medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Paving the way for precision treatment of psychiatric symptoms with functional connectivity neurofeedback

Taylor JE[1], Oka T[1,2], Murakami M[1], Motegi T[1,3], Yamada T[1,4,5], Kawashima T[6], Kobayashi Y[6], Yoshihara Y[6], Miyata J[6,7], Murai T[6], Kawato M[1], & Cortese A[1].

[1] The Department of Decoded Neurofeedback, Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International, Kyoto, Japan

[2] The Department of Clinical Psychology, Graduate School of Human Sciences, Osaka University, Suita, Japan

[3] Saint-Pierre Hospital, Takasaki, Japan

[4] Department of Cognitive, Linguistic and Psychological Sciences, Brown University, Providence, USA

[5] Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan

[6] Department of Psychiatry, Graduate School of Medicine, Kyoto University, Kyoto, Japan

[7] Department of Psychiatry, Aichi Medical University, Aichi, Japan

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

Despite the prevalence of Major depressive disorder (MDD), a large proportion of patients do not respond well to its existing treatments. Patients with MDD have heterogeneous transdiagnostic subsets of symptoms with differing underlying neural aberrations. Therefore, better treatment response might be achieved using more customizable treatments. Showing promise for this, brain-machine interfaces (BMIs) can be used to directly target patient-specific underlying neural aberrations. As a major step in this direction, here we reproduce and extend, with a larger sample, our previous findings that a BMI technique called Functional Connectivity Neurofeedback (FCNef) can normalize neural aberrations related to specific MDD symptoms. For the first time, we show that normalization of the target neural activity (here, connectivity between the dorsolateral prefrontal cortex and the precuneus) corresponds meaningfully more to reductions in corresponding than non-corresponding symptoms (here, significantly more to reductions in rumination than anxiety symptoms). Furthermore, we showed for the first time that this depended on the specific parameters that FCNef was run with. Specifically, normalization of the targeted neural activity and a corresponding reduction in related symptoms was greater with *more external reward* and with *consecutive* (compared to non-consecutive) days of training, but did not differ depending on whether participants were given shorter or longer time-windows to manipulate their neural activity. Overall, these findings demonstrate the promise of FCNef for precision medicine and highlight the importance of BMI parameter testing for enhancing the feasibility of actual clinical trials. Hereby, we inch closer to a future where signals from our own brains are used to guide our own individual medical interventions.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

The World Health Organization estimates that, globally, 5% of adults suffer from depression[1]. They project that major depressive disorders will become the top cause of global disease burden by 2030[2]. Of the patients that do receive treatment, a whopping number- with estimates in the range of 29-46%[3]- are estimated to not respond in full. One likely contributor to this underwhelming response rate is that individual patients, despite often having heterogeneous and transdiagnostic subsets of symptoms related to different underlying neural mechanisms[4], usually receive relatively homogenous treatment. For example, all or most clinical practice guidelines recommend Selective Serotonin Reuptake Inhibitors as first-line treatment regardless of depressive subtype[5]. Clearly, individual differences need to be better taken into consideration for more effective treatment. In the not-so-distant future, treatment may become more individualized using Brain-Machine Interfaces (BMIs), which can be used to identify and target patients' own underlying individual neural aberrations and thereby guide their own individual medical intervention. However, to date, only a handful of BMIs have received approval from local medical regulatory agencies for human trials[6–8], and even fewer have received full market authorization[9,10]. A key step towards approval of a given BMI technique by regulatory agencies is demonstrating the optimality of the chosen parameters. Here, we describe our efforts to discover an optimal set of parameters to improve outcomes with a specific form of BMI called Functional-Connectivity Neurofeedback (FCNef), which has proven promising for treating specific subsets of psychiatric symptoms[11].

FCNef is a functional magnetic resonance imaging (fMRI) neurofeedback technique where participants have their brains scanned and receive external feedback in real-time. Importantly, this feedback depends on the current state of functional connectivity between targeted brain regions (measured as the correlational relationship between time-courses of BOLD activity from these regions). The feedback is used to train participants to move a targeted functional connection in a specific direction (e.g., make it more positive), a result that has been successfully and repeatedly demonstrated in previous studies[11–18]. Because aberrations in brain connectivity patterns may relate to specific psychiatric symptoms[19–26], any method that can normalize these patterns (restore them to a healthy state) may aid in symptom amelioration. Showing great promise for this, recent FCNef studies have shown precise correspondence between normalization of functional connections and reductions in specifically related

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

symptoms[11,16,17,27]. Importantly, these effects endured for at least one-two months after FCNef[11]. We therefore previously proposed that FCNef has high promise for use in precision medicine[11,28].

Critically, given its potential use as a medical tool in the future, FCNef should primarily maximize health outcomes while minimizing patient burden. However, to date and to the best of our knowledge, no study has examined the parameters under which FCNef can most successfully do so. Here, we therefore sought to rigorously investigate the core parameters of our previously reported FCNef paradigm, which was designed with the goal of reducing depressive symptoms (hereby, our "FCNef for depression" paradigm)[11,18]. In this paradigm, we targeted functional connectivity between the dorsolateral prefrontal cortex and the precuneus/posterior cingulate cortex (DLPFC-PCC FC) with the goal of making this more anticorrelated, as is seen in healthy controls relative to participants with melancholic depression [11,29]. Importantly for precision medicine, we expected normalizing this specific functional connection to reduce maladaptive rumination (brooding[30]) symptoms but not anxiety symptoms, given their differences in underlying neural circuitry (e.g. see Williams et al., 2016, 2017[23,24]). In the current paper, we extend our previously reported results with new analyses in a far larger sample size. We further extend previous results by, for the first time, outlining our efforts to find parameters that would optimize beneficial outcomes while also keeping participant fatigue to a minimum (because preliminary testing in a clinical sample resulted in the fatigue-related drop-out of one out of six patients[31]). With these goals in mind, we focused on the following three parameters.

1) Reward schedule: During real-time neurofeedback tasks, the feedback has conventionally been provided simply as scores that reflect how similar the induced brain activity is to the target brain activity[32–34]. However, recent evidence suggests that the target neural activity may be better reinforced during neurofeedback when, in addition to feedback scores, external reward (e.g. money) is also used[35]. Here, we manipulated how participants were assigned bonus money to their feedback scores so that different groups of participants could earn less/more overall external reward.

2. Induction time-window: During preliminary studies with our FCNef design, participants were instructed to manipulate their brain activity on each trial during a 40s induction period[11,18]. However, newer neurofeedback studies have shown success with shorter induction time-windows[13–15,36] and reanalysis of some of our preliminary data indicated that participants were most effective in manipulating their targeted functional connectivity within the first 20s of

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

our 40s induction time-window (see Supplementary Results and Supplementary Table 2). We therefore set out to examine if we could still obtain successful FCNef results when the induction time-window on each trial of our experiment was shortened from 40s to 20s. This would half the cognitive effort required by participants on each trial and reduce the overall length of the experiment by ten minutes each day.

3. Experimental schedule: In our original FCNef for depression studies[11,18], for consistency with most past FCNef studies (e.g. Megumi et al., 2015[14], but see Koush et al., 2017[13]), participants came in for multiple consecutive days of experimentation. Coming in for *consecutive* days can be exhausting and requires motivation and organization skills, all of which can make this paradigm particularly difficult for people with psychiatric symptoms. We therefore tested whether a more flexible schedule, over non-consecutive days, could yield similar results to the consecutive training schedule.

Overall, we ran 68 participants in our FCNef for depression paradigm while manipulating reward schedule (low/high reward), induction time-windows (20s/40s), and experimental schedule (continuous/non-continuous training days). Our goal was to fine-tune our FCNef for depression paradigm so that it could be optimally successful. FCNef success was operationalized here by normalization of the targeted functional connectivity and a related reduction in symptoms. First, using a larger sample size, we strengthen our previous findings showing that normalization of DLPFC-PCC functional connectivity relates to reductions in brooding rumination symptoms (which are thought to relate to this targeted functional connectivity specifically) significantly more than to reductions in anxiety symptoms (which are thought to have different underlying neural mechanisms). Furthermore, we found (1) that a high compared to low reward feedback schedule led to greater FCNef success, (2) no meaningful significant differences in FCNef success regardless of whether a 20s or 40s induction time-window was used, and (3) that the aforementioned promising results did not replicate when FCNef took place over non-consecutive days. The reproduction of our previous results shows high promise for the use of FCNef for precision medicine in the future. The parameter findings pave the way for improved/optimized FCNef interventions. Furthermore, they highlight the importance of BMI parameter testing for enhancing the feasibility of actual clinical trials, thereby bringing BMIs one step closer to the clinic.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

Methods

Ethics statement

Participants all provided written informed consent on each day of screening and experimentation prior to commencement. This study was approved by the Ethics Committee of the Review Board of Advanced Telecommunications Research Institute International, Japan (Ethics No.132, 172) and by the Kyoto University Certified Review Board (UMIN000015249, jRCTs052180169). All experiments were performed in accordance with the guidelines and regulations of this Ethics Committee.

Participants

Participants were screened twice using the Beck Depression Inventory-II (BDI) questionnaire (see Supplementary Methods for more detail)[37]. Only those with an average score >8, who indicated no intent for suicide, who spoke Japanese, who held no current clinical diagnosis, and who were not currently receiving treatment for a psychiatric illness were invited to participate. Overall, 69 people passed the screening and participated in the main experiment. These participants had an average BDI score of 14.33 (std = 5.26), which puts them generally in the category of “mild depression”[37] . Based on these BDI scores and the fact that these participants held no current psychiatric diagnoses, we will hereby describe them as having “subclinical” symptoms of depression.

Participants were put into six experimental groups- each with differing conditions related to the experimental parameters of interest (see Table 1). Which participants were run in which group depended largely on the schedule and length of availability of the participant, experimenter, and MRI machine at the time. For data analysis, the participant groups were then split into two datasets (see Table 1). Because no meaningful effects of interest were found for the parameter of Induction Time-Window, participant groups within these datasets were then collapsed across this parameter in the figures (this is also described in Table 1). It should be noted that the data of 20 participants (those from the C-40-H and C-40-L groups) has been reported elsewhere[11]. Here, as in this previous report, the experimental data of one of these participants was excluded from data analysis because it was revealed - subsequent to

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

experimentation- that they held a current clinical diagnosis. Further detail about recruitment, participant demographics and payment can be found in the Supplementary Methods.

	Dataset	Group Name	Experiment Schedule	Induction Time-Window	Reward Schedule	Main Experiment Sample Size	1-Month Follow-Up Sample Size	2-Month Follow-Up Sample Size
Consec/ High Rew	1	C-40-H	Consec	40s	High	9	N/A	N/A
	1	C-20-H	Consec	20s	High	12	11	9
Consec/ Low Rew	1 & 2	C-40-L	Consec	40s	Low	10	9	8
	1 & 2	C-20-L	Consec	20s	Low	13	11	9
Non/ Low Rew	2	N-40-L	Non-Consec	40s	Low	12	11	9
	2	N-20-L	Non-Consec	20s	Low	12	11	11
						Total = 68	Total = 53	Total = 46

Table 1.

This table displays the six experimental conditions of interest and the specifics of the different parameters employed within each condition. It shows the number of participants who participated in each condition for the main experiment, and for the one- and two-month follow-up tests. Furthermore, this table show how the six experimental conditions were split into two overlapping datasets for analysis. It also shows how- because no effects of interest were found for the Induction Time-Window parameter- this data can be collapsed into three major groups of participants (shown in color on the left- these colors correspond to the colors used in the figures throughout this article). N/A = not applicable, because follow-up tests were not conducted for the C-40-H group. Consec = Consecutive. Non-Consec = Non-Consecutive.

Experimental procedure, materials, and imaging data acquisition

An outline of the experimental procedure is displayed visually in Figure 1 alongside details of the FCNef task itself. These were largely the same as our previous report[11], except for a few differences which were designed to allow for testing between experimental conditions of interest. These differences are highlighted in the sub-section below entitled “Differences in experimental procedure”. The protocol and imaging data acquisition details are identical to those in our previous report[11] and are summarized in the Supplementary Methods.

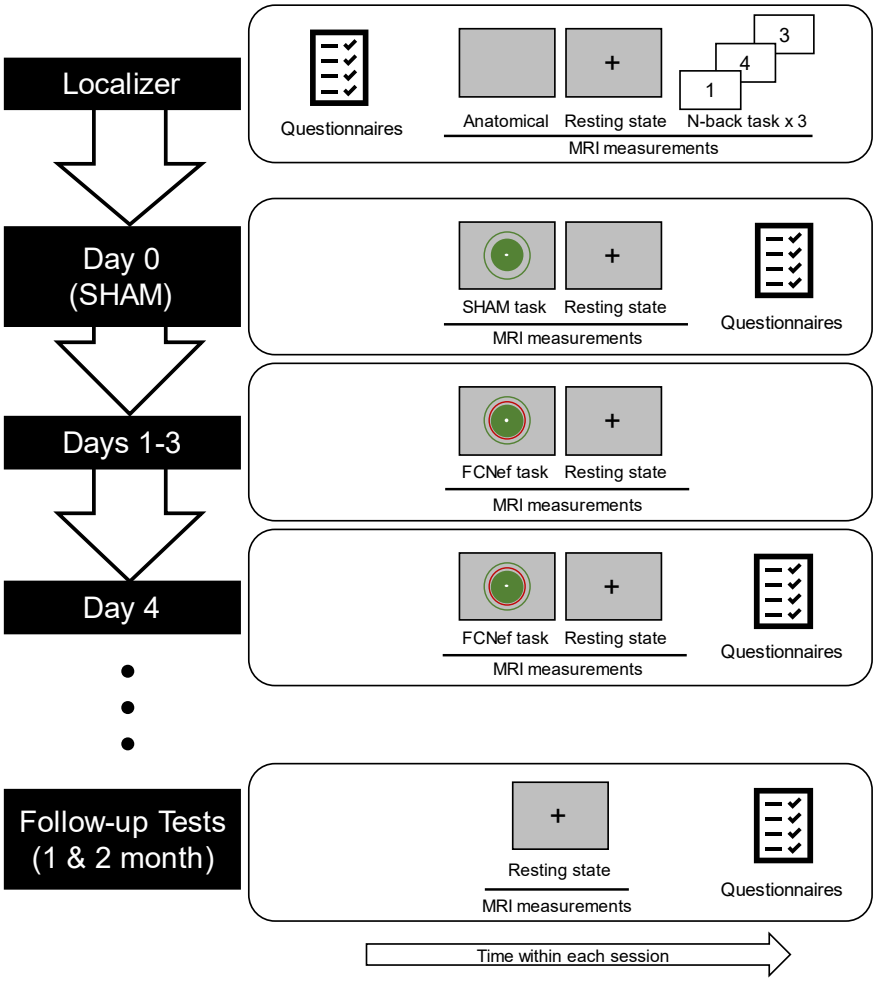
Details of the symptoms questionnaires can be found in the Supplementary Methods, but overall general depressive symptoms were measured with the BDI[37], brooding rumination symptoms were measured with a subscale of the Rumination Response Scale (RRS)[30,38], and trait anxiety symptoms were measured with the trait anxiety subscale of the State-Trait Anxiety

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

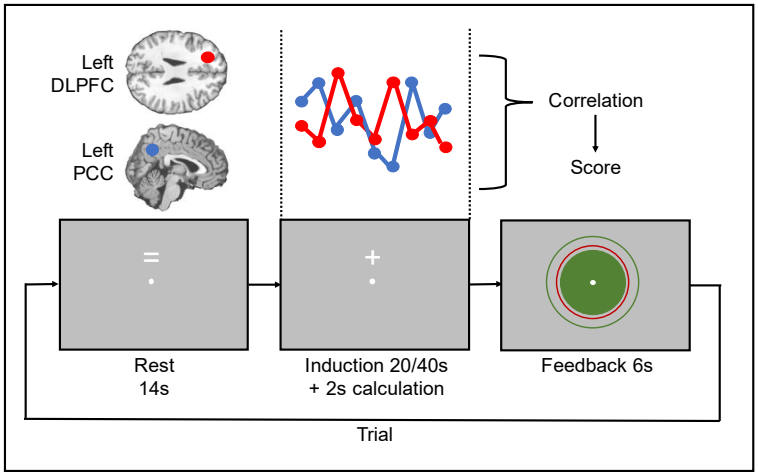
Inventory (STAI-Y2)[39]. As can be seen on Figure 1, because there would be limited clinical meaning, these symptom questionnaire scores were not measured on all days of the main experiment. Instead, they were only measured on the first day (Day 0) and last day (FCNef Day 4) of the main experiment and during the one- and two- month follow-up tests.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

a.



b.



medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

Figure 1. Experimental procedure and example FCNef trial.

(a) The order of events that happened on each day of experimentation. *Questionnaires = the Beck Depression Inventory-II[37], the Rumination Response Scale[30], and the State-Trait Anxiety Inventory[39]. Anatomical = T1-weighted structural MRI. N-back = a well-known executive control task[40], used here as a functional localizer[11].*

*(**b) An example FCNef trial.** During the rest period, participants should simply relax. During the induction period, they should "somehow" manipulate their brain activity to get the best possible feedback. Participants were told that different strategies of brain activity manipulation might work for different people. Unbeknown to participants (nothing changed on screen), there was a 2s calculation period at the end of the induction period. During FCNef, DLPFC-PCC connectivity (from the induction period) was calculated during the calculation period and this determined the feedback presented during the feedback period (but during SHAM, feedback was just random). Feedback was presented on screen as a green circle and participants had been clearly instructed that the larger this was the more monetary reward they would receive on that trial. During FCNef, they were instructed to try to make the green circle bigger than a red circle that was additionally presented on screen. The circumference of this red circle represented the participant's baseline DLPFC-PCC connectivity (the average from SHAM). During SHAM, there was no red circle and participants were simply instructed to try to make this green circle as big as possible. Modified with permission from Taylor et al. (2022)[11].*

Differences in experimental procedure

The following three parameters were manipulated to differ between experimental conditions. This means that some of the conditions here had different parameters to those used in our previous report, which were specifically the C-40-L and C-40-H groups from Table 1[11].

Reward schedule:

All participants received a baseline reward bonus of ¥500 on each day for both the SHAM and FCNef tasks. This is the maximum reward that they could receive in the SHAM task, but they could receive an additional reward bonus in the FCNef task depending on their average FCNef scores from that day. The specific way in which scores were calculated in the SHAM and FCNef tasks is the same as in our previous report[11] and is described in the Supplementary Methods. Importantly for here, when calculating the additional reward bonus, different calculation methods were used in different conditions for assigning money to the FCNef scores. In reality, this meant that participants in some groups got *higher* reward overall (usually,

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

baseline bonus + some additional bonus) and participants in other groups got *lower* reward overall (usually only the baseline bonus, but in a sparse number of trials some participants also got some small additional bonus). For ease of expression, we therefore refer to these as the high and low reward schedule conditions. Participants in the high reward schedule conditions (C-20-H and C-40-H; the *H* refers to *high* reward) could achieve an additional reward bonus of ¥50 for each average FCNef score point over 50. Participants in the low reward schedule conditions (C-20-L, C-40-L, N-20-L, N-40-L; the *L* refers to *low* reward) could achieve an additional reward bonus of ¥100 for each average FCNef score point over 75. See Supplementary Table 1 for specific examples.

Induction time-window:

Three groups of participants completed FCNef and SHAM with a 40s induction time-window; These were the groups: C-40-L, N-40-L, and C-40-H (the *40* refers to the *40s* induction time-window). Three other groups of participants completed FCNef and SHAM with a 20s induction time-window; These were the groups: C-20-L, N-20-L, and C-20-H (the *20* refers to the *20s* induction time-window)

Experimental schedule:

Four groups of participants completed SHAM, and FCNef Days 1-4 over 5 consecutive days (Monday-Friday); These were the groups: C-20-L, C-40-L, C-20-H, and C-40-H (the *C* refers to the *consecutive* experimental schedule). Two other groups of participants completed SHAM, and FCNef Days 1-4 over 5 non-consecutive days over a period of several weeks to months (mean of 18.5 days ± std of 15.3 days). These were the groups: N-20-L, and N-40-L (the *N* refers to the *non-consecutive* experimental schedule).

Data analyses

Linear Mixed Effect (LME) models

In each dataset separately, we ran analyses to investigate changes across time in: (1) mean FCNef scores, (2) FCNef score variance, (3) symptoms, and (4) resting-state functional connectivity (rs-FC). We also ran LME models to examine the relationships between changes in

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

different symptoms and changes in rs-FC. Each of these analyses included multiple LME models, which are described and compared in tables in the Supplementary Results. We briefly describe here the process of determining the best-fit model for each analysis. In level one, we started with a simple model which included just one independent variable and an intercept (e.g. *'Score ~ Day'*, in Wilkinson notation; Score = FCNef task score; Day = FCNef day 1-4). In level two, we ran an additive model which included the independent variable from level one plus regressors for the manipulated parameters for the given dataset (e.g. for Dataset 1 *'Score ~ Day+TW+Reward'*; TW= induction time-window (20s/40s); Reward = reward schedule (low/high)). In level three, we ran several different models, each with the same regressors as in level two plus one additional interaction (e.g. *'Score ~ Day*TW+Reward'*, *'Score ~ Day+TW*Reward'*, and *'Score ~ TW+Reward*Day'*). In level four, we ran a model with all possible interactions (e.g. *'Score ~ Day*TW*Reward'*). We next used likelihood ratio tests to compare models. Initially, the level one model was tested against the level two model. Then, going up through the levels, each model was tested against the best-fit model from the levels below it. At each stage, the best-fit model was determined by choosing the model with the lowest Akaike Information Criterion (AIC, a measure of prediction error, which can be used to assess a model's relative quality [41]) if there was a significant difference between models, or by choosing the simplest model if there was no significant difference between models. In the case where there were no significant differences and the models were of equal complexity, then we chose the model with the lowest AIC. Finally, we used a likelihood ratio test to determine whether the best-fit model found would be improved by adding a random intercept for experimental subject (e.g. *'Score ~ Day*TW*Reward+(1|Subject)'*).

Follow-up statistical testing

For each best-fit LME, if a significant main effect or interaction was found then follow-up t-tests or correlations were run. We applied False Discovery Rate (FDR) correction[42] whenever there were multiple comparisons. We had strong directional hypotheses that from before- to after- FCNef symptoms should reduce and rs-FC should become more negative (in line with healthy people). We therefore ran the related t-tests with one-tail. Likewise, we had strong directional hypotheses that symptoms should reduce as the rs-FC became more negative (i.e. a positive correlation between changes in symptoms and changes in rs-FC) and so we ran the related correlations with one-tail.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Results

We examined participants' FCNef success in terms of task performance, changes in rs-FC, changes in symptomatology, and, most importantly, in terms of how changes in rs-FC related to changes in symptomatology. Note here that "changes" refers to changes in the data from before- to after- FCNef and is defined as the baseline data from Day 0 (SHAM) subtracted from the data from Day 4 or the follow-up tests. For each analysis, we examined how the manipulated parameters (induction time-window and reward schedule for Dataset 1; induction time-window and experimental schedule for Dataset 2) may have impacted FCNef success.

FCNef scores

We examined task performance by using LME models to assess FCNef scores. We first sought to determine if the FCNef scores increased significantly across training days because this would indicate that participants had responded to the FCNef task successfully. We second sought to determine if the variance in FCNef scores decreased significantly across training days because this would indicate that participants had gained better control of the DLPFC-PCC FC. Importantly, we examined whether these effects were impacted by the manipulated parameters for each dataset. Overall, results from the best-fit LME models and follow-up analyses (Supplementary Tables 3-10) only revealed evidence for ***learning effects*** (i.e. changes that occurred ***over the days of FCNef***) in Dataset 1. No such learning effects were found for Dataset 2. Specifically, in Dataset 1, results showed that, over training days, average scores significantly increased and score variance significantly decreased in the high reward condition (see Figure 2). These results were dependent on time-window, with significant changes over time only being found in the 40s time-window high reward condition (C-40-H; $p_{FDR} < 0.01$); however, it should be noted that results were in the same direction for the 20s time-window high reward condition (C-20-H). Additional (non-learning effect) results are reported in the Supplementary Results.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

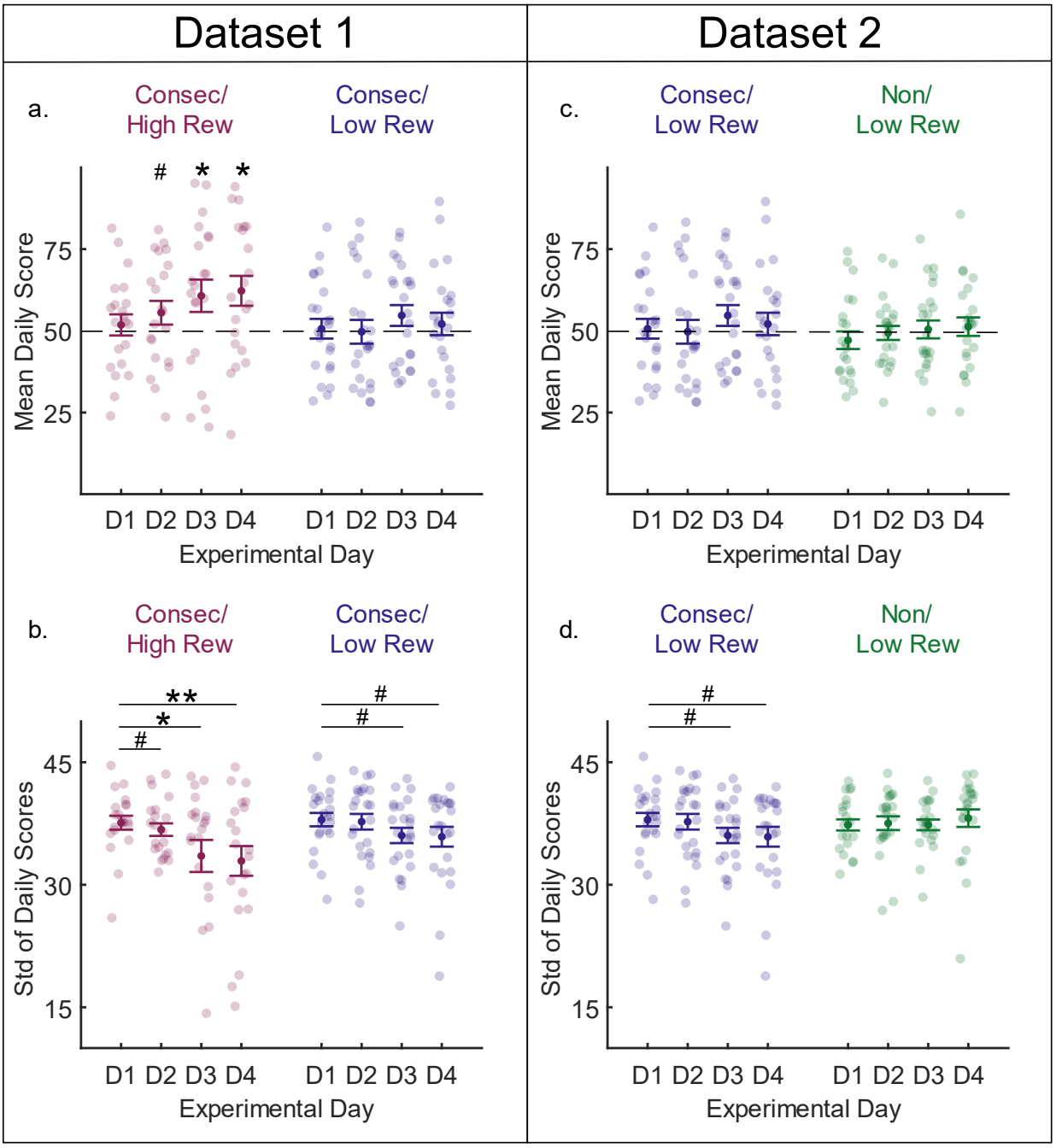


Figure 2. Means and Standard Deviations of FCNef scores over days of FCNef

*Results are collapsed across the different induction time-window conditions; See Table 1 for details. ** represents $p_{FDR} < 0.01$, * represents $p_{FDR} < 0.05$, # represents $p_{FDR} < 0.1$. Std = standard deviation. D = Day.*

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

(a) Mean daily FCNef scores for Dataset 1. *Investigative t-tests, for each group separately, were conducted to examine whether scores on each FCNef day were higher than a baseline score of 50 (which was the average score given during SHAM on Day 0). Scores were found to significantly improve over days of FCNef only for the Consec/High Rew group.*

(b) Daily Stds in FCNef scores for Dataset 1. *Investigative t-tests, for each group separately, were conducted to examine whether variance in scores on each subsequent day of FCNef was lower than that on the Day 1 of FCNef. Score variance was found to significantly reduce over days of FCNef only for the Consec/High Rew group, but a similar trend was found for the Consec/Low Rew group.*

(c) Mean daily FCNef scores for Dataset 2. *These did not significantly improve- relative to baseline- over FCNef days for either group.*

(d) Daily Stds in FCNef scores for Dataset 2. *Participants in the Consec/Low Rew group showed the same results as reported in (b). Participants in the Non/Low Rew group showed no significant differences from Day 1 on any of the other days.*

Self-report symptom questionnaire scores

For each dataset and symptom type (general depression, brooding rumination, and anxiety) separately, post-FCNef symptoms (i.e. those from FCNef Day 4 and from the one- and two- month follow-up tests) were compared with pre-FCNef symptoms (i.e. those from Day 0) to ensure that none had worsened. See the Supplementary Results for detailed t-test results and Supplementary Table 11 for the average and standard deviation of questionnaire/subscale scores from each measurement. Overall, participants' symptom scores for both datasets for all symptom types were significantly lower on FCNef Day 4 than they were on Day 0 ($p_{FDR} < 0.05$). For Dataset 1, participants' symptom scores for all symptoms types remained significantly lower at the follow-up tests one-month later ($p_{FDR} < 0.05$). For Dataset 2, a similar pattern was found (this was significant for general depressive and brooding scores, but $p_{FDR} = 0.058$ for anxiety scores). For both datasets, symptoms only remained significantly lower at the follow-up tests two-months later for brooding ($p_{FDR} < 0.05$), but general depressive and anxiety symptoms were still numerically lower. Overall, these results provide assurance that our paradigm was not affecting our participants negatively.

Next, we wished to see if changes in symptoms changed themselves in the long-term. To examine this, we ran LME models to predict "post-FCNef symptom change" (post-FCNef symptom data - SHAM symptom data) over "post-FCNef Days" (FCNef D4, one-month later,

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and two-months later). The best-fit model for each dataset showed no significant main effects of or interactions with post-FCNef Days on general depressive, brooding, or anxiety scores (see Supplementary Tables 12-20). These results therefore indicate that once symptoms had decreased (from before- to after- FCNef) they then remained lowered in the long-term. The only significant effect that was found in any of these LMEs was a main effect of time-window (TW) in the best-fit LME to predict post-FCNef symptom change for Brooding in Dataset 1 (details in Supplementary Results and Supplementary Table 16); However, because TW did not interact with post-FCNef Day here (meaning it wasn't a learning effect), and because no effect of TW was found in analyses where symptom changes were related to changes in rs-FC (below), this finding may be spurious and is not likely meaningful for our paradigm.

Resting-state functional connectivity

Given the right parameters (Dataset 1), rs-FC changes with FCNef training

For each dataset separately, we first used LME models to examine if DLPFC-PCC rs-FC changed over the course of the main experiment (over Days 0-4). If our FCNef intervention was successful, then we would specifically expect DLPFC-PCC rs-FC to decrease over the experimental days because this is what we reinforced during training. See Supplementary Table 21 for the average and standard deviation of rs-FC from each measurement. Overall, for Dataset 1, the results of the best-fit LME (see Supplementary Table 22) ***did*** show this hypothesized main effect of Day (see Supplementary Table 23). Follow-up t-tests confirmed that the DLPFC-PCC rs-FC was significantly more negative on the days of the main experiment (Days 1-4) when compared to the baseline rs-FC taken after SHAM on Day 0 ($p_{FDR} < 0.05$; see Supplementary Results). This LME model also showed a significant main effect of Reward (see Supplementary Table 23), but this was not supported by any significant results in the follow-up t-tests (see Supplementary Results) and so remains to be further investigated. Overall, for Dataset 2, the results of the best-fit LME (see Supplementary Table 22) ***did not*** show this hypothesized main effect of Day (see Supplementary Table 24). Follow-up t-tests confirmed that the DLPFC-PCC rs-FC did not significantly differ from Day 0 to the subsequent FCNef days of the main experiment ($p_{FDR} > 0.05$; see Supplementary Results).

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

rs-FC changes in the long-term

For each dataset separately, we second used LME models to examine if any post-FCNef changes in DLPFC-PCC rs-FC differed or remained stable over time. Post-FCNef changes in rs-FC were defined as rs-FC data from Day 0 (baseline from SHAM) subtracted from rs-FC data from the post-FCNef Day (i.e. from FCNef Day 4 and the one- and two-month follow-up tests). By examining how these changes themselves may have differed over time, we were able to examine the persistence (or lack thereof) of long-term effects. Overall, for both datasets, the results of the best-fit LMEs (see Supplementary Table 25) showed no main effect of post-FCNef Day (see Supplementary Tables 26 and 28), which indicates that rs-FC change itself did not increase or decrease over the multiple post-FCNef Days. This result actually has different implications for each dataset. For Dataset 1, rs-FC ***was*** more negative in the post-FCNef Days than on Day 0 (see Supplementary results); Therefore, no change across post-FCNef Days means that, overall, rs-FC became more negative from before- to after-FCNef and then remained more negative in the long-term. For Dataset 2, however, rs-FC ***was not*** more negative in the post-FCNef Days than on Day 0 (see Supplementary results); Therefore, no change across post-FCNef measurements means that, overall, rs-FC did not change from before- to after-FCNef and remained unchanged in the long-term.

Here it is important to note (a) that stark differences in results were found between Dataset 1 (where FCNef seemed to work) and Dataset 2 (where FCNef did not seem to work), and (b) that the only consistent difference between these datasets was the deliberate manipulations in FCNef parameters. The best-fit LME models to examine post-FCNef changes in DLPFC-PCC rs-FC ***did not*** actually include regressors for these parameters; this indicates that the ***within***-dataset effects of these regressors were not large enough to account for a significant proportion of the variance. However, because we wished to examine the effects of these parameters in general and because we also wished to examine why the aforementioned differences in results ***between-***datasets arose, we therefore next examined, for each dataset, the best-fit of the LME models which ***did*** include regressors for these parameters of interest. The best-fit of these models for Dataset 1 revealed a significant main effect of Reward (see Supplementary Table 27). Follow-up t-tests clearly showed that rs-FC was reduced significantly in the long-term for only the high reward condition ($p_{FDR} < 0.05$; see Supplementary Results). On the contrary, the best-fit of these models for Dataset 2 did not show any results of significance

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

(see Supplementary Table 29). These results indicate that the FCNef success found in Dataset 1 but not in Dataset 2, may have been driven by the high reward condition of Dataset 1. See Figure 3.

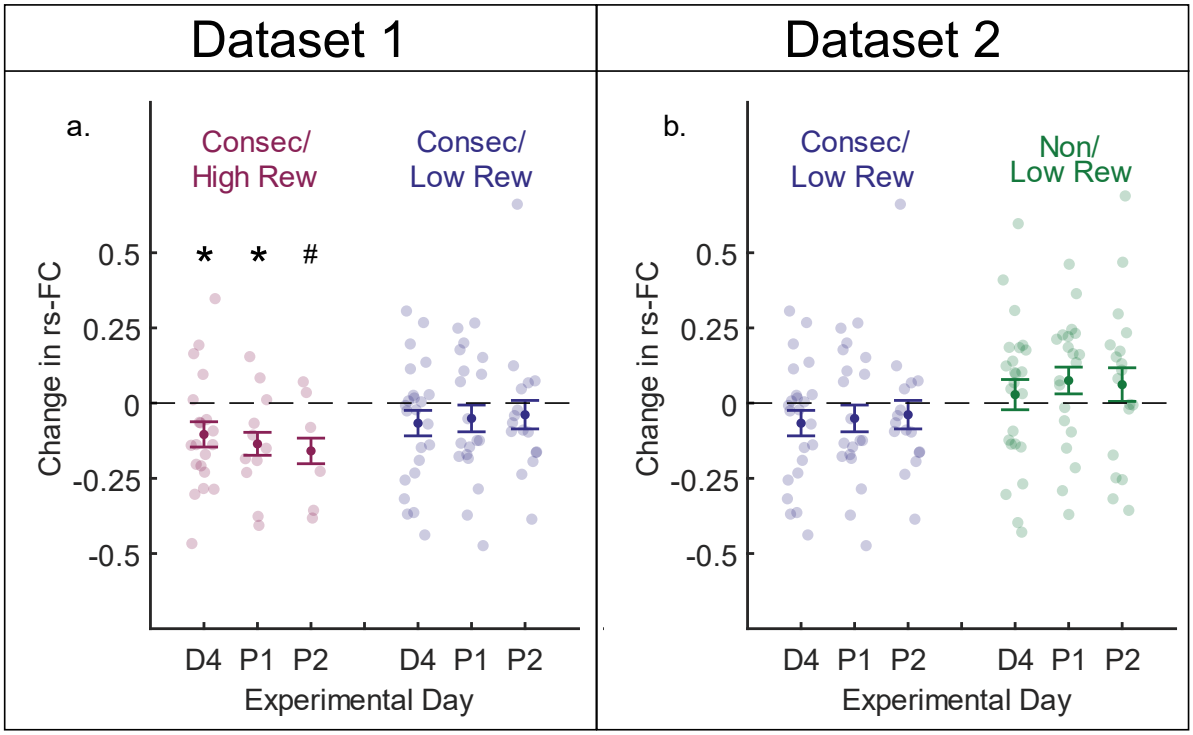


Figure 3. Long-term changes in rs-FC.

*Displayed results are collapsed across the different induction time-window conditions; See Table 1 for details. Note that the different bars on each subplot represent different sample sizes (see Table 1). D = Day, M = Month, * represents $p_{FDR} < 0.05$, # represents $p_{FDR} = 0.05$*

(a) *For Dataset 1, investigative t-tests revealed that for the Consec/High Rew group, overall rs-FC was significantly more negative than baseline immediately after FCNef had been completed (D4) and one-month later (1M), and that it trended towards this even two-months later (2M). Nothing of significance was found for the Consec/Low Rew group.*

(b) *For Dataset 2, overall rs-FC did not significantly change over experimental days for either group. However, visually the rs-FC appears to shift in the non-targeted direction (positive) for the Non/Low Rew group. These results should be further explored using larger sample sizes for the long-term data.*

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

Relationship between changes in self-report symptom questionnaire scores and changes in rs-FC

We sought to examine how changes in participants' DLPFC-PCC rs-FC from pre- to post- FCNef related to changes in general depressive, brooding, and anxiety symptoms. Here, changes were defined as Day 0 data subtracted away from FCNef Day 4 data. If the FCNef paradigm was successful, then we would expect changes in the DLPFC-PCC rs-FC to relate positively to changes in general depressive symptoms (i.e. the more the rs-FC became negative, the more the symptoms decreased). If, as previously hypothesized[11], the targeted functional connection (FC) is *specifically* related to maladaptive symptoms of rumination, then we would expect changes in the DLPFC-PCC rs-FC to positively relate to changes in brooding symptoms, but not to changes in anxiety symptoms (which we used as a control). Importantly, we examined whether these effects were impacted by the manipulated parameters for each dataset

Overall success for Dataset 1 but not Dataset 2

For Dataset 1, in line with the success of our FCNef for depression task, the best-fit model to predict the dependent variable of BDI Change (changes in general depressive symptoms) showed a significant main effect of rs-FC Change (see Supplementary Tables 30 and 31). Corresponding to this, a significant positive correlation between BDI Change and rs-FC Change was found (p_{FDR} = 0.002). See Figure 4a. Note that the same results were found regardless of whether the data from one extreme outlier was included/excluded (see Supplementary Results). Contrary to the results of Dataset 1, Dataset 2 results showed no evidence for the success of our FCNef for depression task. Specifically, the best-fit model to predict the dependent variable of BDI Change showed no significant main effect of rs-FC Change (see Supplementary Tables 30 and 32). Correspondingly, there was no overall correlation between BDI Change and rs-FC Change (p_{FDR} = 0.700). Interestingly, for Dataset 2, the best-fit model did not include any regressors for our manipulated parameters but this was almost significantly superseded by two models which did include such regressors ($\chi^2(3) = 7.37$, $p = 0.061$ for the first and $\chi^2(6) = 11.36$, $p = 0.081$ for the second; see Supplementary Table 30). When both of these models were further investigated they were found to show significant

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

interactions between rs-FC Change and Experimental Schedule (see Supplementary Table 33 and Supplementary Table 34). Follow-up BDI Change/rs-FC Change correlations were therefore examined for the two Experimental Schedule conditions separately. A non-significant correlation with a ***positive*** coefficient was found for participants in the consecutive condition (n = 23, r = 0.300, p_{FDR} = 0.170), whereas a non-significant correlation with a ***negative*** coefficient was found for participants in the non-consecutive condition (n = 24, r = -0.409, p_{FDR} = 0.977). This can be clearly seen in Figure 4d.

Evidence for specificity of the targeted functional connection in Dataset 1 but not Dataset 2

In Dataset 1, highlighting the potential of FCNef for precision medicine, evidence was found that normalization of the DLPFC-PCC rs-FC with FCNef training related to changes in Brooding but not Anxiety symptoms. This evidence supports the idea that the targeted FC is *specifically* related to maladaptive symptoms of rumination. Here, the best-fit model to predict the dependent variable of Brooding Change showed a significant main effect of rs-FC Change (see Supplementary Tables 35 and 36). Corresponding to this, a significant positive correlation between Brooding Change and rs-FC Change was found (p_{FDR} = 0.004). See Figure 4b. On the contrary, the best-fit model to predict the dependent variable of Anxiety Change did not show anything even close to a significant main effect of rs-FC Change (see Supplementary Tables 39 and 40). Corresponding to this, no significant positive correlation between Anxiety Change and rs-FC Change was found (p_{FDR} = 0.498). See Figure 4c. In Dataset 2, on the contrary, these promising results were not replicated. Specifically, no significant main effect of rs-FC Change was found in either the best-fit model with a dependent variable of Brooding Change or in the best-fit model with a dependent variable of Anxiety Change (see Supplementary Tables 35, 38, 39, and 41). See Figure 4e and 4f.

For each dataset, we directly tested whether there was a significant difference between coefficients from Brooding Change/rs-FC Change and Anxiety Change/rs-FC Change correlations (using MatLab code available freely online by Takeuchi, R. F., 2023, with one tail due to our strong directional hypothesis). The results indicated that, for Dataset 1 (z = 2.07, p = 0.019), but not for Dataset 2 (z = 0.17, p = 0.432), there was a meaningful difference between how rs-FC Change related to Brooding Change versus Anxiety Change.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The best-fit model to predict the dependent variable of Brooding Change for Dataset 1 was almost significantly superseded by a more complex model which included regressors for our manipulated parameters ($\chi^2(3) = 6.61$, $p = 0.085$; see Supplementary Table 35). This more complex model showed a significant interaction between Reward Change and rs-FC Change (see Supplementary Table 37). Exploratory follow-up testing revealed a significant correlation between Brooding Change and rs-FC Change in the high ($p_{FDR} = 0.002$) but not in the low reward condition ($p_{FDR} = 0.183$). Correspondingly, we found a significant difference between coefficients from the Brooding Change/rs-FC Change and Anxiety Change/rs-FC Change correlations for the high ($z = 2.076$, $p = 0.019$) but not the low reward condition ($z = 0.881$, $p = 0.189$). Note, the results here and in other relevant analyses did not change dependent on whether the data for one extreme Brooding Change outlier was included/excluded (see Supplementary Materials). Overall, the above results combined show that- given the right FCNef parameters (high reward and consecutive days of experimentation)- targeting a specific FC with FCNef (DLPFC-PCC FC) can lead to very precise decreases in only related symptoms (brooding rumination but not anxiety).

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

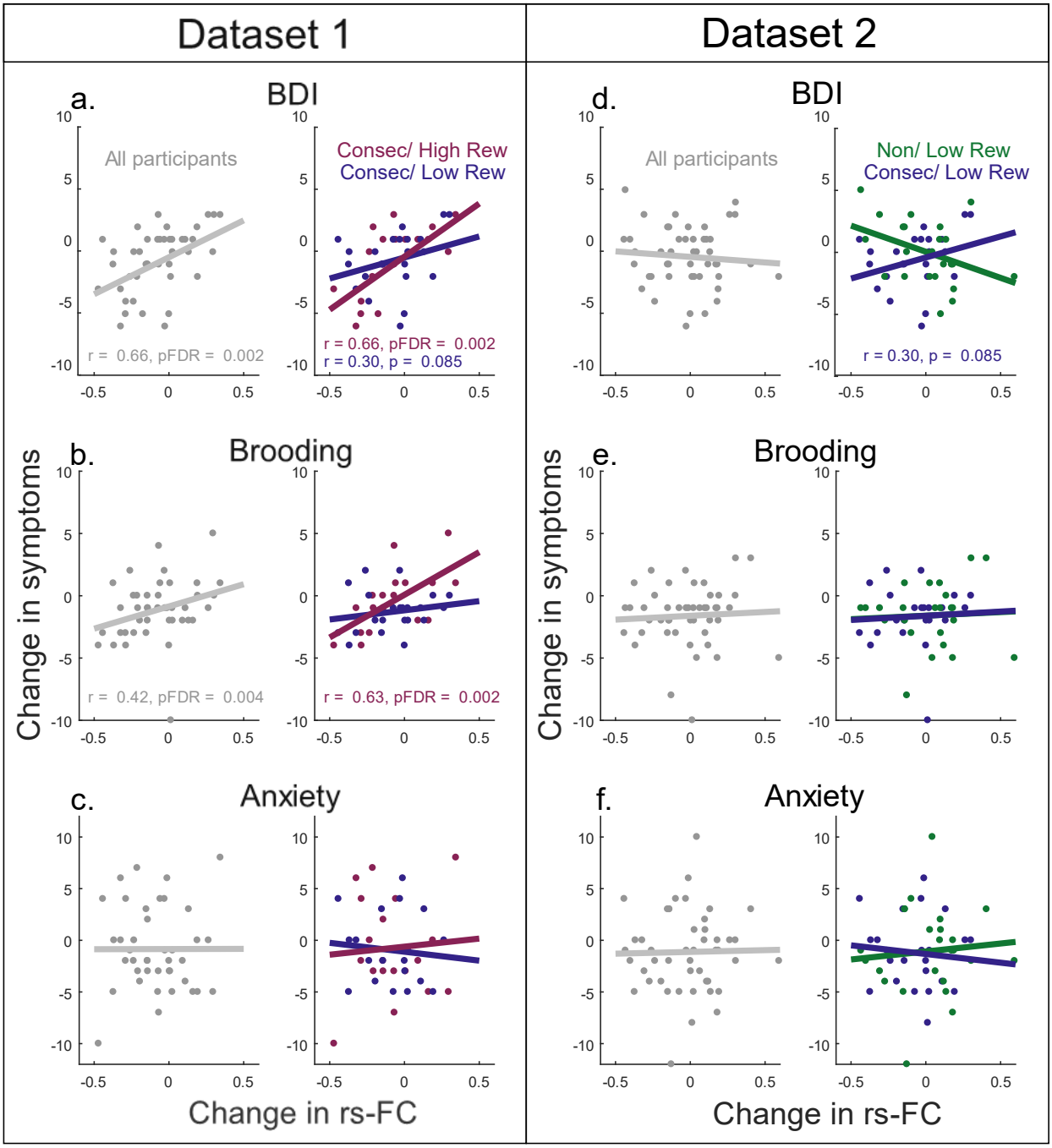


Figure 4. How changes in rs-FC relate to changes in symptomatology.

Changes in DLPFC-PCC rs-FC are shown plotted against changes in BDI scores (subplots a and d), Brooding scores (subplots b and e), and Anxiety scores (subplots c and f). On the left subplots of each dataset panel these relationships are shown for all participants and on the right subplots of each dataset panel these are shown split by parameter condition (see Table 1 for details). For illustrative purposes,

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

plotted relationships whose correlations reached significance are reported on the appropriate plots. Overall, promisingly for precision medicine, Dataset 1 showed significant positive correlations between changes in DLPFC-PCC rs-FC and changes in ***related*** *(BDI and Brooding) but not unrelated (Anxiety) symptoms. These effects were greatly enhanced under conditions of high reward (Consec/ High Rew group), but still in the same direction (reaching a trend with the uncorrected p-value for the BDI change correlation) under conditions of low reward (Consec/ Low Rew group). On the contrary, promising results were not found overall with Dataset 2, which was comprised of the Consec/ Low Rew group (results reported above) and the Non/ Low Rew group (for whom the relationship between DLPFC-PCC rs-FC changes and BDI score changes was actually numerically in the opposite direction to expectation).*

Expected Effect	Consec/ High Rew	Consec/ Low Rew	Non-Consec/ Low Rew
Increase in mean FCNef score	*d* = -0.58	*d* = -0.09	*d* = -0.31
Decrease in FCNef score STD	*d* = 0.72	*d* = 0.42	*d* = -0.19
Normalization of rs-FC	*d* = 0.47	*d* = 0.30	*d* = -0.11
Positive relationship between BDI changes and rs-FC changes	*r* = 0.66	*r* = 0.30	*r* = -0.41
Positive relationship between Brooding changes and rs-FC changes	*r* = 0.63	*r* = 0.21	*r* = 0.05

Large	Medium	Small	Trivial	Trivial	Small	Medium	Large

Hypothesized direction ⟷ Non-hypothesized direction

Table 2: Effect sizes summarized.

We collapsed the results across the different induction time-window conditions (see Table 1). "Expected effects" are effects that we would expect to see if our FCNef for depression training was successful. Comparisons for "Increase in mean FCNef scores" and "Decrease in FCNef score STD" were t-tests comparing data from the first and the last day of FCNef (Days 1 and 4). Comparisons for "Normalization of rs-FC" were t-tests comparing rs-FC from before to after FCNef (Days 0 and 4). The two "Positive relationship…" comparisons were correlations calculated between changes in rs-FC and changes in symptoms, where "changes" were defined as Day 4 - Day 0 data. Effect sizes for each comparison are shown in the relevant cell. These are shown here as Cohen's d (d) for t-tests and Pearson's correlation coefficient (r) for correlations. Following convention, these can be described on a scale ranging from trivial (d < 0.20 or r < 0.10), small (0.2 ≥ d < 0.5 or 0.1 ≥ r < 0.3), medium (0.5 ≥ d < 0.8 or 0.3 ≥ r < 0.5), to large (0.8 ≥ d or 0.5 ≥ r) (when considered here as absolute values). Effects in the hypothesized direction

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

are shown in shades of red and effects in the non-hypothesized direction are shown in shades of grey. The larger the effect size, the stronger the shade of the color in the relevant cell.

Discussion

Showing great promise for precision medicine, we found evidence that normalization of the DLPFC-PCC rs-FC specifically relates to reductions in symptoms related to this neural activity (those of brooding rumination) but not to reductions in symptoms unrelated to this neural activity (those of anxiety). Furthermore, with the goal of moving our FCNef technique onto clinical trials, we here tested the parameters under which it might optimally function to treat symptoms. Overall, we found that (1) FCNef works best using a liberal reward schedule as reinforcement for the targeted shift in functional connectivity, (2) FCNef may function best when participants come in for consecutive days of experimentation, and (3) the length of the induction period did not influence changes in rs-FC or corresponding reductions in symptoms, indicating that we could shorten this to minimize participant fatigue.

FCNef for precision medicine. Under the right parameters, the current results strengthen our previous finding that normalization of a target FC from before- to after FCNef relates to a specific reduction in only related symptoms. The current report shows these results with a far greater sample size than previously reported[18,11], which indicates that previous results were unlikely to have just been spurious. This greater sample size allowed the statistical power for direct testing of correlation coefficients, which confirmed that normalization of the DLPFC-PCC rs-FC was significantly more related to changes in related symptoms (brooding rumination) than to changes in unrelated symptoms (anxiety). Overall, although further testing is required, these results look very promising for the future use of FCNef for precision medicine. They bring us one step closer to a future where patients could one day enter the clinic, have their brains scanned, and then have targeted treatment to normalize their own specific neural aberrations related to their own specific subset of symptoms. Currently, these results are shown with functional connectivity in fMRI, and we plan to apply for approval of our local regulatory agency with this kind of paradigm. However, in the future, we expect this to evolve further so that FCNef may be conducted using EEG signatures (see Keynan et al., 2019[43]) of target FCs. This would allow it to eventually be conducted with portable EEG headsets, possibly even away from the

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

clinic in the privacy of the patient's own home (for more detailed discussion see Taylor et al., 2021[28]).

Parameter testing. The current report also extends our previous report by clarifying the parameters under which FCNef for depression can best be achieved. These results should help guide the design of future neurofeedback and other BMI studies. When selecting parameters for neurofeedback, past studies have tended to follow convention or have gone with what seems best in terms of cost/benefit trade-offs and/or in terms of making things easy for participants. Often, this is the only feasible way to design a BMI study because the cost of testing all possible parameters is enormous. However, our current results show that certain parameters can make a huge difference in BMI effectiveness. This means that without knowing the optimal parameters for a given BMI design, researchers may be finding null results simply because they are not running their designs in the best way they could (which would mean they are wasting time and money anyway). There is no simple solution to this complicated problem, especially because the optimal parameters could differ for different participant populations, target neural activity, goals, etc. Nonetheless, we hope that the current results provide some evidence that can be used to help future researchers be more informed when selecting their own parameters. The specific results for these parameter analyses, as well as their implications, are discussed in the following paragraphs.

Reward schedule. Participants in conditions with the high reward schedule had better FCNef success than those with the low reward schedule (see Table 2 and Figures 2a, 2b, 3a, 4a, and 4b). These results support the proposal that external reward might work as reinforcement that is additional to that provided by the feedback scores during neurofeedback[44]. Based on these results, we recommend using liberal external reward in future neurofeedback studies. Furthermore, because BMIs generally do not use external rewards for reinforcement, this result might be worthy of consideration beyond the realm of neurofeedback.

Of course, when running interventions for the reduction of depressive symptoms, as we did here, it is important to consider whether or not reward will actually be useful for the reinforcement of the target neural activity. Disturbances to reward circuitry and disturbances in reward processing (usually reductions) are commonly reported in depressive and other psychiatric disorders [45–52]. This means that the effect of external reward on the reinforcement of the target neural activity might be diminished when neurofeedback is conducted in patients with

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

such disorders. Here, our results with subclinical patients did not corroborate this, but it remains worthy of further investigation in clinically depressed patients.

Induction time-window. Analysis of both datasets showed no consistent or meaningful statistical evidence that indicated that induction time-window mattered for symptom reduction, rs-FC normalization, or the relationship between the two. However, there was some indication that induction time-window mattered for FCNef task performance. Specifically, within the high reward schedule condition, the average FCNef task scores increased significantly more from Day 1 to Day 4 with a longer compared to a shorter induction time-window (i.e. in C-40-H compared to C-20-H). However, this result was not mirrored in changes in FCNef task variance, as all other FCNef score results were. It may therefore have been spurious, especially given the small sample sizes when the data is broken down this far. Either way, this result did not matter for our operationalized measures of FCNef success (since there was no effect of induction time-window on symptom reduction, rs-FC normalization, and the relationship between the two). This could be because, even with lower overall FCNef scores, there was enough variance in daily FCNef scores for high reward schedule participants with the lower induction time-window (C-20-H) to experience sufficient reinforcement of the target FC. Overall, therefore, these results are promising because they indicate that we can use the shorter induction time-window to achieve our definition of success in our FCNef for depression paradigm. This will reduce the overall daily length of our paradigm by about 10 minutes and it will half the cognitive effort required for induction. We therefore hope it will make our paradigm less strenuous overall for participants. This is also a promising result because it means that other FCNef experiments run in the future may not need such long induction time-windows as have been conventionally used. This result highlights the importance of testing conventional parameters, when trying to make BMIs more palatable for the clinic.

Experimental schedule. FCNef appears to be more effective when participants come in for consecutive, as opposed to non-consecutive, days of FCNef. As seen in Table 2, all expected effects were the strongest in the consecutive condition with high reward, in the same direction (albeit with weaker significance) in the consecutive condition with low reward, and reduced or in the opposite direction for the non-consecutive condition. Of course, the non-consecutive condition that we tested was with low reward and it would have been nice to fully balance this by also testing the non-consecutive condition with high reward. Nonetheless, while these results are not fully conclusive, things do not look good for non-consecutive days of

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef. They indicate that, although possibly more tiring for participants, consecutive days of FCNef seem necessary to achieve positive outcomes. The results do not give any hints as to why this may be the case. However, it is possible that consecutive days of reinforcement are needed to drive learning effectively and/or that more non-controllable confounding personal circumstances can occur between non-consecutive days of training (an idea which researchers designing future BMIs and clinical treatments might benefit from considering).

One point to consider here is the possibility that neural plasticity related to the learning process might cause dynamic rs-FC changes in strength and direction that occur over time before settling into a new pattern (similar to the rebound effect first documented by Kluetsch et al., 2014[53]). If so, then our analyses may not have fairly tested between consecutive/non-consecutive conditions because they may not have been comparing the same snapshots of learning effects. Specifically, learning could begin from FCNef Day 1 and therefore post-FCNef measurements (from FCNef Day 4, and 1- and 2- months later), which were a different number of days later than FCNef Day 1 for consecutive/non-consecutive conditions (and even within the non-consecutive condition), might thus have been taken at different points in the ongoing dynamic rs-FC changes.

Limitations of the current design: One limitation of our study was that it was conducted in participants with only subclinical levels of depression. Nonetheless, preliminary studies using this FCNef technique with clinical patients with MDD have shown promise[18,31]. Results with clinical patients may be improved if the right parameters are employed. Another major limitation of the current study is that there was no control group or within-subject control condition. Therefore, it is possible that our target rs-FC changed and that symptoms improved for reasons such as the placebo or Hawthorne effects. These and other such possibilities are discussed in more detail in our previous paper[11]. Overall, however, we think that because our findings are specific to symptoms related only to the targeted FC (brooding rumination but not anxiety) this means that our results are unlikely to have completely risen from such effects. Finally, although we did not have a problem using reward to reinforce the target neural activity, this does not necessarily mean that the same might be found if FCNef was conducted to target a functional connection related to anhedonic symptoms of depression or in a subset of participants who have the anhedonic biotype (see [23,24]), for whom reward circuit disturbances are most commonly reported [54]. This issue needs yet to be tested. Regardless, the use of external reward in a clinical setting still needs to be further considered and perhaps explored

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

more creatively because there are obvious ethical problems related to rewarding patients with money.

Summary: Overall, here, in a larger dataset, we strengthen our previous results, showing a significant correlation between the normalization of a neural network and a reduction in related but not unrelated symptoms. We directly compared the related correlation coefficients for the first time. Because we found these coefficients to significantly differ, this provides evidence that FCNef might be appropriate for treatment of specific psychiatric symptoms. It therefore brings us one step closer to a future where psychiatric treatment might be tailored to the individual patient. Here, we additionally extended our previous work by investigating the parameters under which our FCNef for depression paradigm is most effective. We found that the effectiveness of FCNef changed greatly depending on the parameters with which it was run. The specifics and implications of some of the parameter-related results that we found might be relevant beyond neurofeedback to BMIs in general. Furthermore, some of the results we found highlight the benefits of testing conventional parameters. Overall, these results should be informative for the design of future BMI testing and for inspiring new interpretations of existing data. More broadly, by documenting how parameter optimization can increase beneficial outcomes and reduce patient burden, we hope to inspire more of this in the future, with the ultimate goal of bringing optimized BMIs to the medical clinic.

Data accessibility

The data and code supporting this study's findings will be publicly available on our GitHub at publication.

Funding

This work was supported by the contract research "Brain/MINDS Beyond" (Grant Number JP18dm0307008, JP21dm0307102h0003), from the Japan Agency for Medical Research and Development (AMED). It was also supported by “Application of DecNef for Development of Diagnostic and Cure System for Mental Disorders and Construction of Clinical Application Bases” under “Development of BMI Technologies for Clinical Application” of the Strategic Research Program for Brain Sciences, (Grant Number JP17dm0107044; AMED), and Innovative Science and Technology Initiative for Security, ATLA, Japan (Grant Number JPJ004596).

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

Competing interests

MK is an inventor of patents related to the functional connectivity neurofeedback method. The original assignee of the patents is ATR, with which the authors are affiliated. We have no other conflicts of interest.

Author Contributions

JET, TM, TY and MK designed the experiments; JET, TO, MM and TM acquired the data; JET analyzed the data; MK and AC supervised the data analysis; JET prepared the original draft; TO, MM, MT, TY, TKa, TKo, YY, YM, JM, TK, and AC reviewed and edited the manuscript. All authors gave final approval for submission and agreed to take responsibility for the manuscript.

Acknowledgements

We thank Kaori Nakamura for her help in scheduling and conducting the experiments; Rumi Yorizawa for her help in screening participants; Toshinori Yoshioka for his continued support with regards to the experimental scripts.

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

FCNef for Precision Medicine.

References

1. Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression.
2. Patel, V. *et al.* Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. *Lancet* **387**, 1672–1685 (2016).
3. Fava, M. & Davidson, K. G. Definition and epidemiology of treatment-resistant depression. *Psychiatr. Clin. North Am.* **19**, 179–200 (1996).
4. Herrman, H. *et al.* Time for united action on depression: a Lancet–World Psychiatric Association Commission. *Lancet* **399**, 957–1022 (2022).
5. Gabriel, F. C. *et al.* Pharmacological treatment of depression: A systematic review comparing clinical practice guideline recommendations. *PLoS One* **15**, e0231700 (2020).
6. The tech: Restore. *Synchron* https://synchron.com/technology/brain-io (2021).
7. Neuralink announces first-in-human clinical trial of BCI device. https://www.fdanews.com/articles/211989-neuralink-announces-first-in-human-clinical-trial-of-bci-device?v=preview.
8. Paradromics. Paradromics raises $33 million in funding, achieves breakthrough medical device designation from FDA. *PR Newswire* https://www.prnewswire.com/news-releases/paradromics-raises-33-million-in-funding-achieves-breakthrough-medical-device-designation-from-fda-301827969.html?tc=eml_cleartime (2023).
9. Neurolutions Receives FDA De Novo Market Authorization for IpsiHand™. *Neurolutions* https://www.neurolutions.com/post/neurolutions-receives-u-s-food-and-drug-administration-de-novo-market-authorization-for-ipsihand (2021).
10. Rustamov, N., Souders, L., Sheehan, L., Carter, A. & Leuthardt, E. C. IpsiHand Brain-Computer Interface Therapy Induces Broad Upper Extremity Motor Recovery in Chronic

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Stroke. *medRxiv* (2023) doi:10.1101/2023.08.26.23294320.

11. Taylor, J. E. *et al.* Depressive symptoms reduce when dorsolateral prefrontal cortex-precuneus connectivity normalizes after functional connectivity neurofeedback. *Sci. Rep.* **12**, 2581 (2022).
12. Koush, Y. *et al.* Connectivity-based neurofeedback: dynamic causal modeling for real-time fMRI. *Neuroimage* **81**, 422–430 (2013).
13. Koush, Y. *et al.* Learning Control Over Emotion Networks Through Connectivity-Based Neurofeedback. *Cereb. Cortex* **27**, 1193–1202 (2017).
14. Megumi, F., Yamashita, A., Kawato, M. & Imamizu, H. Functional MRI neurofeedback training on connectivity between two regions induces long-lasting changes in intrinsic functional network. *Front. Hum. Neurosci.* **9**, 160 (2015).
15. Ramot, M. *et al.* Direct modulation of aberrant brain network connectivity through real-time NeuroFeedback. *Elife* **6**, (2017).
16. Tsuchiyagaito, A. *et al.* Prevent breaking bad: A proof of concept study of rebalancing the brain's rumination circuit with real-time fMRI functional connectivity neurofeedback. *Hum. Brain Mapp.* **42**, 922–940 (2021).
17. Tsuchiyagaito, A. *et al.* Real-Time fMRI Functional Connectivity Neurofeedback Reducing Repetitive Negative Thinking in Depression: A Double-Blind, Randomized, Sham-Controlled Proof-of-Concept Trial. *Psychother. Psychosom.* **92**, 87–100 (2023).
18. Yamada, T. *et al.* Resting-State Functional Connectivity-Based Biomarkers and Functional MRI-Based Neurofeedback for Psychiatric Disorders: A Challenge for Developing Theranostic Biomarkers. *Int. J. Neuropsychopharmacol.* **20**, 769–781 (2017).
19. Drysdale, A. T. *et al.* Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* **23**, 28–38 (2017).

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

20. Roelfs, D. *et al.* Genetic overlap between multivariate measures of human functional brain connectivity and psychiatric disorders. *Nature Mental Health* 1–11 (2024).
21. Tokuda, T. *et al.* Identification of depression subtypes and relevant brain regions using a data-driven approach. *Sci. Rep.* **8**, 14082 (2018).
22. van den Heuvel, M. P., Scholtens, L. H. & Kahn, R. S. Multiscale Neuroscience of Psychiatric Disorders. *Biol. Psychiatry* **86**, 512–522 (2019).
23. Williams, L. M. Precision psychiatry: a neural circuit taxonomy for depression and anxiety. *Lancet Psychiatry* **3**, 472–480 (2016).
24. Williams, L. M. Defining biotypes for depression and anxiety based on large-scale circuit dysfunction: a theoretical review of the evidence and future directions for clinical translation. *Depress. Anxiety* **34**, 9–24 (2017).
25. Woodward, N. D. & Cascio, C. J. Resting-State Functional Connectivity in Psychiatric Disorders. *JAMA psychiatry* vol. 72 743–744 (2015).
26. Yahata, N., Kasai, K. & Kawato, M. Computational neuroscience approach to biomarkers and treatments for mental disorders. *Psychiatry Clin. Neurosci.* **71**, 215–237 (2017).
27. Zhao, Z. *et al.* Real-time functional connectivity-based neurofeedback of amygdala-frontal pathways reduces anxiety. *bioRxiv* 308924 (2018) doi:10.1101/308924.
28. Jessica Elizabeth Taylor, Itamar Jalon, Toshinori Chiba, Tomokazu Motegi, Mitsuo Kawato, and Talma Hendler. Chapter 12: Translation to the clinic and other modalities. in *fMRI Neurofeedback* (ed. Hampson, M.) 265–286 (Academic Press, Elsevier, 2021).
29. Ichikawa, N. *et al.* Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. *Sci. Rep.* **10**, 3542 (2020).
30. Treynor, W., Gonzalez, R. & Nolen-Hoeksema, S. Rumination Reconsidered: A Psychometric Analysis. *Cognit. Ther. Res.* **27**, 247–259 (2003).

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

31. Takamura, M. *et al.* Application of functional connectivity neurofeedback in patients with treatment-resistant depression: A preliminary report. *Journal of Affective Disorders Reports* **14**, 100644 (2023).
32. deCharms, R. C. *et al.* Learned regulation of spatially localized brain activation using real-time fMRI. *Neuroimage* **21**, 436–443 (2004).
33. deCharms, R. C. *et al.* Control over brain activation and pain learned by using real-time functional MRI. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 18626–18631 (2005).
34. Scharnowski, F. *et al.* Connectivity changes underlying neurofeedback training of visual cortex activity. *PLoS One* **9**, e91090 (2014).
35. Sepulveda, P. *et al.* How feedback, motor imagery, and reward influence brain self-regulation using real-time fMRI. *Hum. Brain Mapp.* **37**, 3153–3171 (2016).
36. Yamashita, A., Hayasaka, S., Kawato, M. & Imamizu, H. Connectivity Neurofeedback Training Can Differentially Change Functional Connectivity and Cognitive Performance. *Cereb. Cortex* **27**, 4960–4970 (2017).
37. Beck, A.T., Steer, R.A., & Brown, G.K. *Beck Depression Inventory (BDI-II): Manual and Questionnaire*. (1996).
38. Hasegawa, A. Translation and initial validation of the Japanese version of the Ruminative Responses Scale. *Psychol. Rep.* **112**, 716–726 (2013).
39. Spielberger, C. D., Gorsuch, R. L., Lushene, R. & Vagg, P. R. Manual for the State-Trait Anxiety Inventory; Palo Alto, CA, Ed. *Palo Alto: Spielberger* (1983).
40. Kirchner, W. K. Age differences in short-term retention of rapidly changing information. *J. Exp. Psychol.* **55**, 352–358 (1958).
41. Akaike, H. Information theory and an extension of the maximum likelihood principle. in *2nd international symposium on information theory.* (ed. Csáki, B. N. P. &.) 267–281 (Akadémia

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Kiadó., 1973).

42. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
43. Keynan, J. N. *et al.* Electrical fingerprint of the amygdala guides neurofeedback training for stress resilience. *Nat Hum Behav* **3**, 63–73 (2019).
44. Watanabe, T., Sasaki, Y., Shibata, K. & Kawato, M. Advances in fMRI Real-Time Neurofeedback. *Trends Cogn. Sci.* **21**, 997–1010 (2017).
45. Nestler, E. J. & Carlezon, W. A., Jr. The mesolimbic dopamine reward circuit in depression. *Biol. Psychiatry* **59**, 1151–1159 (2006).
46. Russo, S. J. & Nestler, E. J. The brain reward circuitry in mood disorders. *Nat. Rev. Neurosci.* **14**, 609–625 (2013).
47. Zhang, W.-N., Chang, S.-H., Guo, L.-Y., Zhang, K.-L. & Wang, J. The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *J. Affect. Disord.* **151**, 531–539 (2013).
48. Baskin-Sommers, A. R. & Foti, D. Abnormal reward functioning across substance use disorders and major depressive disorder: Considering reward as a transdiagnostic mechanism. *Int. J. Psychophysiol.* **98**, 227–239 (2015).
49. Whitton, A. E., Treadway, M. T. & Pizzagalli, D. A. Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Curr. Opin. Psychiatry* **28**, 7–12 (2015).
50. Keren, H. *et al.* Reward Processing in Depression: A Conceptual and Meta-Analytic Review Across fMRI and EEG Studies. *Am. J. Psychiatry* **175**, 1111–1120 (2018).
51. Knowland, D. & Lim, B. K. Circuit-based frameworks of depressive behaviors: The role of reward circuitry and beyond. *Pharmacol. Biochem. Behav.* **174**, 42–52 (2018).
52. Fox, M. E. & Lobo, M. K. The molecular and cellular mechanisms of depression: a focus on

medRxiv preprint doi: https://doi.org/10.1101/2024.04.03.24304187; this version posted April 3, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

reward circuitry. *Mol. Psychiatry* **24**, 1798–1815 (2019).

53. Kluetsch, R. C. *et al.* Plastic modulation of PTSD resting-state networks and subjective wellbeing by EEG neurofeedback. *Acta Psychiatr. Scand.* **130**, 123–136 (2014).

54. Borsini, A., Wallis, A. S. J., Zunszain, P., Pariante, C. M. & Kempton, M. J. Characterizing anhedonia: A systematic review of neuroimaging across the subtypes of reward processing deficits in depression. *Cogn. Affect. Behav. Neurosci.* **20**, 816–841 (2020).