


Deterministic
Probabilistic
Random
a
FCz
HC
Amplitude (μV)
-2
-1
0
1
2
3
4
-200
300
800
b
FCz
MDD
-5μV
0μV
5μV
Incorrect
Correct
Difference wave