

Neural predictors of treatment response to brain stimulation and psychological therapy in depression: a double-blind randomized controlled trial

Camilla L Nord^{1,2*}, D Chamith Halahakoon^{1,3}, Tarun Limbachya⁴, Caroline Charpentier^{1,5}, Níall Lally^{1,6,7}, Vincent Walsh¹, Judy Leibowitz⁴, Stephen Pilling^{4,8}, Jonathan P Roiser¹

Affiliations:

¹ Institute of Cognitive Neuroscience, University College London, London, UK.

² MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK.

³ Department of Psychiatry, University of Oxford, Oxford, UK

⁴ Camden and Islington NHS Foundation Trust, London, UK.

⁵ Division of the Humanities and Social Sciences, California Institute of Technology, California, USA.

⁶ Warwick Medical School, University of Warwick, Coventry, UK.

⁷ Experimental Therapeutics & Pathophysiology Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland, USA.

⁸ Department of Clinical, Educational, and Health Psychology, University College London, London, UK.

*To whom correspondence should be addressed:

Dr. Camilla L Nord

Address: MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, CB2 7EF Cambridge, UK

Email: Camilla.Nord@mrc-cbu.cam.ac.uk

Phone: +44 223355294

Word count: 3997

No. references: 45

No figures and tables: 5

Abstract: Standard depression treatments, including antidepressant medication and cognitive behavioural therapy (CBT) are ineffective for many patients. Prefrontal transcranial direct current stimulation (tDCS) has shown inconsistent efficacy as an experimental treatment for depression, and its mechanisms are poorly understood. We recruited unmedicated patients with major depressive disorder (N=71 approached; N=39 randomised) for a mechanistic, double-blind, randomized controlled trial consisting of eight weekly sessions of prefrontal tDCS administered to the left prefrontal cortex prior to CBT. We probed (1) whether tDCS improved the efficacy of CBT relative to sham stimulation; and (2) whether neural measures predicted clinical response. We found a modest and non-significant effect of tDCS on clinical outcome over-and-above CBT (active: 50%; sham: 31.6%; odds ratio: 2.16, 95% CI=0.59—7.99), but a strong relationship, predicted *a priori*, between baseline activation in the stimulated prefrontal region and symptom improvement. Repeating our analyses of symptom outcome splitting the sample according to this biomarker revealed that tDCS was significantly superior to sham in individuals with high L-DLPFC activation at baseline; we also show 86% accuracy in predicting clinical response using this measure. Exploratory analyses revealed several other regions where activation at baseline was associated with subsequent response to CBT, irrespective of tDCS. This mechanistic trial revealed variable, but predictable, clinical effects of prefrontal tDCS combined with CBT for depression. We have discovered a potential explanation for this variability: individual differences in baseline activation of the region stimulated. Such a biomarker could potentially be used to pre-select patients for trials and, eventually, in the clinic.

Trial Registration: The trial was registered at clinicaltrials.gov, registration number NCT01875419; URL: <https://clinicaltrials.gov/ct2/show/NCT01875419?term=NCT01875419&rank=1>

Introduction

Cognitive behavioural therapy (CBT) is a common, effective treatment for depression[1]. Nevertheless, only 60% of patients show an adequate response to therapy; even fewer achieve remission[2]. Some researchers have suggested improving CBT response with augmentative strategies, such as cognitive-enhancing drugs[3]. More recently, noninvasive brain stimulation, including repetitive transcranial magnetic stimulation (rTMS)[4] and transcranial direct current stimulation (tDCS)[5], has been suggested as a strategy to enhance response to CBT.

Both rTMS and tDCS have been applied as treatments for depression over the dorsolateral prefrontal cortex (DLPFC), a region disrupted in both currently-depressed[6] and remitted patients[7,8]. Stimulation of the DLPFC directly affects its activation, as well that of distal regions including the cingulate cortex[9,10]. Previous trials have shown that prefrontal tDCS has a moderately strong antidepressant effect[11–16].

Recently, a large multi-centre trial failed to show any evidence in support of tDCS, even finding a potentially detrimental effect for unipolar depression[17]. This inconsistency echoes non-clinical experiments. Initially, many results suggested that prefrontal tDCS enhances cognitive performance, including planning[18], insight [19], and cognitive control during and after tDCS, in both healthy[20,21] and depressed[22] individuals. However, the cognitive effects of tDCS also appear highly variable [21,23,24]; a quantitative review suggested there were no consistent effect of tDCS on any cognitive process in healthy individuals[23] (for a critique of this review, see[25]).

If prefrontal tDCS improves cognitive control, it may improve patients' ability to benefit from CBT, which itself entails challenging cognitive processing, including planning and working memory[1]. This is supported by preliminary evidence that tDCS enhances the antidepressant effects of cognitive control training[26] and, in a small feasibility trial (N=10), internet-based CBT[27]. However, given the inconsistent evidence for tDCS in both cognition[23] and depression[17], it is also likely there are substantial individual differences in response to tDCS-enhanced CBT.

A key priority for psychiatric research is identifying objective measures to help optimize treatment selection[28]. There is preliminary evidence that pre-treatment neural activation (measured with functional magnetic resonance imaging: fMRI) can predict response to antidepressant medication and CBT[29]. However, studies investigating such associations invariably lack comparison groups and randomization to groups. To identify biomarkers unambiguously, mechanistic measurements need to be incorporated into randomized controlled trials (RCTs).

A strong candidate biomarker for response to noninvasive brain stimulation in depression is activation of the region stimulated (usually the DLPFC): experimentalists have hypothesized that baseline DLPFC activity might mediate the cognitive effects of prefrontal tDCS on cognition[24]; and one small study in depressed patients reported that higher pre-treatment DLPFC resting-state perfusion was associated with better clinical response to DLPFC rTMS[30].

Therefore we conducted a mechanistic, proof-of-principle RCT of prefrontal tDCS to augment CBT in unmedicated patients with unipolar depression. We hoped to capitalise on the sustained (~90 minute) aftereffects of tDCS on cortical excitability by delivering stimulation immediately prior to weekly CBT sessions[31]. Our study had two key aims: 1) to assess whether tDCS augments the ability of CBT to treat depression over and above sham (placebo) stimulation; and 2) to explore putative pre-treatment biomarkers. We also aimed to identify the cognitive and neural changes resulting from combined CBT and tDCS. We hypothesized that 1) tDCS would improve response to CBT; and 2) that higher pre-treatment activation in the DLPFC would predict better response to combined CBT and tDCS.

Materials and Methods

Participants

We randomized 39 unmedicated patients with major depressive disorder (Hamilton Rating Scale for Depression (HAM-D)>17) to receive real (N=20) or sham (N=19) tDCS. Patients were recruited through the Camden and Islington NHS Foundation Trust Improving Access to Psychological Treatment (IAPT) service within the UK National Health Service (NHS). Patients, investigators, and therapists were blind to tDCS condition for the duration of the trial. Six patients did not complete the intervention (defined as $\geq 7/8$ sessions: 1 active; 5 sham); dropout rates **differed marginally (non-significantly)** between groups ($p=0.091$); to account for potential non-random dropout we employed intention-to-treat analyses as our primary analysis. Thirty-one patients received eight sessions of CBT and two patients received seven sessions of CBT. See Figure 1 for CONSORT diagram and procedure.

Only patients meeting criteria for at least moderate depression where the IAPT service assigned one-on-one cognitive behavioural therapy (CBT) for depression were included in the trial (see Supplemental Materials for full participant criteria). The study was approved by the London Queen Square NHS Research Ethics Committee (ID: 13/LO/1028); subjects' consent was obtained in accordance with the Declaration of Helsinki.

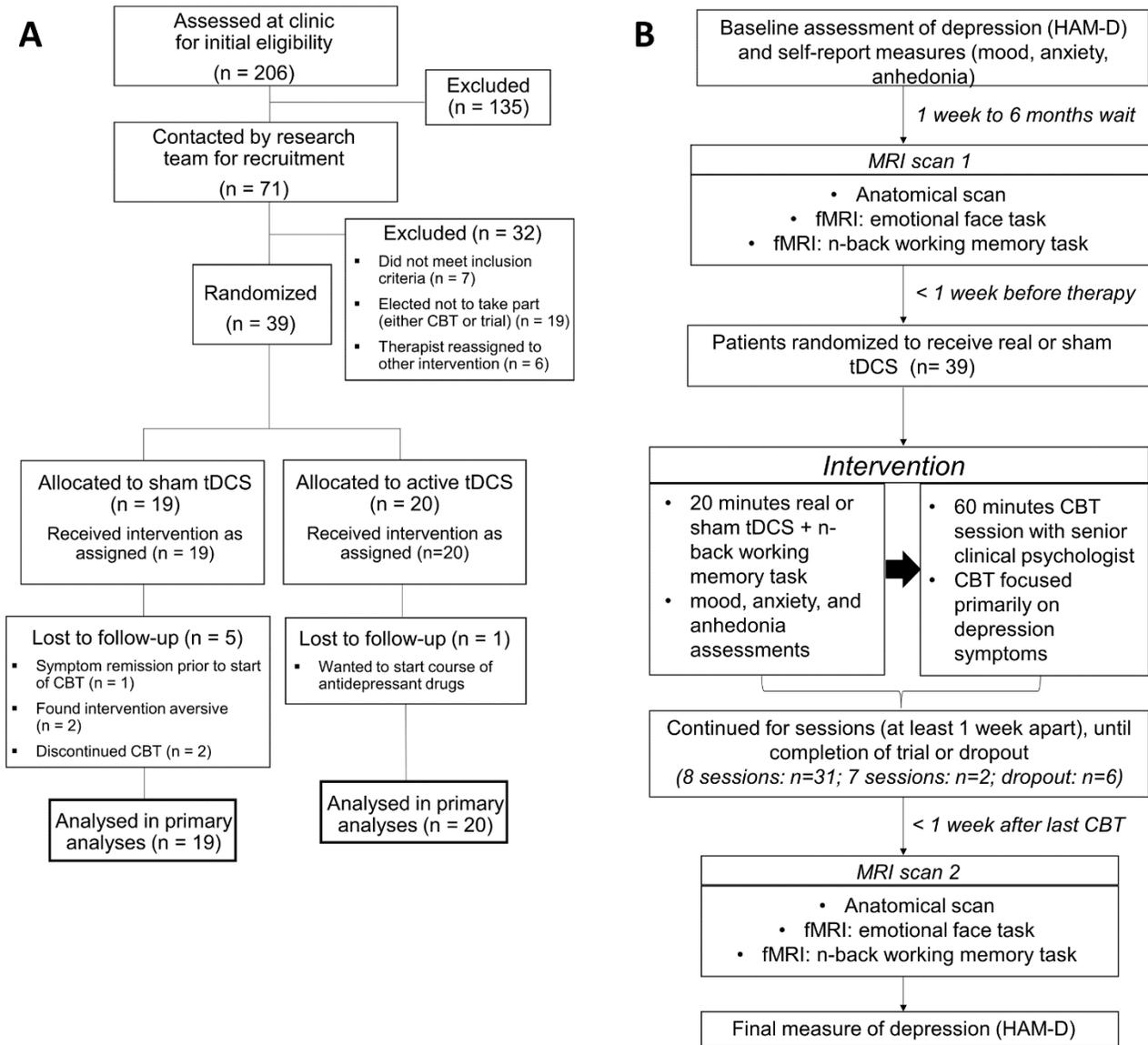


Fig 1. CONSORT diagram describing recruitment, randomization, and attrition in the clinical trial (A) and depiction of trial procedure (B). CBT=cognitive behavioural therapy; tDCS=transcranial direct current stimulation; HAM-D=Hamilton Rating Scale for Depression; fMRI=functional magnetic resonance imaging.

Protocol

The study protocol was registered on *clinicaltrials.gov* (NCT01875419). See Figure 1A for CONSORT diagram. See Figure 1B for trial protocol, consisting of a screening session, fMRI scans (before and after intervention), and randomization (using custom-written MATLAB code by a researcher not involved in the trial). See Supplemental Materials for the full protocol.

Imaging details

The MRI protocol included a T1-weighted anatomical scan and two T2*-weighted functional scans (n-back working memory task [21], practiced at the screening session to ensure participants could perform the task; and an emotion processing task [32] where participants discriminated the gender of fearful, happy, and neutral faces). See Supplemental Materials for full details.

Stimulation procedure

Immediately preceding each CBT session, a 1mA constant current fronto-extracerebral montage was delivered to the left prefrontal cortex (anode on F3 on the international 10-20 electroencephalography (EEG) system (using an EEG cap for placement), secured in place with an elasticated head-strap; cathode on ipsilateral deltoid). Although this delivers relatively diffuse prefrontal stimulation[33]), [an open-label, preliminary study comparing bifrontal and fronto-extracerebral tDCS in depression found that frontoextracerebral tDCS was more effective\[34\] motivating our choice \(note, however, that this study may have been underpowered: N=11\)](#). We were restricted to 1mA by our ethics committee; however, a recent study found that only unilateral 1mA stimulation (and not bilateral 1mA or 2mA) caused sustained increases in corticospinal excitability[35].

The stimulation, delivered using a neuroConn DC-stimulator, lasted for 20 minutes. Sham stimulation involved 30 seconds of direct current followed by 1,170 seconds without, which produces a similar sensation to active stimulation (most sensations are felt in the first 30 seconds during ramping up), resulting in effective blinding at 1mA[36]. [Both electrodes were 35 cm² \(7 cm x 5 cm\), rubber, and placed inside electrode sponges that had been briefly soaked in saline to conduct the current.](#) The polarity of the device was independently verified using a resistor and volt-meter. For practicality, tDCS was administered prior to CBT. Following stimulation, patients completed the Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), and Snaith-Hamilton Pleasure Scale (SHAPS).

Patients performed the n-back working memory task during stimulation (in real and sham stimulation conditions), as previous studies reported enhanced cognitive effects of tDCS when delivered concurrently with a working memory paradigm[26]. The n-back task was identical to the version patients were previously trained on and performed inside the scanner.

Therapy procedure

Immediately following each tDCS session, patients received a one-hour CBT intervention for depression, as described previously[37]. Sessions occurred a minimum of six days apart. We decreased variability in CBT delivery by working with a selected team of therapists, doctoral-level clinical psychologists working in the NHS IAPT service, which itself has standardised training and CBT delivery. To ensure adherence to protocol, therapists also received regular supervision of CBT sessions by supervisors accredited by the British Association of Behavioural and Cognitive Psychotherapists (BABCP). To increase consistency of CBT, we ran a day-long expert-led training for all therapists before recruitment.

Side effects and blinding

At each session, following tDCS, patients filled out the tDCS Adverse Events Questionnaire [38], and guessed their stimulation condition. At four sessions, the Adverse Events questionnaire was not completed due to time constraints.

Analyses

Our primary clinical outcome measure was response rate according to the HAM-D, defined as a reduction of at least 50% from baseline; we also report remission rate, defined as a HAM-D score of 7 or below. We analysed our primary outcome with an intent-to-treat analysis (accounting for missing values at follow-up using last-observation-carried-forward: N=39) as we could not assume attrition occurred completely at random. We analysed secondary outcome measures (BDI, BAI, SHAPS, and working memory performance d' (previously described [21])) with linear mixed models, which accounts for missing data by estimating the trajectory of change for the subjects who did provide data (see Supplemental Materials for details).

For all measures, we repeated this analysis using the sample of patients who completed the treatment.

fMRI analyses (on sample who completed) were performed in SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/>). Our primary analyses used regions of interest (ROIs): left and right amygdala, subgenual anterior cingulate cortex, and L-DLPFC defined anatomically (emotion processing task; contrasts: fearful>neutral faces and happy>neutral faces); and R- and L-DLPFC (n-back task). L-DLPFC was identified by placing a 10mm sphere on coordinates from a meta-analysis of working memory tasks, the majority of which used the n-back (-44,20,30)[6]; R-DLPFC was defined using the corresponding contralateral coordinate (44,20,30). We additionally conducted an exploratory whole-brain flexible factorial analysis (with an initial cluster-forming threshold of $p < 0.001$, uncorrected). [For these exploratory analyses, we use a Bonferroni-corrected significance threshold of \$p = 0.025\$ \(FWE-corrected\), reporting significance at the cluster- or voxel-level.](#)

Outcome prediction analysis

We first mean-corrected all predictor variables: R- and L-DLPFC activation; R- and L-amygdala activation; sgACC activation; n-back d' ; SHAPS; BAI; and HAM-D. Note all predictor variables were acquired prior to randomisation.

Our primary analyses tested whether pre-randomization activation averaged within each ROI (the L-DLPFC for the n-back task; the amygdalae and sgACC for the emotional faces task; all analysed in separate regressions) could predict the degree of clinical improvement (HAM-D percent change). Each regression also included the fixed effect of stimulation condition and [the interaction between stimulation condition and baseline ROI activation](#). If this model was significant, we then conducted sensitivity analyses that tested whether the prediction remained significant following inclusion of additional variables in the model (HAM-D, BAI, and SHAPS

scores, as well as n-back performance, [and their interactions with stimulation group](#)) (see Supplemental Materials for details).

Power calculation

This was a mechanistic, proof-of-principle trial powered to detect a large effect size, as reported in a previous trial of tDCS in antidepressant-free patients (Hedges' $g=0.988$)[11] and report of tDCS enhancing n-back performance (Cohen's $d=1.42$)[21]. We had 80% power to detect $w=0.45$ (X^2 test; $\alpha=0.05$) for our primary clinical outcome variable (HAM-D), which was binary. For continuous variables, including our mechanistic (fMRI) measures, we had 80% power to detect $d=1.0$ (independent samples t-test; $\alpha=0.05$ (two-tailed)).

Results

Clinical and demographic data are presented in Table 1. There were no differences in baseline generalised anxiety disorder (GAD) between patients subsequently assigned to real and sham conditions: 14/20 patients assigned to real stimulation, and 13/19 patients assigned to sham stimulation, met criteria for GAD ($X^2=0.011$, $p=0.915$).

Table 1. Participant characteristics by stimulation condition.

	active tDCS	sham tDCS
N	20	19
% female	45.00	57.89
Age	35.60 (12.91)	31.05 (8.17)
FSIQ	110.10 (7.49)	105.15 (7.54)
Baseline HAM-D	21.95 (3.20)	21.05 (3.27)
Baseline BDI	25.70 (8.01)	27.79 (5.34)
Baseline SHAPS	19.85 (7.25)	18.05 (7.25)
Baseline BAI	25.70 (13.57)	25.47 (12.06)
Age of onset	22.80 (10.09)	18.37 (8.23)
No. episodes	2.50 (1.67)	3.05 (1.58)
% hospitalized	15.00	15.79
% suicide attempt	30.00	21.05
% past ADM	55.00	42.10
% past PT	50.00	78.90
% baseline GAD	75	68

Other than N and percentage data, figures represent means (standard deviations); FSIQ = Full Scale Intelligence Quotient (calculated from Weschler Test of Adult Reading); HAM-D = Hamilton Rating Scale for Depression; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; SHAPS = Snaith-Hamilton Pleasure Scale; No. = number; % past ADM = percent of patients with previous history of antidepressant medication use (no patients were currently-medicated); % past PT = percent of patients with a previous history of psychological therapy (all patients were engaged in CBT during the trial); % baseline GAD=percent of patients who met criteria for generalised anxiety disorder (GAD) at baseline.

Assessment of blinding and side effects

38.5% of patients correctly guessed the stimulation condition, with no significant difference in the proportion of active and sham guesses between groups ($X^2=1.64$, $p=0.301$), suggesting that blinding was effective. See Supplemental Materials details and full side-effects.

Primary clinical outcome

Using an intention-to-treat analysis (last observation carried forward), a greater proportion of patients responded (active: 50%; sham: 31.6%; odds ratio: 2.16, 95% CI=0.59—7.99) and remitted (active: 30%; sham: 10.5%; odds ratio: 3.65, 95% CI=0.63—20.96) following CBT with active than with sham tDCS (defined as a 50% or greater reduction on the Hamilton Rating Scale for Depression, HAM-D, and a HAM-D of less than 7, respectively). However, these differences did not achieve statistical significance (response: $X^2=1.37$, $p=0.24$, $\varphi=0.187$; remission: $X^2=2.27$, $p=0.13$, $\varphi=0.241$ (see Figure 2A). We repeated these analyses for the sample of patients who completed the trial (response: active: 52.6%; sham: 42.9%; odds ratio: 1.48; $X^2=0.308$, $p=0.579$, $\varphi=0.184$; remission: active: 31.6%; sham: 14.3%; odds ratio: 2.76; $X^2=1.313$, $p=0.252$, $\varphi=0.090$).

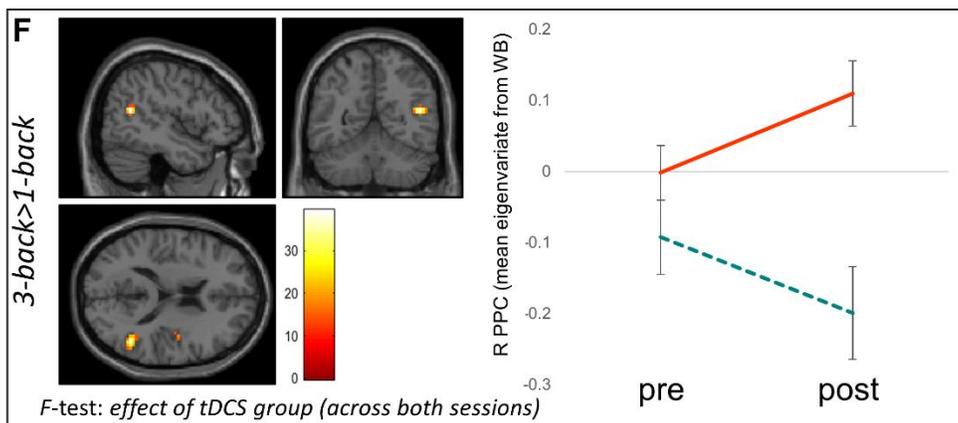
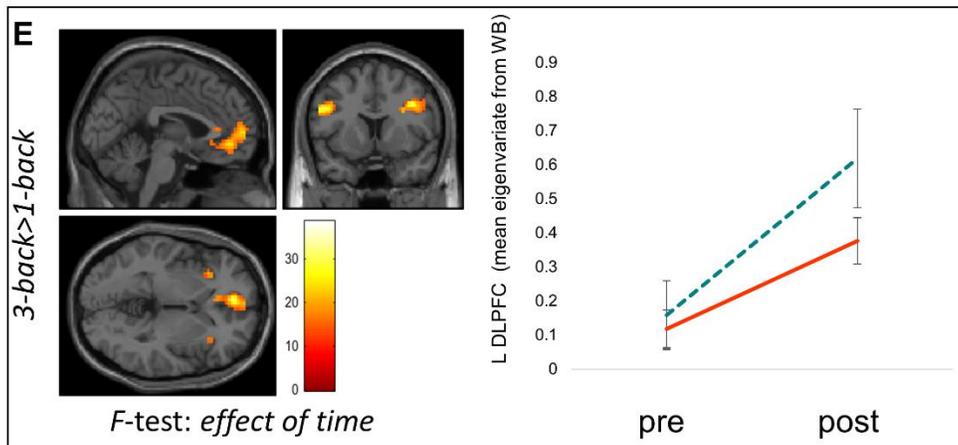
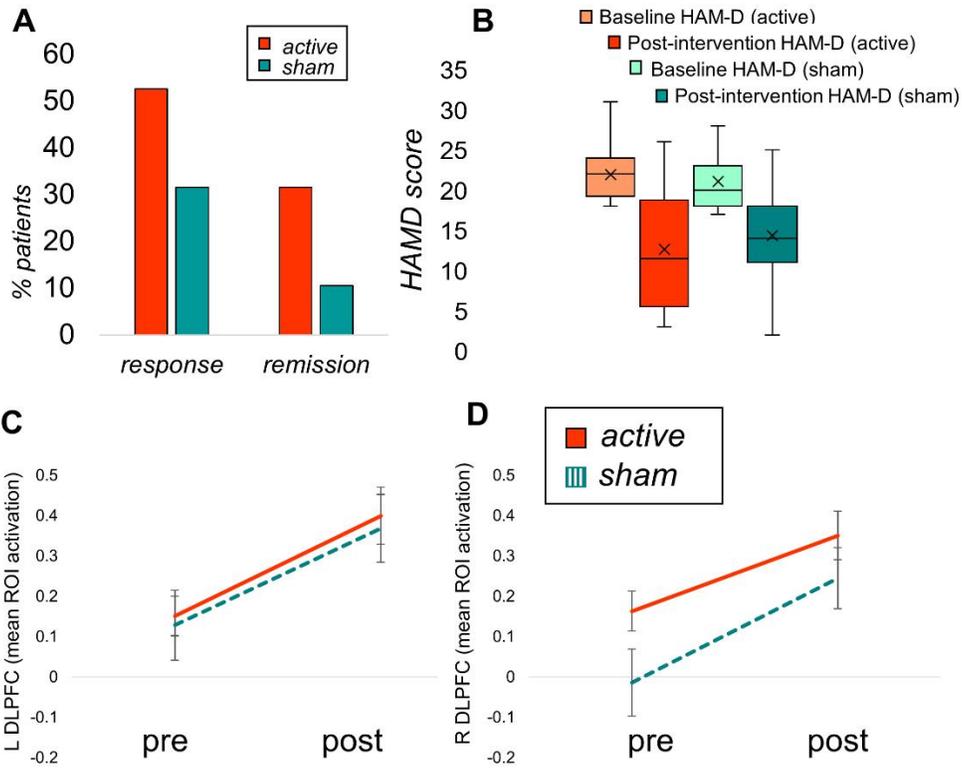


Fig 2. Effect of intervention on depressive symptoms and brain activation. A: Proportion of patients responding and remitting following CBT with active and sham tDCS. These differences did not achieve statistical significance. B: HAM-D scores reduced significantly from pre- to post-intervention, but the main effect of stimulation condition and the interaction were non-significant. C&D: **Effect of intervention on neural activation during working memory** (contrast: 3-back > 1-back). Activation in left (C) and right (D) dorsolateral prefrontal cortex (DLPFC) *a priori* regions-of-interest pre- and post-intervention (orange solid lines = active tDCS; green dotted lines = sham tDCS). E: Whole brain (WB) analysis of the effect of time (pre- versus post-intervention) and plotted eigenvariate of this main effect. E, left panel: we found significant changes in activation over time in bilateral DLPFC (increased over time: coronal section; $p < 0.001$ (uncorrected) cluster-forming threshold, WB cluster-level FWE-corrected) and medial PFC/perigenual anterior cingulate cortex (decreased over time: sagittal and axial sections; $p < 0.001$ (uncorrected) cluster-forming threshold, WB cluster-level FWE-corrected). E, right panel: for illustrative purposes, eigenvariate values extracted from the L-DLPFC cluster displayed in the left panel are plotted. F: WB effect of group (greater in active than sham tDCS condition) in the right posterior parietal cortex (rPPC) ($p = 0.02$ WB voxel-level FWE corrected; F, left), and plotted eigenvariate of this rPPC cluster (F, right). Overlays are displayed at $p < 0.001$ (uncorrected) and colour bars indicate F-values. Error bars indicate standard errors of the mean.

Secondary clinical outcomes

HAM-D scores reduced significantly from pre- to post-intervention ($F(1,37) = 56.09$, $p < 0.001$, $\eta_p^2 = 0.603$) (mean difference 9.30 (SD=6.82) in active, 6.68 (SD=6.49) in sham), but the main effect of stimulation condition ($F(1,37) = 0.112$, $p = 0.739$) and the stimulation-by-time interaction ($F(1,37) = 1.50$, $p = 0.228$) were non-significant (Figure 2B). The results were similar for completers (N=33). Including therapist as a covariate in this analysis produced similar results (see Supplemental Materials). There were no significant effects of stimulation condition on questionnaire measures or working memory performance (all $p > 0.1$) (see Supplemental Materials for full results, and in-scanner task performance)

Neural activation: n-back working memory task

ROI analysis (*a priori*): Bilateral DLPFC activation increased after the intervention (main effect of time: $F(1,31) = 20.95$, $p < 0.001$, $\eta_p^2 = 0.403$), but the interaction with group was non-significant ($F(1,31) = 0.363$, $p = 0.551$) (Figure 2). At baseline patients showed significantly lower bilateral DLPFC activation than 30 matched healthy controls ($F(1,61) = 8.24$, $p = 0.006$, $\eta_p^2 = 0.119$), but there was no significant difference from healthy controls post-intervention ($F(1,61) = 0.191$, $p = 0.664$). There was no association between increased bilateral DLPFC activation and n-back improvement ($r(33) = 0.064$, $p = 0.725$). Full results in Supplemental Materials.

Whole-brain analysis (exploratory): Activation increased significantly from pre- to post-intervention in bilateral parietal cortex and bilateral DLPFC, and decreased significantly in medial prefrontal cortex, extending into perigenual ACC (all $p < 0.001$, cluster-level FWE-corrected; Figure 2E and Supplemental Table 2). Patients receiving active stimulation also had significantly greater activation in right posterior parietal cortex (rPPC) compared to sham ($p = 0.02$ voxel-level FWE-corrected; Figure 2F). rPPC activation increased over time in the active condition ($t(18) = 2.96$, $p = 0.008$) but decreased numerically in the sham condition ($t(13) = 1.98$, $p = 0.069$) (group-by-time interaction: $F(1,31) = 11.74$, $p = 0.002$; Figure 2F). Note that since the contrast used to identify this result was the main effect of group, this group-by-time interaction is orthogonal, meaning this is not a circular analysis.

Neural activation: emotion processing task

Amygdala ROI analysis (*a priori*): We found a significant time-by-simulation condition interaction in the amygdala ($F(1,31)=5.04$, $p=0.032$, $\eta_p^2=0.140$): those in the sham group showed decreased amygdala activation (averaged across contrasts) at post- relative to pre-intervention ($t(14)=2.63$, $p=0.021$), while those in the active group did not ($t(19)=0.605$, $p=0.553$).

Independent-samples t-tests revealed marginal (non-significant) differences between the active and sham conditions pre- ($t(31)=1.74$, $p=0.092$) and post-intervention ($t(31)=1.96$, $p=0.059$), such that there was numerically lower baseline amygdala activation and numerically higher post-intervention amygdala activation in patients subsequently assigned to the active condition (compared to those assigned to sham). No other main or interaction effects reached significance (see Figure 3; full results described in Supplemental Materials).

L-DLPFC ROI analysis (*a priori*): Similarly, there was a significant interaction between time and stimulation condition ($F(1,31)=10.95$, $p=0.002$, $\eta_p^2=0.261$). In patients receiving active stimulation, left DLPFC activation (collapsing across fear and happy contrasts) marginally (non-significantly) increased from pre-to post-intervention ($t(18)=2.06$, $p=0.055$); in patients receiving sham stimulation, left DLPFC activation decreased significantly ($t(13)=2.62$, $p=0.021$).

Independent-samples t-tests revealed a marginal (non-significant) difference between the stimulation conditions at baseline ($t(31)=1.88$, $p=0.07$), but post-intervention there was significantly greater left DLPFC activation in the active compared to the sham stimulation group ($t(31)=2.08$, $p=0.046$). No other main or interaction effects reached significance (see Supplementary Materials).

sgACC ROI analysis (*a priori*): There were no significant effects.

Whole-brain analysis (exploratory): There were no significant main or interaction effects of stimulation condition on emotion processing at our corrected threshold of $p=0.025$ (cluster- or voxel-level) (see Supplemental Materials and Supplemental Table 3).

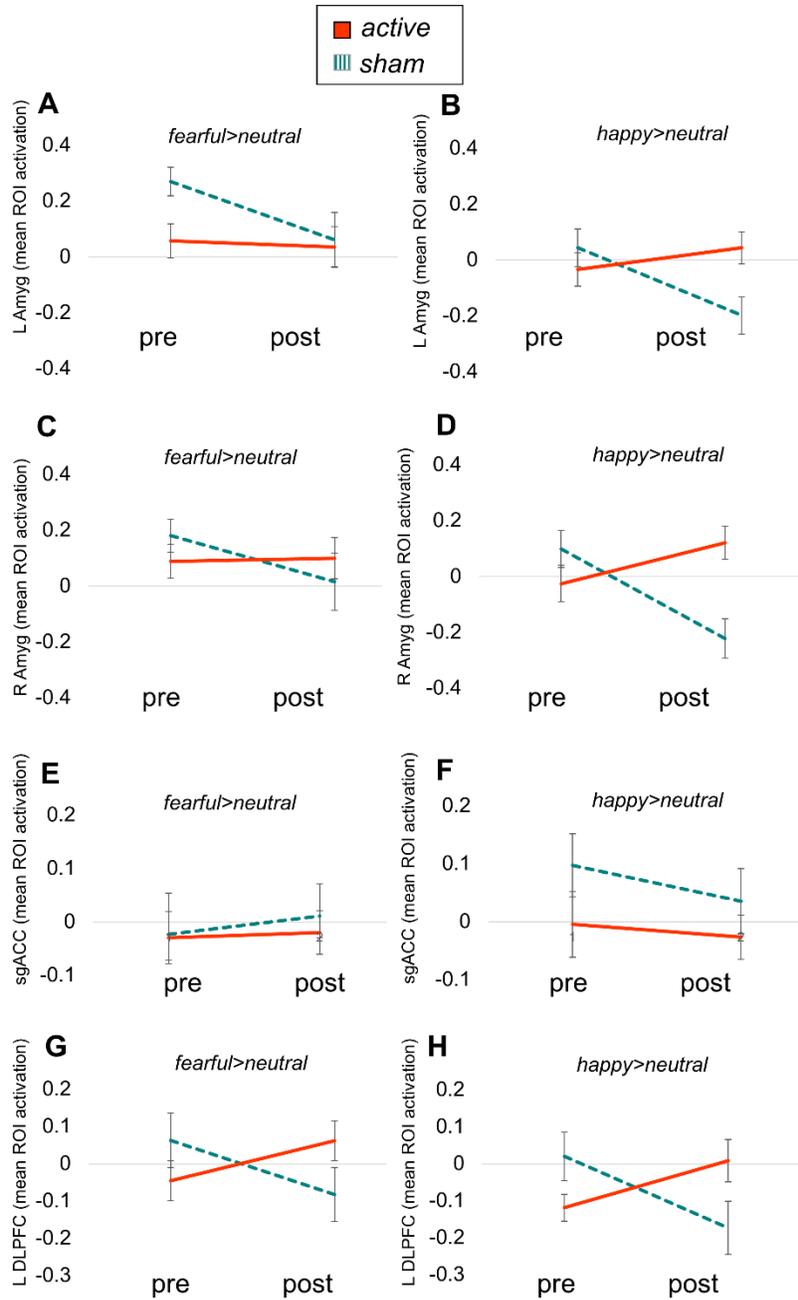


Fig 3. Effect of intervention on neural activation during emotion processing. Activation in *a priori* regions of interest pre- and post-intervention (orange solid lines = active tDCS; green dotted lines = sham tDCS) for the fearful vs neutral (left column) and happy vs neutral (right column) contrasts: left amygdala (L Amyg, A-B); right amygdala (R Amyg, C-D); subgenual anterior cingulate cortex (sgACC, E-F); and left dorsolateral prefrontal cortex (L DLPFC, G-H). In the amygdalae (A-D) there was a significant time-by-stimulation condition interaction, with activation decreased from pre- to post-intervention in the sham condition only, with a similar pattern of results for fearful and happy stimuli. No significant effects were detected in the sgACC (E-F). In the left DLPFC (G-H), there was a similar time-by-stimulation condition interaction to that in the amygdala.

Prediction of outcome

ROI analysis (*a priori*): In the L-DLPFC model (n-back: 3-back>1-back contrast), pre-treatment activation was significantly positively associated with subsequent percentage improvement in HAM-D score ($F(1,29)=12.77, p=0.001, \eta_p^2=0.306$). There was also a significant interaction between activation and stimulation condition ($F(1,29)=6.83, p=0.014, \eta_p^2=0.191$). This remained significant after controlling for baseline symptom and working memory scores (main effect: $F(1,26)=29.90, p<0.001, \eta_p^2=0.535$); interaction: ($F(1,26)=10.87, p=0.003, \eta_p^2=0.295$)) (see Supplemental Materials).

Pre-treatment activation in L-DLPFC was significantly positively associated with subsequent symptomatic improvement in the active ($r=0.711, p=0.001$) but not sham ($r=0.205, p=0.482$) condition (Figure 4A). Note, n-back performance did not predict percentage improvement ($r=0.112, p=0.536$). Baseline L-DLPFC activation was not associated with baseline HAM-D severity ($p>0.1$).

In the right amygdala model (emotion processing: fearful>neutral contrast), pre-treatment activation was significantly positively associated with subsequent percentage improvement in HAM-D score across both groups ($F(1,29)=7.69, p=0.010, \eta_p^2=0.210$); the interaction with stimulation group was non-significant ($F(1,29)=0.54, p=0.469$) (Figure 4B).

No other regions or their interaction with stimulation condition were significantly associated with clinical outcome (all $p>0.1$); therefore, we did not proceed with sensitivity or reliability analyses.

Whole-brain analysis (exploratory): Higher pre-treatment activation in the left posterior parietal cortex during working memory (PPC, $p=0.027$, FWE cluster-corrected) and higher left rostral PFC during emotion processing (fearful>neutral contrast; $p=0.001$, FWE cluster-corrected) were positively associated with greater subsequent percentage improvement in HAM-D across both groups (Figure 4C-F); **however, only the latter survived the Bonferroni corrected threshold of $p=0.025$ FWE**. No interactions with stimulation group survived whole-brain correction for multiple comparisons (see Supplemental Tables 4-6).

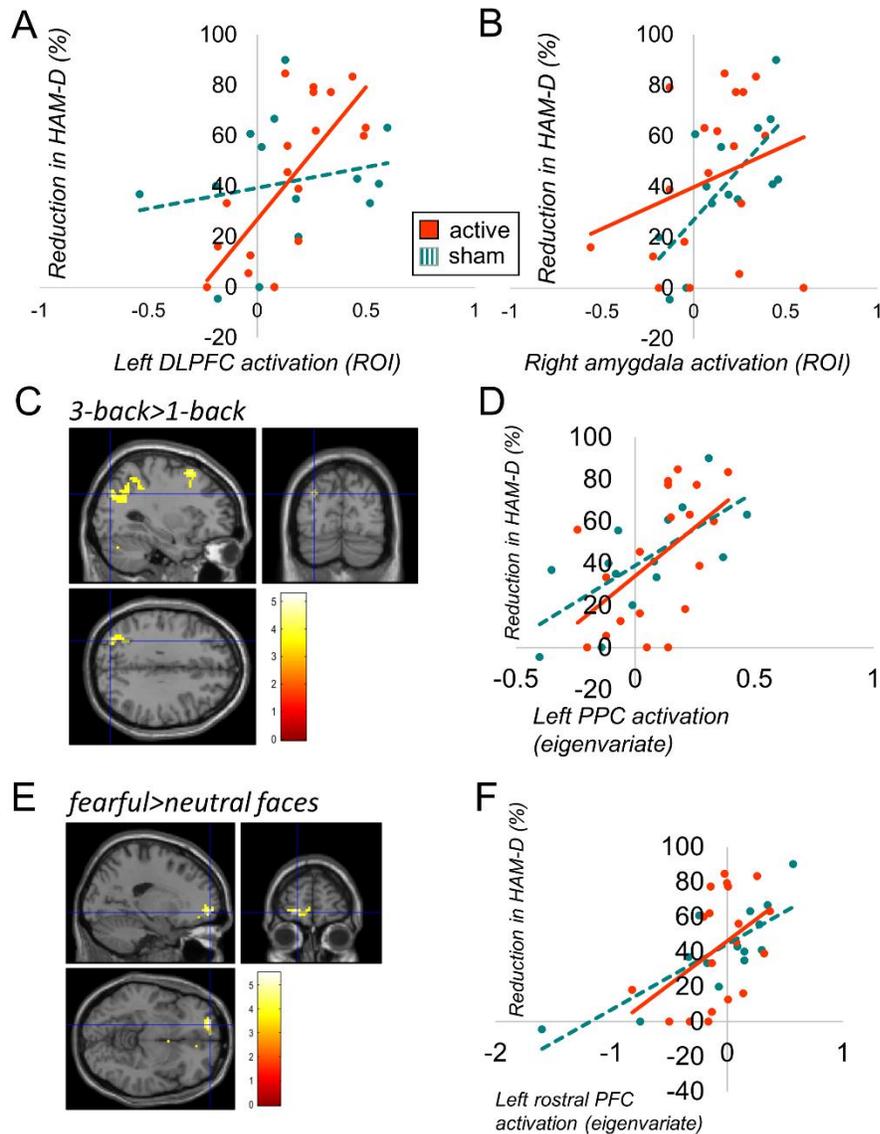


Fig 4. ROI analysis: imaging biomarkers associated with clinical response. A: Clinical response (percentage reduction in HAM-D) was associated with greater pre-treatment left DLPFC activation during working memory (3-back>1-back contrast) specifically in the active group ($r=0.711$, $p=0.001$) but not the sham group ($r=0.205$, $p=0.482$); interaction ($F(1,29)=6.83$, $p=0.014$). B: Clinical response was associated with greater pre-treatment right amygdala activation during emotion processing (fearful>neutral faces contrast) across both groups ($F(1,29)=7.69$, $p=0.010$). C: Results of whole-brain exploratory analysis for regions predictive of outcome in the n-back task (3-back>1-back contrast) ($p=0.027$, cluster-corrected; **note this does not fall below our stringent significance threshold of $p=0.025$**). Crosshairs located at peak voxel in left posterior parietal cortex (L PPC). D: For illustrative purposes only, scatterplot of relationship between L PPC activation (eigenvariate of cluster) and percent reduction in Hamilton Depression Rating Scale (HAM-D) over the trial in active (orange solid line) and sham (teal dotted line) groups. E: Results of whole-brain exploratory analysis for regions predictive of outcome in the emotion processing task (fearful>neutral contrast) ($p=0.001$, cluster-corrected). Crosshairs located at peak voxel in left rostral prefrontal cortex (PFC); activation extends into bilateral OFC. F: For illustrative purposes only, scatterplot of relationship between left rostral PFC activation (eigenvariate of cluster) and percent reduction in Hamilton Depression Rating Scale (HAM-D) over the trial in active and sham groups. For 5D and F, note that responses were extracted from clusters identified through a voxelwise analysis and therefore the magnitude of the correlation will be inflated (and thus is shown for illustrative purposes only).

Post-hoc area-under-curve (AUC) analysis: L-DLPFC prediction of clinical response

We repeated our primary analyses after dividing the sample into low and high L-DLPFC activation groups using a median split. Low L-DLPFC patients (N=19, 9 real; 10 sham) showed no difference between sham and real stimulation for either response ($X^2=0.540$, $p=0.463$) or remission ($X^2=0.006$, $p=0.937$). However, high L-DLPFC activation patients (N=20, 11 real; 9 sham) were significantly more likely to respond to tDCS than sham stimulation ($X^2=5.69$, $p=0.017$, $\phi=0.382$; remission: $X^2=2.78$, $p=0.095$, $\phi=0.267$). A receiver operator characteristic analysis of the accuracy of predicting clinical response using baseline L-DLPFC activation revealed that the area-under-curve (AUC) for real stimulation was 0.856 (good) while for the sham group AUC=0.417 (no better than chance).

Reliability

Activation in the L-DLPFC ROI (3-back>1-back contrast) showed very good reliability between pre- and post-trial scans (combined ICC=0.67, 95%CI=0.33—0.84, $p=0.001$). This was the case for both active (ICC=0.69, 95%CI=0.19—0.88) and sham (ICC=0.65, 95%CI=-0.08—0.89) groups. However, activation in the R-amygdala ROI (fearful>neutral contrast) showed poor reliability between scans (combined ICC=-1.630, 95%CI=-4.325— -0.299, $p=0.996$; active: ICC=-1.83, 95%CI=-6.35— -0.09; sham: ICC=-1.26, 95%CI=-6.04— -0.27) (Supplemental Figure 2).

Discussion

We conducted a mechanistic RCT of prefrontal tDCS combined with CBT for depression. We hypothesised that tDCS would enhance the therapeutic effect of CBT, compared with sham stimulation. Our results did not provide clear evidence for this; this may in part be due to the small sample size. On our mechanistic outcomes, we found that tDCS+CBT increased DLPFC activation relative to sham only during emotion processing, but not during working memory. Finally, we found a strong and specific association between pre-treatment L-DLPFC activation and subsequent clinical response, shedding light on the variability in our tDCS findings, and suggesting that it may be possible to identify which patients are most likely to respond to tDCS+CBT.

We did not find clear evidence that tDCS enhances the efficacy of CBT. While it is possible that our trial was underpowered (we identified moderately-sized odds ratios between active and sham groups (response: 2.16; remission: 3.65)), we found no effect of tDCS self-reported depressive symptoms. Thus, the clinical relevance of tDCS combined with CBT remains to be determined in future, larger studies. It is also possible that the intervention itself is not clinically efficacious on average because of inter-individual variability in response to tDCS. Indeed, a recent large-scale international trial of tDCS for depression reported a null result (the trial used a more typical daily or near-daily administration schedule, and a higher dosage)[17]. Therefore, one of the most important aspects of our trial was its mechanistic measures, which we could use to begin to explore such variability.

This trial was unusual in its inclusion of fMRI measurements. This design is central to understanding the possible mechanisms of tDCS+CBT, and can also illuminate sources of inter-individual variability. Here, DLPFC activation during working memory increased substantially over the trial (irrespective of condition), consistent with a smaller study[39], to a level comparable with healthy volunteers. Additionally, specifically under active tDCS, right PPC activation increased from pre- to post-intervention. This might reflect long-range effects of tDCS on the parietal cortex[9], which is robustly activated by the n-back task[40]. Similarly, L-DLPFC and bilateral amygdala activation during emotion processing generally decreased over time under sham stimulation, but increased under active stimulation. Thus, although DLPFC activation during working memory did not increase more during active than during sham stimulation, we did show treatment-dependent effects on the DLPFC during emotion processing.

Most importantly, we found that greater L-DLPFC activation prior to treatment was associated with greater subsequent symptomatic improvement, consistent with results from a small rTMS study[30]. Critically, this was only the case in patients receiving active stimulation, and this measure showed good reliability, a marker of potential utility as an fMRI “biomarker”[41]. Additionally, pre-treatment responses in the PPC (during working memory) and rostral PFC and right amygdala (during emotion processing) were associated with subsequent response to CBT, irrespective of stimulation condition (although note that the PPC result marginally missed statistical significance at our stringent FWE-corrected level of $p=0.025$). The amygdala result replicates the results of two previous neuroimaging investigations of CBT response[39,42], although puzzlingly the right amygdala also shows extremely poor test-retest reliability[41].

Our results suggest the DLPFC may have potential in predicting response to neurostimulation in depression, a result confirmed by significant effects on our primary outcome (HAM-D response) for the high-L-DLPFC group only, and by high accuracy of an ROC analysis using baseline L-DLPFC activation. In contrast, n-back working memory performance was not associated with improved clinical response. The high within-subject reliability of L-DLPFC activation fulfils a vital criterion for a clinically useful predictor, and suggests L-DLPFC BOLD response is more stable than other regions proposed as ‘biomarkers’, such as the sgACC and amygdala, which appear to have poor within-subject reliability[41], a finding we replicated here.

Individual differences in baseline DLPFC activation may contribute to the notable inconsistencies in the broader field of prefrontal tDCS[23]. Speculatively, while clinical response to our low dose of tDCS (1mA) is associated with high L-DLPFC activation at baseline, other dosages (for instance, the 2.5 mA used in a recent negative international clinical trial[17]) might show a different, or even reversed, relationship with DLPFC activation, making incorporation of mechanistic measures in future trials essential.

Limitations and future directions

Despite the strengths of our mechanistic design, there are important limitations, including relatively weak current strength and a fronto-extracerebral electrode montage. This montage results in comparatively diffuse frontal stimulation according to current flow models[33]. We chose this montage because in a [preliminary open-label study](#), it was substantially more effective for depression than a bifrontal tDCS montage[34]. We have previously shown an identical montage and current strength to have cognitive effects in healthy controls[43]. Recent work has confirmed the physiological validity of unilateral 1mA montages, failing to show physiological validity for bilateral stimulation or 2mA intensity conditions[35], despite their efficacy in previous depression trials[12,44]. It also bears mentioning that we found increased redness and itching observed in the real versus sham conditions; however, patients were worse than chance at guessing their stimulation condition (other side-effects were significantly more common in the sham condition).

Our weekly stimulation schedule could also have contributed to our null clinical result[11–14] (weekly CBT sessions is typical for NHS outpatients with moderate-to-severe depression). Furthermore, there is inherent heterogeneity between CBT therapists (confirmed here, despite efforts to standardise treatment); although including therapist as a covariate did not change our analyses.

Finally, we discovered a potential predictor of treatment outcome (baseline L-DLPFC activation), but it is as-yet unclear how this potential predictor relates to cognitive factors; importantly, baseline working memory was not predictive of treatment outcome. It is possible that DLPFC activation during the n-back task is reflective of a more general cold cognitive capacity than working memory ability *per se*; other cognitive measures of executive function (or perhaps a combined measure) may be better predictors of outcome than the n-back. One of the most important factors to establish, if tDCS is an effective intervention for depression, is the particular neural target it engages. While the most intuitive candidate is the region stimulated, our result failed to show any differential treatment effects on this region. Instead, we found a general increase in bilateral DLPFC activation regardless of tDCS condition. Future work should explore the contribution of baseline variables to this change in DLPFC activation. However, it is certainly possible that while the immediate neural effects of tDCS occur within the stimulated region (as previous work would suggest[45]), the longer-term mechanisms driving its clinical effect may be more distally located (for example, in the parietal cortex, as suggested by our exploratory whole-brain analysis).

Conclusion

Our data do not provide clear support for the use of tDCS to augment CBT in depression in all patients, but we discovered a biomarker explaining variability in tDCS response: baseline activation in the area of stimulation (L-DLPFC). This biomarker appears relatively strong and specific to active stimulation. Our data speak to the potential utility of neuroscience techniques in developing novel mental health interventions; incorporating cognitive neuroscience measures into clinical trials could allow simultaneous assessment of clinical efficacy and detection of potential biomarkers of treatment response. Ultimately, these data could guide patient selection in larger clinical trials, and, if confirmed, inform clinical use of tDCS in depression.

Funding and Disclosures: This work was supported by the Brain and Behavior Research Foundation (grant number 20162) to JPR and a Brain Research Trust PhD studentship awarded to CLN, and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre (SP). JPR consults for Cambridge Cognition, Takeda Ltd and GE. The other authors report no conflict of interest.

Acknowledgments: The authors would like to thank research assistants Jessica Aylward and Alan Gray for their help with data collection. They are also extremely grateful to the NHS clinicians Michaela Thomas, Dr China Eziefula, Carlo Gittens, Laura Franklin, Dr Karen Katz, Dr Robin Weedon, Dr Jeremy Handel, Dr Flavia Cigolla, Akbar Jamil, Dr Isabella Foustanos, Dr Abi Harris, Dr Sukhvinder Virdee, Dr Jessica Wilner-Reid, and Dr Joshua Buckman; as well as all the patients who generously volunteered their time for the trial. **Author contributions:** CLN, JPR, NL, SP, VV, and CC designed the study, CLN, CH, TL, and JL collected data, CLN analysed data, CLN and JPR wrote the paper with contributions from all the other authors.

Data and materials availability: All data discussed in the paper will be made available to any reader who enters into a managed access arrangement. Ethical and data governance restrictions prevent the data from being shared in a public repository (all data is patient data collected under UK National Health Service ethics and patients did not give consent for open sharing). All data will be shared freely following a formal Data Sharing Agreement with University College London. Researchers who wish to access the data should contact the senior author (j.roiser@ucl.ac.uk).

References

1. Kendall PC, Hollon SD. Cognitive-behavioral interventions: Theory, research, and procedures. vol. 21. Academic Press; 2013.
2. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry*. 2005;62:409–416.
3. Frye MA, Grunze H, Suppes T, McElroy SL, Keck Jr PE, Walden J, et al. A placebo-controlled evaluation of adjunctive modafinil in the treatment of bipolar depression. *American Journal of Psychiatry*. 2007;164:1242–1249.
4. Vedeniapin A, Cheng L, George MS. Feasibility of simultaneous cognitive behavioral therapy (CBT) and left prefrontal rTMS for treatment resistant depression. *Brain Stimulation*. 2010;3:207.
5. D’Urso G, Mantovani A, Micillo M, Priori A, Muscettola G. Transcranial direct current stimulation and cognitive-behavioral therapy: evidence of a synergistic effect in treatment-resistant depression. *Brain Stimul*. 2013;6:465–467.
6. Wang X-L, Du M-Y, Chen T-L, Chen Z-Q, Huang X-Q, Luo Y, et al. Neural correlates during working memory processing in major depressive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. 2015;56:101–108.
7. Hooley JM, Gruber SA, Scott LA, Hiller JB, Yurgelun-Todd DA. Activation in dorsolateral prefrontal cortex in response to maternal criticism and praise in recovered depressed and healthy control participants. *Biological Psychiatry*. 2005;57:809–812.
8. Kerestes R, Ladouceur C, Meda S, Nathan P, Blumberg H, Maloney K, et al. Abnormal prefrontal activity subserving attentional control of emotion in remitted depressed patients during a working memory task with emotional distracters. *Psychological Medicine*. 2012;42:29–40.
9. Stagg CJ, Lin RL, Mezue M, Segerdahl A, Kong Y, Xie J, et al. Widespread modulation of cerebral perfusion induced during and after transcranial direct current stimulation applied to the left dorsolateral prefrontal cortex. *The Journal of Neuroscience*. 2013;33:11425–11431.

10. Nord CL, Lally N, Charpentier CJ. Harnessing electric potential: DLPFC tDCS induces widespread brain perfusion changes. *Frontiers in Systems Neuroscience*. 2013;7.
11. Boggio PS, Rigonatti SP, Ribeiro RB, Myczkowski ML, Nitsche MA, Pascual-Leone A, et al. A randomized, double-blind clinical trial on the efficacy of cortical direct current stimulation for the treatment of major depression. *The International Journal of Neuropsychopharmacology*. 2008;11:249–254.
12. Brunoni A, Valiengo L, Baccaro A, Zanao TA, de Oliveira JF, Goulart A, et al. The sertraline vs electrical current therapy for treating depression clinical study: results from a factorial, randomized, controlled trial. *JAMA Psychiatry*. 2013;70:383–391.
13. Fregni F, Boggio PS, Nitsche MA, Marcolin MA, Rigonatti SP, Pascual-Leone A. Treatment of major depression with transcranial direct current stimulation. *Bipolar Disorders*. 2006;8:203–204.
14. Loo CK, Alonzo A, Martin D, Mitchell PB, Galvez V, Sachdev P. Transcranial direct current stimulation for depression: 3-week, randomised, sham-controlled trial. *The British Journal of Psychiatry*. 2012;200:52–59.
15. Nord CL, Roiser J. Non-invasive direct current brain stimulation: the evidence behind the hype. *Advances in Clinical Neuroscience and Rehabilitation*. 2015;15:9–11.
16. Shiozawa P, Fregni F, Benseñor IM, Lotufo PA, Berlim MT, Daskalakis JZ, et al. Transcranial direct current stimulation for major depression: an updated systematic review and meta-analysis. *The International Journal of Neuropsychopharmacology*. 2014:1–10.
17. Loo CK, Husain MM, McDonald WM, Aaronson S, O'Reardon JP, Alonzo A, et al. International randomized-controlled trial of transcranial Direct Current Stimulation in depression. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*. 2018;11:125–133.
18. Dockery CA, Hueckel-Weng R, Birbaumer N, Plewnia C. Enhancement of planning ability by transcranial direct current stimulation. *The Journal of Neuroscience*. 2009;29:7271–7277.
19. Metuki N, Sela T, Lavidor M. Enhancing cognitive control components of insight problems solving by anodal tDCS of the left dorsolateral prefrontal cortex. *Brain Stimulation*. 2012;5:110–115.
20. Fregni F, Boggio PS, Nitsche M, Berman F, Antal A, Ferdoes E, et al. Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Experimental Brain Research*. 2005;166:23–30.
21. Lally N, Nord CL, Walsh V, Roiser JP. Does excitatory fronto-extracerebral tDCS lead to improved working memory performance? *F1000Research*. 2013;2.
22. Oliveira JF, Zanao TA, Valiengo L, Lotufo PA, Benseñor IM, Fregni F, et al. Acute working memory improvement after tDCS in antidepressant-free patients with major depressive disorder. *Neuroscience Letters*. 2013;537:60–64.
23. Horvath JC, Forte JD, Carter O. Quantitative Review Finds No Evidence of Cognitive Effects in Healthy Populations from Single-Session Transcranial Direct Current Stimulation (tDCS). *Brain Stimulation*. 2015. 2015.
24. Tremblay S, Lepage J-F, Latulipe-Loiselle A, Fregni F, Pascual-Leone A, Théoret H. The uncertain outcome of prefrontal tDCS. *Brain Stimulation*. 2014;7:773–783.
25. Price AR, Hamilton RH. A Re-evaluation of the Cognitive Effects From Single-session Transcranial Direct Current Stimulation. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation*. 2015. 2015.
26. Segrave R, Arnold S, Hoy K, Fitzgerald P. Concurrent cognitive control training augments the antidepressant efficacy of tDCS: a pilot study. *Brain Stimulation*. 2014;7:325–331.
27. Welch ES, Weigand A, Hooker JE, Philip NS, Tyrka AR, Press DZ, et al. Feasibility of Computerized Cognitive-Behavioral Therapy Combined With Bifrontal Transcranial Direct Current Stimulation for Treatment of Major Depression. *Neuromodulation: Technology at the Neural Interface*. 2018. 2018.
28. Dunlop BW, Mayberg HS. Neuroimaging-based biomarkers for treatment selection in major depressive disorder. *Dialogues in Clinical Neuroscience*. 2014;16:479.
29. Roiser JP, Elliott R, Sahakian BJ. Cognitive mechanisms of treatment in depression. *Neuropsychopharmacology*. 2012;37:117–136.
30. Weiduschat N, Dubin MJ. Prefrontal cortical blood flow predicts response of depression to rTMS. *Journal of Affective Disorders*. 2013;150:699–702.
31. Nitsche MA, Paulus W. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology*. 2001;57:1899–1901.
32. O'Nions EJ, Dolan RJ, Roiser JP. Serotonin transporter genotype modulates subgenual response to fearful faces using an incidental task. *Journal of Cognitive Neuroscience*. 2011;23:3681–3693.
33. Laakso I, Tanaka S, Mikkonen M, Koyama S, Sadato N, Hirata A. Electric fields of motor and frontal tDCS in a standard brain space: a computer simulation study. *Neuroimage*. 2016;137:140–151.

34. Martin DM, Alonzo A, Mitchell PB, Sachdev P, Gálvez V, Loo CK. Fronto-extracerebral transcranial direct current stimulation as a treatment for major depression: an open-label pilot study. *Journal of Affective Disorders*. 2011;134:459–463.
35. Parkin BL, Bhandari M, Glen JC, Walsh V. The physiological effects of Transcranial Electrical Stimulation do not apply to parameters commonly used in studies of Cognitive Neuromodulation. *Neuropsychologia*. 2018. 2018.
36. Ambrus GG, Al-Moyed H, Chaieb L, Sarp L, Antal A, Paulus W. The fade-in–short stimulation–fade out approach to sham tDCS–reliable at 1 mA for naive and experienced subjects, but not investigators. *Brain Stimulation*. 2012;5:499–504.
37. Hollon SD, Stewart MO, Strunk D. Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annu Rev Psychol*. 2006;57:285–315.
38. Brunoni AR, Amadera J, Berbel B, Volz MS, Rizzerio BG, Fregni F. A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. *International Journal of Neuropsychopharmacology*. 2011;14:1133–1145.
39. Siegle GJ, Thompson W, Carter CS, Steinhauer SR, Thase ME. Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biological Psychiatry*. 2007;61:198–209.
40. Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*. 2005;25:46–59.
41. Nord CL, Gray A, Charpentier C, Robinson O, Roiser J. Unreliability of putative fMRI biomarkers during emotional face processing. *NeuroImage*. 2017. 2017.
42. Siegle GJ, Carter CS, Thase ME. Use of fMRI to predict recovery from unipolar depression with cognitive behavior therapy. *American Journal of Psychiatry*. 2006. 2006.
43. Nord CL, Forster S, Halahakoon DC, Penton-Voak IS, Munafò MR, Roiser JP. Prefrontal cortex stimulation does not affect emotional bias, but may slow emotion identification. *Social Cognitive and Affective Neuroscience*. 2017. 2017.
44. Brunoni AR, Moffa AH, Sampaoi-Junior B, Borriero L, Moreno ML, Fernandes RA, et al. Trial of Electrical Direct-Current Therapy versus Escitalopram for Depression. *The New England Journal of Medicine*. 2017;376:2523–2522.
45. Stagg CJ, Lin RL, Mezue M, Segerdahl A, Kong Y, Xie J, et al. Widespread Modulation of Cerebral Perfusion Induced during and after Transcranial Direct Current Stimulation Applied to the Left Dorsolateral Prefrontal Cortex. *The Journal of Neuroscience*. 2013;33:11425–11431.

Figure legends

Fig 1. CONSORT diagram describing recruitment, randomization, and attrition in the clinical trial (A) and depiction of trial procedure (B). CBT=cognitive behavioural therapy; tDCS=transcranial direct current stimulation; HAM-D=Hamilton Rating Scale for Depression; fMRI=functional magnetic resonance imaging.

Fig 2. Effect of intervention on depressive symptoms and brain activation. A: Proportion of patients responding and remitting following CBT with active and sham tDCS. These differences did not achieve statistical significance. B: HAM-D scores reduced significantly from pre- to post-intervention, but the main effect of stimulation condition and the interaction were non-significant. C&D: **Effect of intervention on neural activation during working memory** (contrast: 3-back > 1-back). Activation in left (C) and right (D) dorsolateral prefrontal cortex (DLPFC) *a priori* regions-of-interest pre- and post-intervention (orange solid lines = active tDCS; green dotted lines = sham tDCS). E: **Whole brain (WB) analysis of the effect of time** (pre-versus post-intervention) and plotted eigenvariate of this main effect. E, left panel: we found significant changes in activation over time in bilateral DLPFC (increased over time: coronal section; $p < 0.001$ (uncorrected) cluster-forming threshold, WB cluster-level FWE-corrected) and medial PFC/perigenual anterior cingulate cortex (decreased over time: sagittal and axial sections; $p < 0.001$ (uncorrected) cluster-forming threshold, WB cluster-level FWE-corrected). E, right panel: for illustrative purposes, eigenvariate values extracted from the L-DLPFC cluster displayed in the left panel are plotted. F: WB effect of group (greater in active than sham tDCS condition) in the right posterior parietal cortex (rPPC) ($p = 0.02$ WB voxel-level FWE corrected: F, left), and plotted eigenvariate of this rPPC cluster (F, right). Overlays are displayed at $p < 0.001$ (uncorrected) and colour bars indicate F-values. Error bars indicate standard errors of the mean.

Fig 3. Effect of intervention on neural activation during emotion processing. Activation in *a priori* regions of interest pre- and post-intervention (orange solid lines = active tDCS; green dotted lines = sham tDCS) for the fearful vs neutral (left column) and happy vs neutral (right column) contrasts: left amygdala (L Amyg, A-B); right amygdala (R Amyg, C-D); subgenual anterior cingulate cortex (sgACC, E-F); and left dorsolateral prefrontal cortex (L DLPFC, G-H). In the amygdalae (A-D) there was a significant time-by-stimulation condition interaction, with activation decreased from pre- to post-intervention in the sham condition only, with a similar pattern of results for fearful and happy stimuli. No significant effects were detected in the sgACC (E-F). In the left DLPFC (G-H), there was a similar time-by-stimulation condition interaction to that in the amygdala.

Fig 4. ROI analysis: imaging biomarkers associated with clinical response. A: Clinical response (percentage reduction in HAM-D) was associated with greater pre-treatment left DLPFC activation during working memory (3-back>1-back contrast) specifically in the active group ($r=0.711$, $p=0.001$) but not the sham group ($r=0.205$, $p=0.482$); interaction ($F(1,29)=6.83$, $p=0.014$). B: Clinical response was associated with greater pre-treatment right amygdala activation during emotion processing (fearful>neutral faces contrast) across both groups ($F(1,29)=7.69$, $p=0.010$). C: Results of whole-brain exploratory analysis for regions predictive of outcome in the n-back task (3-back>1-back contrast) ($p=0.027$, cluster-corrected; **note this does not fall below our stringent significance threshold of $p=0.025$**). Crosshairs located at peak voxel in left posterior parietal cortex (L PPC). D: For illustrative purposes only, scatterplot of relationship between L PPC activation (eigenvariate of cluster) and percent reduction in Hamilton Depression Rating Scale (HAM-D) over the trial in active (orange solid line) and sham (teal dotted line) groups. E: Results of whole-brain exploratory analysis for regions predictive of outcome in the emotion processing task (fearful>neutral contrast) ($p=0.001$, cluster-corrected). Crosshairs located at peak voxel in left rostral prefrontal cortex (PFC); activation extends into bilateral OFC. F: For illustrative purposes only, scatterplot of relationship between left rostral PFC activation (eigenvariate of cluster) and percent reduction in Hamilton Depression Rating Scale (HAM-D) over the trial in active and sham groups. For 5D and F, note that responses were extracted from clusters identified through a voxelwise analysis and therefore the magnitude of the correlation will be inflated (and thus is shown for illustrative purposes only).

Supplemental materials

Side effects (active stimulation)	% sessions	Number of instances	% mild	% moderate	% severe	% "unrelated" to stimulation	% "remotely" related	% "possibly" related	% "probably" related	% "definitely" related
Headache	12.1	18	61.1	16.7	22.2	33.3	11.1	22.2	0.0	0.0
Neck pain	6.7	10	90.0	0.0	10.0	50.0	10.0	10.0	20.0	0.0
Scalp pain	18.8	28	60.7	35.7	3.6	17.9	0.0	28.6	39.3	0.0
Tingling	79.9	119	63.0	27.7	9.2	9.2	16.0	17.6	53.8	1.7
Itching	54.4	81	50.6	32.1	17.3	8.6	17.3	17.3	53.1	1.2
Burning	25.5	38	65.8	23.7	10.5	10.5	7.9	18.4	65.8	2.6
Skin redness	28.9	43	69.8	23.3	7.0	14.0	18.6	2.3	62.8	0.0
Sleepiness	34.2	51	41.2	33.3	25.5	11.8	35.3	31.4	9.8	3.9
Trouble concentrating	46.6	68	38.2	38.2	23.5	13.2	38.2	16.2	8.8	2.9
Mood changes	16.1	24	41.7	29.2	29.2	33.3	33.3	12.5	16.7	8.3
Side effects (sham stimulation)	% sessions	No. instances	% Mild	% moderate	% severe	% "unrelated" to stimulation	% "remotely" related	% "possibly" related	% "probably" related	% "definitely" related
Headache	28.3	34	73.5	14.7	11.8	32.4	32.4	14.7	0.0	0.0
Neck pain	12.5	15	66.7	33.3	0.0	33.3	20.0	13.3	0.0	0.0
Scalp pain	19.2	23	56.5	43.5	0.0	8.7	13.0	39.1	39.1	0.0
Tingling	73.3	88	54.5	43.2	2.3	1.1	15.9	45.5	31.8	5.7
Itching	30.0	36	58.3	27.8	13.9	5.6	27.8	16.7	41.7	5.6
Burning	38.3	46	58.7	32.6	8.7	4.3	8.7	47.8	30.4	4.3
Skin redness	1.7	2	50.0	50.0	0.0	100.0	0.0	0.0	0.0	0.0
Sleepiness	60.0	72	41.7	31.9	26.4	27.8	23.6	26.4	8.3	1.4
Trouble concentrating	74.2	89	48.3	40.4	11.2	31.5	27.0	14.6	4.5	0.0
Mood changes	13.3	16	81.3	12.5	6.3	25.0	56.3	6.3	0.0	0.0

Supplemental File 1. Table of side effects in active and sham groups

Participants

Potential patients were initially identified by a member of the IAPT clinical team if they met the following criteria: willingness to take part in research; a patient health questionnaire (PHQ-9) score of 15 or above (indicating at least moderate depression); not currently taking antidepressant medication; and internal NHS criteria (including severity and risk) for one-on-one cognitive behavioural therapy for depression. At the clinic, an assessment was also made about the primary complaint: in many cases, a patient was depressed and met the eligibility criteria for the study, but the assessing therapist believed the focus of the therapy was better placed on another presenting issue, for example, bulimia nervosa or body dysmorphic disorder. Additionally, some patients were given more than one psychotherapeutic option by the assessing clinician, leading to non-eligibility for the trial in cases where the patient elected to receive a different type of therapy (e.g., couples' therapy or dialectical behavioural therapy).

Following initial IAPT assessment, all eligible patients (N=71) were screened in person at the Institute of Cognitive Neuroscience, University College London by the lead researcher (C.N.), where diagnosis was confirmed using the Mini International Neuropsychiatric Inventory [34] and depression severity measured with the HAM-D [35]. Patients were also assessed for the following exclusion criteria: use of any psychotropic medication; past or present substance or alcohol dependence (save for a remote history of abuse or dependence restricted to a prior major depressive episode); illegal drug use within the past month; neurological illness; major health conditions likely to affect cognitive performance; and prior or present manic or psychotic symptoms (with the exception of psychotic depression). Additionally, we excluded patients who did not meet MRI safety criteria, which included presence of any irremovable ferromagnetic metal in or on the body, and medical conditions that might increase the risk of the MRI scan: pregnancy; severe claustrophobia; back pain; or severe asthma.

At this session, patients practiced the working memory task (the n-back task) and completed self-report questionnaires measuring depression (Beck Depression Inventory: BDI [36]), anxiety (Beck Anxiety Inventory: BAI [37]), and anhedonia (Snaith-Hamilton Pleasure Scale: SHAPS [38]). Those who were native English speakers or who had been educated in English during high-school (N=23) were administered the Weschler Test of Adult Reading (WTAR), which we used to estimate intelligence quotient.

Following the screening session, eligible patients were recruited to the study, at which point they were placed on the IAPT waiting list for one-on-one CBT. Between one week and six months after the screening session (mean=58.6 days; SD=40.2 days), patients received their first magnetic resonance imaging (MRI) scan (at the Birkbeck-UCL Centre for Neuroimaging) and began their course of CBT. All MRI scans were collected between 1 and 6 days before starting therapy. In the case of more than one month elapsing between initial assessment and the beginning of CBT, our primary depression measure (the HAM-D) was repeated. We also scanned 30 healthy participants (full case-control analyses at baseline will be reported separately).

After the initial scan, patients were randomized to active or sham conditions using custom-written MATLAB code by a researcher not involved in the trial. Researchers involved in the trial were given a list of five-digit codes to input to the stimulator, half of which corresponded to active stimulation and half to sham.

For patients who completed the trial (at least 7 CBT sessions: N=33), a second MRI scan was scheduled shortly (0-6 days) after their final CBT session. This was identical to the pre-intervention scan, and followed by a final HAM-D interview. All HAM-D interviews were conducted by the same researcher (C.N.).

Imaging details

At each scanning session, following a standard localizer scan, we collected one five-minute T1-weighted anatomical scan (1 mm isotropic magnetization-prepared rapid gradient-echo) using a Siemens Avanto 1.5 Tesla MRI scanner with a 32-channel head coil. Patients then completed the two tasks in the scanner, during which we collected T2*-weighted gradient-echo planar images (EPIs) with 36 slices per volume. Each EPI scan was followed by a corresponding field map scan (two T2*-weighted images with the identical volume and parameters of each EPI scan), to enable offline correction of blood oxygen level-dependent (BOLD) images for distortion. For both tasks, echo time was 50 ms, repetition time per slice was 87 msec, and in-plane resolution was 2 x 2 mm. We varied slice thickness between the tasks: for the emotion processing task, 2 mm slice thickness to allow optimal coverage of the ventral brain; for the n-back task, 2.5 mm slice thickness to allow fuller brain coverage including dorsal prefrontal regions.

We analysed all data using Statistical Parametric Mapping (SPM12; Wellcome Trust Centre for Neuroimaging, London, www.fil.ion.ucl.ac.uk/spm) in MATLAB R2015a. For each time series, we removed the first six volumes to allow for T1 equilibration, realigned the remaining volumes to the seventh volume, coregistered the volumes to each participant's anatomical scan, normalized into standardized space using the Montreal Neurological Institute template, and smoothed using an 8mm full width at half maximum Gaussian kernel. All image sequences were checked for large movements (>1.5mm) or rotations (>1 degree) in any direction following realignment (no images were excluded), and anatomical images were manually checked for artefacts related to overfitting following normalization.

Regressors of interest were convolved with a synthetic hemodynamic response function time-locked to the onset of the corresponding event (for the emotion processing task, each 16-second emotion block; for the n-back task, each 18-second 3-back or 1-back block). We also included six movement regressors of no interest in all subjects. For subjects who made gender discrimination errors in the emotion processing task, an extra regressor was included to model trials where they may not have been attending to the task (22 patients; mean=4.6 out of 96 trials). For the eleven subjects who made no errors, this regressor was not included. For each regressor, we estimated parameter estimate images using the general linear model; these were combined to create contrasts for each task.

We conducted second-level analyses according to the standard summary statistics approach to random effects analysis. [For each subject, and for each contrast, we then extracted mean activation across each ROI using the `spm_sample_vol` function incorporated into a custom-](#)

written Matlab script that returned voxel values from each volume (in this case, each anatomical ROI).

Linear Mixed Models Analysis

Weekly scores were calculated for the Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), and Snaith-Hamilton Pleasure Scale (SHAPS), and weekly performance on the n-back task was measured as d-prime on the n-back task, which incorporates both hits and false alarms in the 3-back condition. For these variables, the model included effects of time (i.e., for the majority of patients, 8 sessions), stimulation group (active/sham), and the interaction between the two as fixed effects, and participant as a random effect. We employed a heterogeneous first order autoregressive covariance structure, which does not assume homogeneous variance between conditions, but which is useful for data where each measurement is most closely correlated with its proximal measurements, with correlations decreasing with distance.

Outcome prediction analyses

Here, we first tested for multicollinearity for each predictor separately, assuming evidence of multicollinearity if the variance inflation factor (VIF) between any two variables was greater than 3. Next, we constructed the model including all six predictor variables (if there was no evidence for multicollinearity) and their interactions with stimulation group. If the overall model was non-significant, we did not proceed with stepwise regression. If the overall model was significant, we used a backwards stepwise regression technique to test which of the measures predicted subsequent clinical improvement. We removed any variables with $p > 0.1$, until all variables, or their interactions with stimulation group, predicted percent change in HAM-D sufficiently (i.e., $p < 0.1$). We then report these remaining predictor variables.

Additionally, for significant single-predictor models, we assessed the within-subject reliability of activation between the baseline scan and the post-intervention scan by calculating intra-class correlation coefficients (ICCs) for each ROI that was significantly associated with clinical response (see Supplemental Materials for details).

For each contrast (n-back: 3-back > 1-back; emotion processing: fearful > neutral, happy > neutral), we also conducted an exploratory whole-brain analysis examining prediction of clinical response. We constructed an independent samples t-test in SPM testing for the effect of group (active or sham) on each contrast and included percent change in HAM-D as a covariate in the model and the interaction. For completeness we report all whole-brain activations associated with response across both groups, and all interactions between group and percent change HAM-D ($p < 0.001$ (uncorrected), minimum cluster size $k=4$). Whole-brain corrected p -values are reported at the cluster and voxel levels.

Reliability analyses

For significant single-predictor models, we assessed the within-subject reliability of activation between the baseline scan and the post-intervention scan by calculating intra-class correlation coefficients (ICCs) for each ROI that was significantly associated with clinical response. Note that although we expected group-level changes between baseline and post-intervention activation, our measure of within-subject reliability assesses only the relative consistency of activation between scans (i.e., the subjects with greater activation on day 1,

relative to the rest of the group, are also those with greater activation on day 2). Particularly in the case of an interventional design such as this, we would not expect absolute agreement between individual subjects' activation. We calculated intra-class correlation coefficients (ICCs) for each ROI that was significantly associated with clinical response. The ICC is a standard method to quantify the stability of measurements between test and retest sessions (Bennett and Miller, 2010). To calculate ICCs, we used the identical approach as in our previous work (Nord *et al.*, 2017b), assessing reliability using ICC(3,1), a 2-way mixed effects ICC. We adhere to a conventional interpretation of ICCs to quantify the degree of reliability: ICC<0.4 = poor reliability; 0.4—0.75 = moderate to good reliability; >0.75 = excellent reliability (Fleiss, 1986; Plichta *et al.*, 2012; Nord *et al.*, 2017b). A negative ICC is usually interpreted as a reliability of zero (Bartko, 1976), since the theoretical limits of the ICC are 0—1 (negative values can occur when the within-groups variance exceeds the between group variance, but are outside the theoretical range of the ICC (Lahey *et al.*, 1983)). We also report *p*-values and 95% confidence intervals for all ICCs, obtained from an *F*-test against the null hypothesis.

Assessment of blinding and side effects

We compared the most frequent guess for each participant with their actual condition to verify blinding. Due to a bias for patients to believe they were in the active condition, 21.1% of the sham group, but 60.0% of the active group, guessed correctly.

All patients reported at least one side effect on at least one stimulation session. The most common symptom reported was tingling, reported on 80% of active and 73% of sham stimulation sessions (no significant difference between stimulation conditions, $X^2=1.60$, $p=0.206$). Headache was reported significantly more often in the sham condition ($X^2=11.26$, $p=0.001$), as was burning sensation ($X^2=5.10$, $p=0.024$), sleepiness ($X^2=17.79$, $p<0.001$), and trouble concentrating ($X^2=22.26$, $p<0.001$; NB, side-effects were assessed immediately following the n-back task). Itching and skin redness were reported significantly more frequently in the active condition ($X^2=16.05$, $p<0.001$; $X^2=35.28$, $p<0.001$). Correcting for multiple comparisons (N=10 side-effects), there was no longer a significant difference in burning sensation between the groups (Bonferroni-adjusted threshold: $p=0.005$), but the other differences remained significant. Note that the Type I error rate is likely to be elevated for these analyses as observations over successive sessions within an individual are probably not independent from one another. See Supplemental Table 1 for full results.

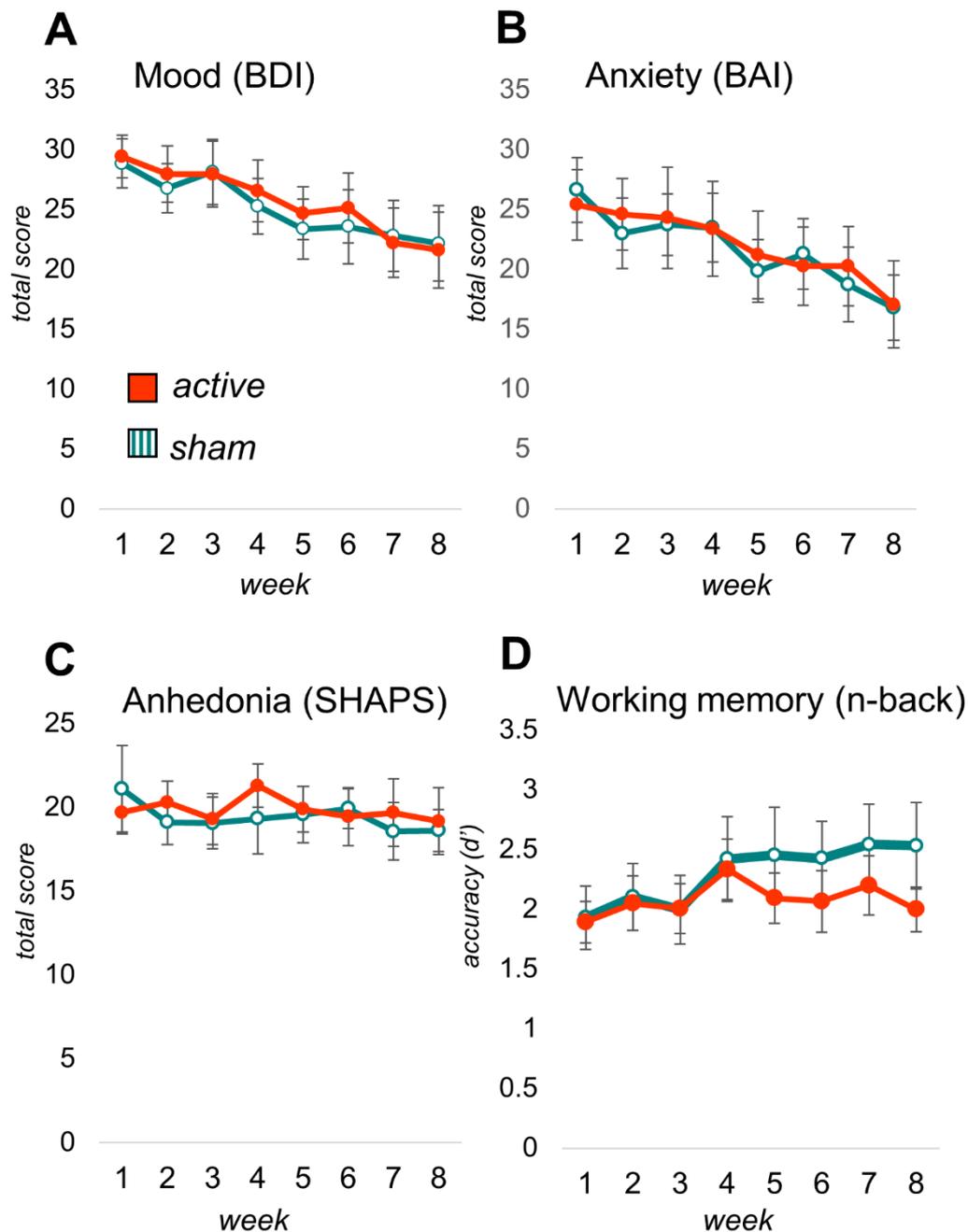
Secondary clinical outcomes

For the questionnaire measures, using linear mixed models we found a significant decrease over time in the questionnaire-based Beck Depression Inventory (BDI: $F(7,108.80)=2.99$, $p=0.007$) and Beck Anxiety Inventory (BAI: $F(7,139.56)=2.16$, $p=0.042$) scores, but not anhedonia (Snaith Hamilton Pleasure Scale (SHAPS) scores ($F(7,131,56)=0.685$, $p=0.684$)). None showed any significant effect of stimulation condition or interaction between time and stimulation condition ($p>0.1$). There was no significant effect of time on working memory performance ($F(7,149.96)=1.99$, $p=0.060$), nor any effect of stimulation condition ($F(1,40.57)=0.340$, $p=0.563$) or interaction ($F(7,149.96)=0.421$, $p=0.888$) (Supplemental Figure 1).

In the per-protocol sample (N=33 completers), HAM-D scores reduced significantly from pre- to post-intervention ($F(1,31)=64.07, p<0.001, \eta_p^2=0.674$), but there was no interaction with stimulation condition ($F(1,31)=0.02, p=0.881$). Again, there was no main effect of stimulation condition ($F(1,31)=0.048, p=0.828$). Here, the active group decreased by a mean difference of 9.42 (SD = 6.99) points; the sham group decreased by a mean difference of 9.07 (SD = 5.92) points.

Including therapist as a covariate in this analysis produced similar results: a reduction in HAM-D following the intervention ($F(1,36)=40.45, p<0.001, \eta_p^2=0.529$), and no main effect of stimulation condition ($F(1,36)=0.109, p=0.743$) or interaction between stimulation condition and the intervention overall ($F(1,36)=1.29, p=0.263$); however, there was a difference in outcomes in general (across both stimulation conditions) for different therapists ($F(1,36)=7.38, p=0.01, \eta_p^2=0.170$).

No questionnaire showed any significant effect of stimulation condition or interaction between time and stimulation condition: BDI: stimulation condition ($F(1, 35.82)=0.081, p=0.777$); interaction ($F(7, 108.80)=0.274, p=0.963$); BAI: stimulation condition ($F(1, 40.10)=0.332, p=0.568$) or interaction ($F(7, 139.56)=0.871, p=0.531$); SHAPS: stimulation condition ($F(1, 39.52)=0.343, p=0.561$); interaction condition ($F(7, 131.56)=0.561, p=0.794$). For working memory, there was no significant effect of time ($F(7,149.96)=1.99, p=0.060$), nor any effect of stimulation condition ($F(1,40.57)=0.340, p=0.563$) or interaction ($F(7,149.96)=0.421, p=0.888$) (see Supplemental Figure 1)



Supplemental Figure 1. Weekly mood (A), anxiety (B), and anhedonia (C) self-report scores, and performance on the working memory (n-back) task (D), for each stimulation condition. Error bars represent standard error of the mean. BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; SHAPS = Snaith-Hamilton Pleasure Scale.

In-scanner working memory and emotion task performance

The n-back behavioural data met assumptions of normality on both sessions (Kolmogorov-Smirnov test, both $p=0.2$), but behavioural data on the emotion processing task (% correct gender classification) did not, either pre- or post-intervention (both $p<0.001$). For the

behavioural data collected inside the scanner (analysed using a per-protocol approach for consistency with the fMRI data), n-back working memory task performance improved significantly over time ($F(1,31)=28.46, p<0.001$) and there was no main effect of group ($F(1,31)=1.01, p=0.323$). Surprisingly, performance improved significantly more in the sham than the active group (group-by-time interaction: $F(1,31)=6.17, p=0.019$). In the sham group, baseline d' was 1.41 (SD = 0.82), rising to 2.76 (SD = 1.37), a mean improvement of 1.35 (SD = 1.27); in the active stimulation group, baseline d' was 1.55 (SD = 0.72), rising to 2.04 (SD = 0.89), a mean improvement of 0.49 (SD = 0.70).

There was no effect of group on performance on the emotion processing task (accuracy of gender identification of emotional faces, assessed with a non-parametric Mann-Whitney U test) ($p>0.1$). In the sham group, baseline accuracy was 98.1% (SD = 0.02), and post-intervention accuracy was 97.5% (SD = 0.03), a mean difference of -0.01 (SD = 0.02). In the active group, baseline accuracy was 96.5% (SD = 0.08), and post-intervention accuracy was 95.0% (SD = 0.08), a mean difference of -0.02 (SD = 0.07).

N-back fMRI results

DLPFC ROI analysis (*a priori*) (full description): To examine the effect of stimulation on DLPFC activation during working memory (3- vs 1-back contrast), we conducted a repeated-measures ANOVA with within-subjects factors of laterality (left, right) and time (pre-, post-intervention) and a between-subjects factor of stimulation (active or sham). With the exception of right DLPFC activation on the post-intervention scan, data met the assumption of normality. For the right DLPFC, we conducted a nonparametric independent samples Kruskal-Wallis test, testing the effect of group on the post-intervention scan, which confirmed the results of the ANOVA ($p>0.1$).

Activation was stronger in the left than the right DLPFC (main effect of laterality: $F(1,31)=5.00, p=0.033, \eta_p^2=0.139$) and increased after the intervention (main effect of time: $F(1,31)=20.95, p<0.001, \eta_p^2=0.403$), but the interaction was non-significant ($F(1,31)=0.363, p=0.551$). There was also no interaction between laterality and stimulation condition ($F(1,31)=2.84, p=0.102$), between time and stimulation condition ($F(1,31)=0.095, p=0.760$), or between laterality, time, and stimulation condition ($F(1,31)=1.330, p=0.258$). The main effect of stimulation condition was non-significant ($F(1,31)=1.29, p=0.266$).

Emotion processing fMRI results

Amygdala ROI analysis (*a priori*) (full description): All distributions of parameter estimates for the amygdalae met the assumption of normality (Kolmogorov-Smirnov test, all $p>0.08$). To examine the effect of stimulation on amygdala activation during emotion processing (contrasts: fearful>neutral and happy>neutral), we conducted a repeated-measures ANOVA, with within-subjects factors of emotion (fear, happy), laterality (left, right), and time (pre-, post-intervention), and a between-subjects factor of stimulation condition.

There was no main effect of laterality ($F(1,31)=0.336, p=0.566$). The interaction between laterality and group narrowly missed significance ($F(1,31)=4.11, p=0.051$) and was not analysed further. Amygdala activation for the fearful>neutral contrast was significantly stronger than for the happy>neutral contrast ($F(1,31)=20.54, p<0.001, \eta_p^2=0.399$). Emotion did not interact with laterality ($F(1,31)=2.84, p=0.102$). The three-way interaction between

emotion, stimulation condition, and laterality narrowly missed significance and was not analysed further ($F(1,31)=3.65, p=0.065$).

There was no interaction between time and laterality ($F(1,31)=0.146, p=0.705$), between time and emotion ($F(1,31)=0.022, p=0.882$), between time, laterality, and stimulation condition ($F(1,31)=0.605, p=0.443$), between time, emotion, and laterality ($F(1,31)=0.699, p=0.409$), or between time, emotion, and stimulation condition ($F(1,31)=1.86, p=0.182$). There was also no four-way interaction between time, emotion, laterality, and stimulation condition ($F(1,31)=2.29, p=0.140$).

There was a significant interaction between emotion and stimulation condition ($F(1,31)=8.41, p=0.007, \eta^2=0.213$), such that those in the sham group showed greater amygdala activation for fearful faces, but lower amygdala activation for happy faces than those in the active group (however, post-hoc contrasts averaging across days and laterality did not reveal significant differences between the groups in either the fearful ($t(31)=1.110, p=0.276$) or happy ($t(31)=2.02, p=0.052$) conditions). Although there was no main effect of time ($F(1,31)=2.00, p=0.168$), importantly there was a significant time-by-stimulation condition interaction (as described in main text, $F(1,31)=5.04, p=0.032, \eta^2=0.140$), such that those in the sham group showed decreased amygdala activation (averaged across fear and happy contrasts) at post-relative to pre-intervention ($t(14)=2.63, p=0.021$), while those in the active stimulation condition did not ($t(19)=0.605, p=0.553$).

sgACC ROI analysis (a priori): We conducted a similar analysis for the sgACC. All but one distribution of parameter estimates met the assumption of normality (Kolmogorov-Smirnov test, all $p>0.2$). In the one that did not (post-intervention sgACC activation in the happy>neutral faces contrast: $p=0.042$), we performed a nonparametric independent samples Kruskal-Wallis test, testing the effect of group on post-intervention activation, which confirmed the results of the ANOVA ($p>0.3$).

There was no significant main effect of stimulation condition ($F(1,31)=1.61, p=0.214$). There was also no main effect of emotion ($F(1,31)=1.41, p=0.244$), no emotion-by-stimulation condition interaction ($F(1,31)=0.851, p=0.363$), no main effect of time ($F(1,31)=0.043, p=0.837$), no time-by-stimulation condition interaction ($F(1,31)=0.006, p=0.937$), no time-by-emotion interaction ($F(1,31)=1.28, p=0.267$), and no three-way interaction between time, emotion, and stimulation condition ($F(1,31)=0.325, p=0.573$).

Left DLPFC ROI analysis (a priori): We conducted a similar analysis for the stimulated region, the left DLPFC. The distributions of parameter estimates all met the assumption of normality (Kolmogorov-Smirnov test, all $p>0.2$). The main effect of emotion narrowly missed significance ($F(1,31)=3.91, p=0.054$), and there was no emotion-by-stimulation condition interaction ($F(1,31)=0.002, p=0.962$), no main effect of time ($F(1,31)=0.361, p=0.552$), no time-by-emotion interaction ($F(1,31)=0.038, p=0.846$), and no three-way interaction between time, emotion, and stimulation condition ($F(1,31)=0.245, p=0.624$). There was no main effect of stimulation condition ($F(1,31)=0.116, p=0.736$).

Whole-brain analysis (exploratory): we constructed two separate flexible factorial models, testing for the effects of stimulation group, time and their interaction on: (1) activation to happy compared to neutral faces; (2) activation to fearful compared to neutral faces. In both cases, F-contrasts for the main effects of time and stimulation condition revealed no significant effects at the whole-brain level; an F-contrast for the interaction between stimulation group and time also did not reveal any significant clusters at the whole-brain level (at our FWE-corrected threshold of $p=0.025$). See Supplementary Table 2.

Effect (3-back>1-back)	p (cluster level)	Extent (k)	p (voxel level)	F(1,31)	X	Y	Z	Region	Direction of effect
<i>day</i>	<0.001	243	0.024	38.31	-30	-61	41	L PPC	post > pre
	<0.001	326	0.040	35.64	27	-64	38	R PPC	post > pre
	<0.001	354	0.061	33.50	-6	47	-1	mPFC/pgACC	pre > post
	<0.001	196	0.062	33.48	39	8	35	R DLPFC	post > pre
	<0.001	333	0.063	33.33	-48	8	29	L DLPFC	post > pre
	0.865	7	0.489	22.77	-57	-13	-16	L lat. temporal	pre > post
	0.164	32	0.544	22.10	51	-37	53	R inferior parietal	post > pre
	0.037	54	0.651	20.89	-42	-37	50	L inferior parietal	post > pre
	0.387	20	0.690	20.46	-45	-61	-13	L fusiform	post > pre
	0.120	34	0.739	19.90	33	23	5	R insula	post > pre
	0.071	44	0.803	19.13	-6	-46	29	L posterior cingulate	pre > post
	0.600	12	0.857	18.42	33	-82	29	R angular gyrus	post > pre
	0.922	5	0.897	17.80	54	-37	14	R sup. temporal	post > pre
	0.621	13	0.935	17.09	-39	47	11	L DLPFC	post>pre
	0.922	5	0.992	14.89	-48	17	-4	L OFC	post > pre
	0.948	4	0.997	14.26	9	-49	26	R posterior cingulate	pre > post
<i>group</i>	0.153	33	0.020	39.40	48	-55	17	R PPC	active > sham
	0.700	11	0.187	27.92	-39	38	-13	L VLPFC	active > sham
	0.740	10	0.479	22.89	-21	53	8	L rostral PFC	sham > active
	0.922	5	0.485	22.82	-15	2	53	L sup. FG	sham > active
	0.740	10	0.533	22.24	33	-46	56	R sup. parietal	active > sham
	0.948	4	0.929	17.21	30	20	-10	R insula	active > sham
	0.948	4	0.963	16.35	42	-7	17	R prim. motor	active > sham
	0.948	4	0.990	15.09	-21	-61	32	L sup. parietal	sham > active
<i>day*group</i>	none								

Supplementary Table 2. N-back task: whole-brain results. Whole-brain results of flexible factorial analysis for the n-back task (all contrasts: 3-back>1-back; cluster-forming threshold $p<0.001$, uncorrected; minimum cluster size=4) for both main effects (day, i.e. pre-intervention vs post-intervention; and group, i.e. active vs sham tDCS) and the interaction effect. R=right; L=left; lat.=lateral; sup.=superior; prim.=primary; FG=frontal gyrus; PPC=posterior parietal cortex; mPFC=medial prefrontal cortex; OFC=orbitofrontal cortex; VLPFC=ventrolateral prefrontal cortex; pgACC=perigenual anterior cingulate cortex; sup.=superior; prim.=primary; DLPFC = dorsolateral prefrontal cortex.

Effect	<i>p</i> (cluster-corrected)	Extent	<i>p</i> (voxel-corrected)	F(1,31)	X	Y	Z	region	direction	
<i>happy > neutral</i>									pre > post	
	0.909	6	0.805	20.38	-30	-22	26	WM		
	0.275	22	0.837	19.95	51	38	-10	R OFC	pre > post	
	0.639	12	0.889	19.17	-24	-16	44	L SFG	pre > post	
	0.941	5	0.957	17.74	3	-40	17	R post. cing.	pre > post	
	0.966	4	0.987	16.48	36	-64	26	R angular gyrus	pre > post	
	0.966	4	0.998	14.99	0	-40	5	Lat. ventricle	pre > post	
<i>group</i>								L premotor	active > sham	
	0.087	35	0.095	32.83	-15	11	50			
	0.941	5	0.676	21.91	24	-28	44	R prim. sensory	active > sham	
	0.966	4	0.832	20.02	15	5	44	R dACC	active > sham	
	0.941	5	0.988	16.44	12	20	53	R SMA	active > sham	
	<i>day * group</i>	0.547	14	0.902	18.95	57	-1	-19	R lat. temp.	active ↑ sham ↓
		0.735	10	0.997	15.40	-42	14	35	L DLPFC	active ↑ sham ↓
<i>fearful > neutral</i>	0.966	4	0.997	15.36	-12	65	8	L FPC	active ↑ sham ↓	
	<i>day</i>							L fusiform	pre > post	
		0.783	9	0.289	27.05	-24	-31	-22		
	<i>group</i>	0.966	4	0.995	15.73	-3	65	17	L FPC	pre > post
		0.303	21	0.588	22.88	30	-7	14	R insula	sham > active
	0.829	8	0.955	17.77	42	-52	8	R fusiform	Sham > active	
	0.909	6	0.987	16.45	30	-52	-16	R fusiform	sham > active	
<i>day * group</i>	0.871	7	0.957	17.72	-33	-1	26	L PCG	active ↑ sham ↓	

Supplementary Table 3. Emotion processing task: whole brain results. Whole-brain results from flexible factorial for emotion processing task (thresholded at $p < 0.001$, uncorrected; minimum cluster size=4) for both main effects (day, i.e. pre- versus post-intervention; group, i.e. active versus sham tDCS) and interaction effect. R=right; L=left; WM=white matter; OFC=orbitofrontal cortex; SFG=superior frontal gyrus; lat.=lateral; prim.=primary; post.=posterior; cing.=cingulate; dACC=dorsal anterior cingulate cortex; SMA=supplementary motor area; temp.=temporal cortex; DLPFC = dorsolateral prefrontal cortex; FPC=frontopolar cortex; PCG=precentral gyrus.

Prediction of outcome analyses

The interaction between baseline left DLPFC activation and stimulation group remained significant after controlling for baseline depression, anxiety, anhedonia, and baseline n-back performance (d') in a sensitivity analysis. We first iteratively tested for multicollinearity between all independent variables. The variance inflation factor (VIF) did not exceed our threshold; therefore, the initial sensitivity model included all six variables: left DLPFC activation, baseline HAM-D, SHAPS, and BAI scores, baseline n-back performance (d'), and stimulation condition (dependent variable: percent change in HAM-D score), as well as all interactions with stimulation condition. Following stepwise elimination of variables with $p > 0.1$, left DLPFC activation remained a significant individual predictor of clinical outcome ($F(1,26)=29.90, p < 0.001, \eta_p^2=0.535$); its interaction with stimulation condition was also still significant ($F(1,26)=10.87, p=0.003, \eta_p^2=0.295$). Baseline HAM-D was also a significant predictor ($F(1,26)=12.38, p=0.002, \eta_p^2=0.323$). BAI was only a marginally significant predictor ($F(1,26)=3.12, p=0.089$), but its interaction with stimulation condition was significant ($F(1,26)=5.38, p=0.029, \eta_p^2=0.169$).

In the right amygdala model, to test the sensitivity of the model, we iteratively tested for multicollinearity between all independent variables. The variance inflation factor (VIF) did not exceed our threshold; therefore, the initial model included all six variables, as well as interactions with stimulation group. Right amygdala activation remained a significant predictor of outcome ($F(1,29)=6.52, p=0.019, \eta_p^2=0.237$) (all other predictors and their interactions with stimulation group were not, $p > 0.2$; therefore, we did not continue with backwards elimination).

Follow-up Generalised Anxiety Disorder analyses

We ran follow-up analyses testing the effect of generalised anxiety disorder (GAD) on our prediction of response models, for both DLPFC and right amygdala models. The interaction between baseline left DLPFC activation and stimulation group remained significant after controlling for baseline GAD ($F(1,27)=4.98, p=0.034$); so did right amygdala activation as a predictor of outcome ($F(1,28)=7.12, p=0.013$). In both models, the effect of GAD and its interaction with stimulation group were non-significant: in the DLPFC model, neither GAD ($F(1,27)=0.073, p=0.789$) nor its interaction with stimulation group ($F(1,27)=0.189, p=0.667$) predicted outcome; likewise, in the amygdala model, neither GAD ($F(1,28)=0.035, p=0.875$) nor its interaction with stimulation group predicted outcome ($F(1,28)=1.14, p=0.295$).

Effect (3-back>1-back contrast)	p (cluster-corrected)	Extent	p (voxel-corrected)	t(29)	Z	Y	Z	region
Response	0.118	43	0.143	5.26	6	-22	-4	L thalamus
	0.074	51	0.451	4.64	-30	14	56	L premotor
	0.158	38	0.580	4.47	24	20	53	R premotor
	0.027	69	0.633	4.40	-30	-79	32	L PPC
	0.132	41	0.755	4.24	-27	-46	41	L parietal
	0.791	10	0.814	4.15	-15	-61	26	L posterior cingulate
	0.912	6	0.847	4.10	-33	-70	-31	L cerebellum
	0.824	9	0.932	3.92	-48	-52	11	L PPC
	0.957	4	0.962	3.82	18	11	8	R caudate
Nonresponse	none							
Interaction: sham>active	0.688	13	0.320	4.85	-3	-16	2	L thalamus
	0.937	5	0.719	4.29	12	-40	-10	R cerebellum
	0.791	10	0.908	3.98	-27	-19	59	L premotor
	0.937	5	0.966	3.80	-9	2	47	L SMA
Interaction: active>sham	none							

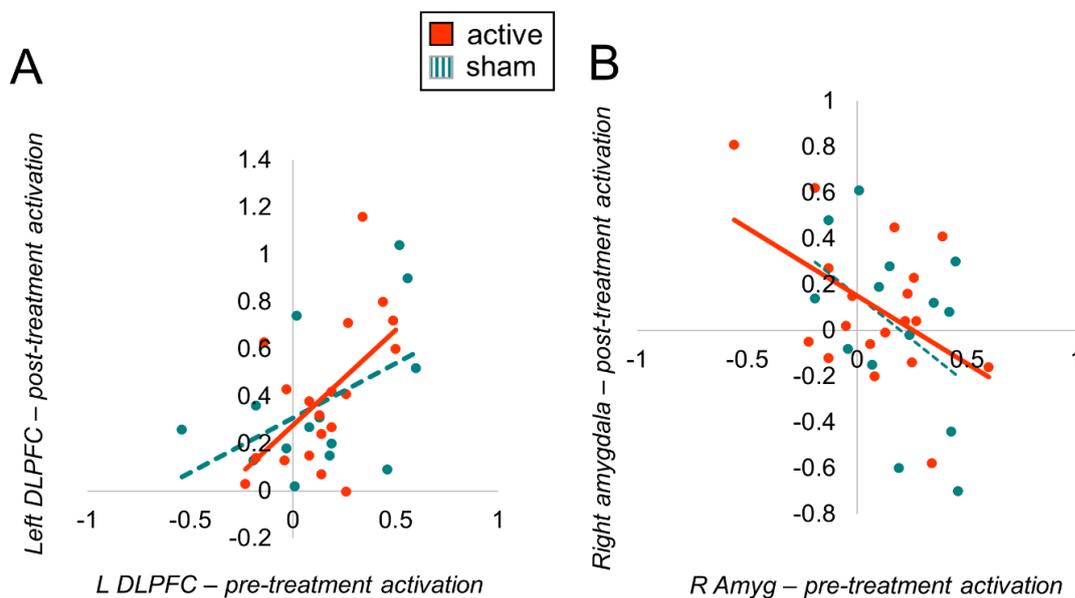
Supplementary Table 4. *N-back task: whole-brain correlates of response.* Whole-brain correlates of response for n-back task (cluster-forming threshold $p < 0.001$, (uncorrected); minimum cluster size=4). R=right; L=left; PPC=posterior parietal cortex; SMA=supplementary motor area.

Effect (fearful>neutral contrast)	p (cluster-corrected)	Extent	p (voxel-corrected)	t(29)	Z	Y	Z	region
Response	0.001	121	0.110	5.49	-18	59	-7	L rostral PFC
	0.966	4	0.932	4.04	-21	-16	-1	L GP
	0.678	12	0.933	4.04	12	-82	-28	L cerebellum
	0.845	8	0.945	4.01	6	-4	-16	R hypothalamus
	0.915	6	0.982	3.84	27	-61	-37	R cerebellum
	0.915	6	0.993	3.73	60	-16	-25	R temporal
Non-response								
Interaction: sham>active	0.882	7	0.383	4.86	6	-25	-10	R substantia nigra
	0.915	6	0.730	4.39	-9	17	11	L caudate
	0.882	7	0.946	4.00	3	5	8	R thalamus
	0.678	12	0.994	3.72	6	47	-10	R rostral PFC
	0.966	4	0.999	3.55	-3	44	-22	L VMPFC
Interaction: active>sham	none							

Supplementary Table 5. *Whole-brain correlates of response (fearful>neutral).* Whole-brain correlates of response for emotion processing task (cluster-forming threshold $p < 0.001$ (uncorrected); minimum cluster size=4). R=right; L=left; PFC=prefrontal cortex; VMPFC=ventromedial prefrontal cortex; GP=globus pallidus.

Effect (happy > neutral contrast)	p (cluster-corrected)	Extent	p (voxel-corrected)	t(29)	X	Y	Z	region
<i>Response</i>								Cingulate gyrus
	0.526	16	0.872	4.16	-18	5	29	
	0.882	7	0.954	3.96	6	2	-7	BNST
<i>Non-response</i>	.915	6	0.805	4.27	-24	2	5	L putamen
<i>Interaction: sham>active</i>	none							
<i>Interaction: active>sham</i>	0.767	19	0.812	4.26	24	32	-4	Basal operculum

Supplementary Table 6. Whole-brain correlates of response (happy>neutral). Whole-brain correlates of response for the happy>neutral faces contrast in the emotion processing task (cluster-forming threshold $p < 0.001$, uncorrected; minimum cluster size=4). BNST=bed nucleus of the stria terminalis.



Supplemental Figure 2. Within-subject reliability of activation between pre- and post-treatment scan days. A: left DLPFC (ICC=0.67, 95%CI=0.33–0.84, $p=0.001$); B: right amygdala (ICC=-1.630, 95%CI=-4.325– -0.299, $p=0.996$).

Exploratory analysis of predictors of change in DLPFC activation

Repeating the methods of our outcome-prediction regression models, we tested whether anhedonia, baseline depression, previous number of depressive episodes, IQ, previous experience of psychological therapy, previous experience of antidepressant medication, previous hospitalisations, or previous suicides could predict subsequent change in DLPFC activation (across both sham and real groups). None were predictive of the degree of increase in our *a priori* ROI-based DLPFC activation (calculated as the average change across the ROI from pre-to-post treatment): all $p > 0.5$; therefore, we did not analyse any individual predictors further.