# EEG-based neurofeedback with network components extraction: data-driven approach by multilayer ICA extension and simultaneous EEG-fMRI measurements

Takeshi Ogawa[1*], Hiroki Moriya[1], Nobuo Hiroe[2], Motoaki Kawanabe[1,3], Jun-ichiro Hirayama[3,4]

[1]Cognitive Mechanisms Laboratories, Advanced Telecommunications Research Institute International, Seika-cho, Soraku-gun, Kyoto, Japan

[2]Neural Information Analysis Laboratories, Advanced Telecommunications Research Institute International, Seika-cho, Soraku-gun, Kyoto, Japan

[3]RIKEN Center for Advanced Intelligence Project, Seika-cho, Soraku-gun, Kyoto, Japan

[4]Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba City, Ibaraki, Japan

*Correspondence:

Takeshi Ogawa, Ph.D

t.ogawa@atr.jp

# Abstract

Advanced treatments for depressive symptoms, such as a real-time functional MRI (fMRI) neurofeedback have been proven by several studies. In particular, regularization of functional connectivity (FC) between executive control network (ECN) and default mode network (DMN) during fMRI neurofeedback have been proposed to reduce depressive symptoms. However, it is difficult to install this system anywhere and repetitively provide the treatment in practice, because the cost is high and no practical signal processing techniques have existed so far to extract FC-related features from EEG, particularly when no physical forward models are available. In this regard, stacked pooling and linear components estimation (SPLICE; Hirayama et al., 2017), recently proposed as a multilayer extension of independent component analysis (ICA) and related independent subspace analysis (ISA; Hyvärinen & Hoyer, 2000), can be a promising alternative, although its application to the resting-state EEG have never been investigated so far. We expected that if the extracted EEG network features were correlated with fMRI network activity corresponding to DMN or ECN, it may help to modulate the target FC in the EEG-based neurofeedback training.

Here, we describe a real-time EEG neurofeedback paradigm for improvement of depressive symptoms by using an EEG network features estimated by SPLICE. We hypothesized upregulation of the dorsolateral prefrontal cortex (DLPFC)/middle frontal cortex or downregulation of precuneus/posterior cingulate cortex (PCC) related to fMRI biomarker for depression (Ichikawa et al., Sci Rep, 2020) should specifically predict decreases in depressive symptoms during the neurofeedback training. We conducted a single-blind design for neurofeedback group (n=8; NF group) and sham group (n=9) groups for three days. To this end, we found large effect size in the rumination response

scale score (total, brooding and reflection) in the comparison between NF and sham groups. Additionally, brain signals of the tasked fMRI (a contrast of 2-back > 0-back) in the neurofeedback group were significantly decreased in the right cuneus and DLPFC compared to the control group. We demonstrated a feasibility of EEG neurofeedback treatment for depressive symptoms using EEG network features extracted by SPLICE in the subclinical trials.

**Keywords: neurofeedback, depression, EEG-fMRI, blind source separation, functional connectivity**

# 1. Introduction

Advanced treatments for depressive symptoms have been studied using brain or physiological measurement instead of traditional treatments of medicine. One of the noninvasive ways to accurately monitor brain activity that may represent depressive symptoms is functional MRI (fMRI) because of its high spatial resolution even around deep brain areas. Several groups have proposed real-time fMRI neurofeedback paradigms that can used to train the brain of people with depressive symptoms to function more like those of healthy individuals (Young et al., 2017; Yamada et al., 2017; Taylor et al., 2021). During the intervention of the real-time neurofeedback, participants can control their own brain activity to be closed to the healthy condition, and then reduce their depressive symptoms.

The effectiveness of real-time fMRI neurofeedback for depressive symptoms has been proven by several studies. However, it is difficult to install this system anywhere and repetitively provide the treatment in practice because of the expensive costs. Instead of fMRI, therefore, alternative approaches to the real-time neurofeedback have been proposed using electroencephalography (EEG), such as those based on frontal alpha asymmetry (Wang et al., 2019), peak alpha frequency (Yu et al., 2020), and alpha-theta ratio (Cheon et al., 2016). The frontal alpha asymmetry EEG neurofeedback has been examined to modulate left-side alpha power associated with brain activity in the left amygdala. EEG has many practical benefits such as reasonable cost, high-temporal resolution, and less physical constraints while it has a detriment of low-spatial resolution. Several ‘proof of concept’ studies have already shown the potential effectiveness of real-time EEG-fMRI neurofeedback paradigms for depression (Zotev et al., 2014; Zotev et al.,

2016; Zotev et al., 2020) and improved the reliability of EEG signature for neurofeedback. In these types of neurofeedback training (NF training), participants are required to imagine their autographic memory of happy events and to increase brain activity in the left amygdala, and frontal alpha asymmetry at the same time. Thus, the purpose of these neurofeedback is to control the left amygdala and related brain regions similar to the healthy state.

Recently, a new type of approach has also been developed for fMRI neurofeedback of depression based on regularization of functional connectivity (FC) between executive control network (ECN) and default mode network (DMN) (Yamada et al., 2017; Taylor et al., 2021; Tsuchiyagaito et al., 2020). In healthy individuals, the time courses of ECN and DMN are often observed to be anticorrelated, and it is associated with the distinction between internal (e.g., own thoughts) and external (e.g., perceptions of external environment) contents. In contrast to the health individuals, the case of depressive disorder tends to dysbalanced, and then it increases self-focus (rumination, guilt, fear) and decreases environment-focus (social withdrawal, psychomotor retardation) (Northoff, 2016; Drysdale et al., 2017). This fact is clarified in the fMRI biomarker, which was identified melancholic depression with data-driven approaches (Ichikawa et al., 2020; Yamashita et al., 2020). In particular, FC between the dorsolateral prefrontal cortex (DLPFC)/middle frontal gyrus (mFG) in ECN and the precuneus/posterior cingulate cortex (PCC) in DMN was shown to be necessary. Given an evidence of fMRI biomarker, this type of technique, called FC neurofeedback (FCNef), with a precise localization of the target connection, has shown a great promise for treatments of depression.

However, in contrast to the real-time fMRI feedback or FCNef, an EEG neurofeedback using network features have been poorly investigated. This is mainly because no practical signal processing techniques have existed so far to extract FC-related features from EEG, particularly when no physical forward models are available. In fact, most studies have used sensor-level features such as time-instant powers of particular frequency bands or sought to extract source-level features by a blind-source separation (BSS) technique such as independent component analysis (ICA). Neither method is not designed to detect EEG features that reflect FC and modular organization related to the resting-state networks. To extract EEG network features related to the FC observed in fMRI, a new BSS-based technique is strongly required, which should be combined with a signal model that explicitly assumes a network structures at the source level. In this regard, stacked pooling and linear components estimation (SPLICE; Hirayama et al., 2017), recently proposed as a multilayer extension of ICA and related independent subspace analysis (ISA; Hyvärinen & Hoyer, 2000), can be a promising alternative, although its application to the resting-state EEG have never been investigated so far. We expected that if the extracted EEG network features were correlated with fMRI network activity corresponding to DMN or ECN, it may help to modulate the target FC in the EEG-based NF training.

Here, we describe a real-time EEG neurofeedback paradigm for improvement of depressive symptoms by using an EEG network features estimated by SPLICE. We hypothesized upregulation of the DLPFC/mFC or downregulation of precuneus/PCC should specifically predict decreases in depressive symptoms during the NF training. To investigate this, we ran subclinical participants with depressive symptoms in an EEG-

126 neurofeedback paradigm for three days targeting regions (DLPFC/mFC or
127 precuneus/PCC) in a single-blind design while referring a FCNef paradigm (Yamada et
128 al., 2017; Taylor et al., 2021). For training of the SPLICE model, we conducted two-days
129 EEG-fMRI experiments for each participant to simultaneously acquire EEG-fMRI data at
130 resting and during N-back task. After modeling of SPLICE from the resting-state data,
131 we selected EEG network features which was correlated with fMRI signal in the DLPFC
132 or precuneus/PCC for NF training. Scores on the Beck Depression Inventory-II (BDI-II;
133 Beck et al., 1996) and the Rumination Response scale (RRS; Hasegawa, 2013) were
134 measured before and after the NF training. To this end, we evaluated effect size of
135 difference between pre- and post-training among neurofeedback and sham groups, then
136 described a feasibility of EEG-neurofeedback using EEG network features relevant to
137 fMRI network information.
138

# 139 2 Materials and Methods

## 140 2.1 SPLICE

141 SPLICE (Hirayama et al., 2017) is a multilayer extension of ICA based on a probabilistic
142 model that simplifies the generative process of EEG. Let $\mathbf{x}$ and $\mathbf{s}$ denote the vectors of
143 EEG measurements and unobserved source signals, respectively, at a single time point.
144 Then, we assume a linear model (first layer) similar to ICA, given by

145 $$\mathbf{x} = \mathbf{A}\mathbf{s}, \qquad (1)$$

146 where mixing matrix $\mathbf{A}$ is square and invertible; the inverse $\mathbf{W}: = \mathbf{A}^{-1}$ is called demixing
147 matrix. Both $\mathbf{x}$ and $\mathbf{s}$ are zero-mean, without loss of generality, by subtracting the sample
148 mean in advance from original data vectors. Note that the sources (i.e., entries in $\mathbf{s}$) lacks
149 explicit physical mapping to the cortex, but they can still be interpretable as reflecting

cortical activations based on their associated topographies (i.e., corresponding columns of $\mathbf{A}$), which is exactly the same case as ICA. For mathematical convenience, the number of sources, $d$, is assumed to be the same as the dimensionality of $\mathbf{x}$; in practice, we apply PCA beforehand to determine an appropriate number of sources as customary done in ICA.

The additional layers, on top of the first linear layer, are introduced to model statistical dependency among the sources and parametric structures behind it. Specifically, SPLICE introduces 1) functional modularity of sources and 2) intrinsic correlations (coactivations) among modules, into the multilayer generative model. The both are reasonable assumptions in particular for the resting-state EEG data, because cortical activity at rest is known to exhibit functional organization with multiple modules (e.g., resting-state network), often mutually correlated or anticorrelated.

Formally, we divide the $d$ first-layer sources into $m$ modules $\mathbf{s}_{[j]}$ without overlapping, where $\mathbf{s}_{[j]}$ denotes the vector consisting of $d_j$ sources belonging to module $j$ ($d = \sum_{j=1}^{m} d_j$). $\mathbf{s}_{[j]}$ represents a $d_j$-dimensional subspace spanned by the corresponding columns in $\mathbf{A}$. Then, we assume that modules $\mathbf{s}_{[j]}$ have mutually correlated (squared) L2-norms $\|\mathbf{s}_{[j]}\|^2$, generated by an additional ICA-like linear model with pointwise nonlinearity, such that

$$\|\mathbf{s}_{[j]}\|^2 = \boldsymbol{F}_j^{-1}([\mathbf{A'}\mathbf{s'}]_j), j = 1, 2, \ldots, m, \quad (2)$$

Where an appropriate link function $\boldsymbol{F}_j$ is used for convention between nonnegative (norm) and real random variables (we set $\boldsymbol{F}_j$ = log in our analysis); $\mathbf{A'}$ and $\mathbf{s'}$ are invertible mixing matrix and source vector (and $\mathbf{W'} := \mathbf{A'}^{-1}$ is demixing matrix) of this layer, and $[\cdot]_j$ denotes $j$-th entry of a vector. Given the norms, every $\mathbf{s}_{[j]}/\|\mathbf{s}_{[j]}\|$ is assumed to be uniformly

distributed on a unit hypersphere, independently of any other random variables, which makes the model non-informative except for the distribution of the norms. Note that the model can actually be recursively extended to have additional layers to further model the dependency in higher-layer sources (Hirayama et al., 2017). In the present study, however, we simply assume Eq.(2) to be the top layer. The entries in $\mathbf{s}$' are thus assumed to be independent of each other, given a typical prior in ICA (e.g., the one corresponding to the standard *tanh* nonlinearity).

To recapitulate the model, we illustrate a conceptual diagram of SPLICE in Figure 1. The model consists of three layers, two linear layers and an intermediate L2-pooling layer calculating the norms. Here, we define $\mathbf{x}' := \mathbf{A}'\mathbf{s}'$. If the highest layer $\mathbf{A}'$ is not diagonal, modules are mutually independent and the model in fact reduces to that of independent subspace analysis (Hyvärinen & Hoyer, 2000). If not diagonal, the modules are not independent and thus expected to be suitable to represent correlated or anticorrelated cortical modules. Note that modeling non-diagonal $\mathbf{A}'$ is not only affect the higher layer but also influences the first layer through the joint learning of the two layers (see below), because it makes a different (possibly improved) prior assumption on the first-layer sources from that of ICA or ISA.

A striking property of SPLICE is that both parameter estimation (learning) and posterior inference on latent variables can be performed in a principled manner, without resorting to any approximative techniques or heavy numerical computation. Learning is done based on conventional maximum likelihood (ML) estimation. The corresponding loss function, i.e., negative log-likelihood for a single datum $\mathrm{L} := -\ln p(\mathbf{x}) + \mathrm{const.}$, is analytically

given by

$$L = \sum_{k=1}^{m} \boldsymbol{H}_k \left( \sum_{k=1}^{m} w'_{kj} \boldsymbol{F}_j \left( \left\| \mathbf{W}_{[j]} \mathbf{x} \right\|^2 \right) \right) - \ln|det \mathbf{W}'| + \sum_{k=1}^{m} \boldsymbol{G}_j \left( \left\| \mathbf{W}_{[j]} \mathbf{x} \right\|^2 \right) - c \ln|det \mathbf{W}| \quad (3)$$

where $\boldsymbol{H}$ and $\boldsymbol{G}$ are functions derived from the prior distribution on $\mathbf{s}$' and link function $\boldsymbol{F}_j$ and the dimensionality of every module (subspace) $d_j$ is adaptively determined before the ML estimation using a clustering-like technique (see Hirayama et al., 2017 for details).

Once demixing matrices $\mathbf{W}$ and $\mathbf{W}$' are learned, one can readily invert the generative process by computing $\mathbf{s} = \mathbf{W}\mathbf{x}$ and $\mathbf{s}' = \mathbf{W}'\mathbf{x}'$, with the L2-pooling $\|\mathbf{s}_{[j]}\|^2$ (followed by $\boldsymbol{F}_j$) between them. In practice, the pooled outputs (squared norm) may fluctuate too heavily due to the noisy nature of EEG. Thus, as a slight extension of the original SPLICE, we also implemented the L2-pooling across successive time points for smoothing, which was done in a principled manner by modifying the model and likelihood, so that the top linear layer determines the L2 norm of $\mathbf{s}_{[j]}$ that now concatenates several instances of the $d_j$ sources across time points. Note that here, the pooling just means taking the sum of squares (of the filtered outputs). Thus, the squared module norm essentially gives the total power or energy of a source module's activity. We therefore focused primarily on the estimated module norms as the main output of the analysis, rather than the top-layer sources $\mathbf{s}$' which the multilayer computation eventually produces, because we expected to associate EEG features with the activity of cortical modules.

(Figure 1 is inserted around here)

## 2.2 Experiment procedures

### 2.2.1 Participants

The flow of this experiment was summarized in Figure 2. All participants were recruited by screening questionnaires and clinician assessment (n = 102: 49 females, 53 males; mean age = 29.88 ± 9.8 years old). To meet criteria for participation in the EEG-fMRI simultaneous recording experiment, participants must have (a) no inclination of suicidal thoughts, as measured by a question on the BDI, (b) no current or recent mental or psychiatric diseases, (c) understanding of the Japanese language. All received cash remuneration for their involvement after the experiment. Twenty eight participants without psychiatric disease in the past could proceed to the EEG-fMRI experiment for two days. We excluded participants who caused body movement in the MRI scanner. Finally, seventeen participants completed the NF training experiment for four days. All participants provided written informed consent prior to the experiment. This study was approved by the ethical committee of the Advanced Telecommunications Research Institute International (ATR) and followed the Declaration of Helsinki.

(Figure 2 is inserted around here)

### 2.2.2 Psychological questionnaires

In the Screening experiment, all participants completed self-report questionnaires that included demographic characteristics, the Beck Depression Inventory II (BDI-II), Rumination Response Scale (RRS), Obsessive-Compulsive Inventory (OCI), Autism-Spectrum Quotient (AQ), State-Trait Anxiety Inventory (STAI2), Snaith-Hamilton Pleasure Scale (SHAPS), Schizotypal Personality Questionnaire (SPQ), Barratt

244 Impulsiveness Scale 11 (BIS-11), Zimbardo Time Perspective Inventory (ZTPI), and were
245 rated on the 21-item Hamilton Depression Rating Scale (HAM-D), the Hamilton Anxiety
246 Rating Scale (HAM-A), and Structured Clinical Interview for DSM-IV-TR Axis disorders
247 (SCID). To evaluate the effectiveness of the NF training, participants completed BDI-II,
248 RRS, OCI, AQ, STAI2, BIS-11, and ZTPI on Day 1 (Visit 4) and Day 4 (Visit 7) in the NF
249 training experiment.
250
251 **2.2.3 N-back task**
252 Participants had performed N-back tasks in the screening (Visit 1) and the NF training
253 experiment (Day 1, Day 4) outside the MRI scanner as behavioral experiments. The N-
254 back task was controlled by using Presentation ® software (Version 18.0,
255 Neurobehavioral Systems, Inc., Berkeley, CA) and can be viewed at a display or a screen.
256 At the beginning of each session, participants had started a N-back task by pressing the
257 response button when they were ready. They had been instructed to simply focus on this
258 and relax. In each session, subsequent to the rest period for 30 s, there were four blocks
259 of the N-back task, with the task 'rule' changing from block-to-block (order randomized
260 within and between sessions). At the beginning of each block, written instructions were
261 first presented on display to inform participants of the current 'rule' and this was followed
262 by 64 trials in which this 'rule' should be applied. On each trial, a fixation for (1300 ms)
263 and then a number between 1-9, for 200 ms was presented centrally on display. In the
264 '0-back' block, the rule was to press the response button on the current trial if the number
265 that appeared on screen was '0'. In the '1-back', '2-back', '3-back' blocks, the rule was
266 to press the response button on the current trial if the number that appeared on display
267 was the same as the number that had been presented on display one, two, or three trials

beforehand, respectively. The participants completed two sessions of this task about 13 min.

In the EEG-fMRI experiment, they also conducted the N-back task with EEG recording only and EEG-fMRI simultaneous recording. At the beginning of each session, they first had a rest-period, where they saw a black fixation cross which was presented on the screen for 30 s. In each session, subsequent to the rest period, there were six blocks of the N-back task (0-, 1-, and 2-back) with the task 'rule' changing from block-to-block (order randomized within and between sessions). At the beginning of each block, written instructions were first presented on display to inform participants of the current 'rule' and this was followed by 20 trials in which this 'rule' should be applied as well as the task for the behavioral experiment. A "rest" period (identical to the one at the beginning of the session) was inserted halfway through (between block 3 and block 4 of) each session. Participants' task was to follow the current 'rule' to make as many correct responses as possible.

## 2.3 Data acquisitions

For a general schematic of the experimental procedure, see Figure 2.

MRI data were required with a 3T Siemens scanner, MAGNETOM Verio (Siemens, Erlangen, Germany) with a Siemens 12-channel head coil. High-resolution, T1-weighted structural images (TR = 2250 ms, TE = 3.06 ms, flip angle = 9 deg, inversion time = 900 ms, matrix = 256 x 256, 208 sagittal slices, 1 mm isotropic) were acquired for normalization to a standard brain for echo planar image (EPI) registration purpose. Images of BOLD signal such as fMRI data, were acquired with an EPI sequence (TR =

292 2450 ms, TE = 30 ms, flip angle = 80 deg, matrix = 64 × 64, field of view = 192 mm, slice
293 thickness = 3.2 mm, 35 axial slices, scan sequence: ascending). The durations of a
294 single session were 6 min 7 s (150 volumes) for the resting state, 5 min 45 s (141
295 volumes) for the N-back task.
296
297 Scalp EEG signals were recorded with an MR-compatible amplifier (BrainAMP MR plus,
298 Brain Products GmbH, Germany) and EEG electrode cap (BrainCap MR, Brain Products
299 GmbH, Germany) providing 63 EEG channels and 1 electrocardiogram (ECG) channel.
300 The EEG electrodes were placed according to the modified International 10-10 system.
301 The ground electrode and on-line reference electrode were placed on AFz and at FCz,
302 respectively. The impedance of all electrodes was kept lower than 10 kΩ throughout the
303 experiment. The ECG electrode was placed on the back of participants to obtain the
304 electrocardiographic data and subsequently correct ballistocardiographic artifacts. Raw
305 recording data was sampled at 5 kHz with a bandpass filter between 0.1 and 250 Hz
306 using the Brain Vision Recorder (Brain Products GmbH, Germany). The amplifier system
307 was set beside the subject's head within the scanner during fMRI scanning. To achieve
308 phase synchronization clocks for digital sampling between the MRI data and the EEG
309 system, the EEG system clock was synchronized with a SyncBox device (Brain Products
310 GmbH, Germany) and the MRI scanner's 10 MHz master synthesizer. The scanner also
311 delivered each TR trigger signal that marked the onset time of every fMRI volume
312 acquisition. These markers were used for fMRI scanning artifact correction of the EEG
313 data.
314
315 In the EEG-fMRI experiment, EEG signals were taken in the soundproof shield room,

316 and then EEG signals and MRI data were simultaneously taken in the MRI scanner for
317 two days. First, each participant completed EEG measurement of the resting-state with
318 two sessions and the N-back task (0-, 1-, 2-back) with two sessions in the soundproof
319 shield room. A single session of the resting-state EEG took five minutes, and a single
320 session of the N-back task took about six minutes. Second, participants moved to the
321 MRI scanner and completed the resting-state (two sessions), the N-back task (two
322 sessions), T1-weighted structural image scan, and the resting state (two sessions).
323 Finally, four sessions of EEG data for the resting-state and the N-back task, eight
324 sessions of EEG-fMRI data for the resting-state, four sessions of the EEG-fMRI data for
325 the N-back task were taken to estimate the SPLICE filter for the neurofeedback.
326
327 In the NF training experiment, EEG signals were taken in the soundproof shield room
328 from Day 1 to Day 3, and EEG-fMRI data were taken in the MRI scanner at Day 4 as
329 well as the EEG-fMRI simultaneous recording experiment. In the NF training (Day 1, 2
330 and 3), the resting-state EEG took five minutes, then each participant conducted the NF
331 training for 9 min in a single session and completed five sessions.
332
333 In the EEG-fMRI data acquisition, the visual stimuli were projected on an opaque screen
334 set inside the scanner via a (DLA-X7-B, JVC; frame rate = 60 Hz) projector and a mirror
335 system. Participants responded to the stimuli using MRI compatible response pads
336 (HHSC-2 × 2, Current Designs, Inc., PA, USA).
337
338 (Table1 is inserted around here)
339

## 2.4 Data analysis

### 2.4.1 Preprocess of fMRI

The flow of preprocessing of EEG and fMRI data and its integration was illustrated in Figure 3.

We followed a preprocessing protocol from a previous study (Ogawa et al., 2018). The data were processed using SPM8 (Wellcome Trust Centre for Neuroimaging). The first four volumes were discarded to allow for T1 equilibration. The remaining data were corrected for slice timing and realigned to the mean image of that sequence to compensate for head motion. Next, the structural image was coregistered to the mean functional image and segmented into three tissue classes in the MNI space. Using the associated parameters, the functional images were normalized and resampled in a 2 × 2 × 2 mm grid. Finally, they were spatially smoothed using an isotropic Gaussian kernel of 8 mm full-width at half maximum.

For each participant, we extracted fMRI time courses within each ROI. To remove several sources of spurious variance along with their temporal derivatives, linear regression was performed, including six motion parameters in addition to averaged signals over gray matter, white matter, and cerebrospinal fluid. Furthermore, to reduce spurious changes caused by head motion, the data were checked by a method that reduces motion-related artifacts. A high-pass filter (< 0.008Hz) was applied to these sets of time courses prior to the following regression procedure.

(Figure 3 is inserted around here)

364

365 **2.4.2 EEG-fMRI integration analysis for selection of a module for NF training**

366 To extract EEG network features, SPLICE model was optimized from the resting-state
367 EEG data of simultaneously recorded with fMRI with leave-one-session out cross
368 validation for each individual. Prior to training of the SPLICE model, the raw EEG data
369 were cleaned as follows: remove the fMRI scanning and ballistocardiographic artifacts
370 using Brain Analyzer (Brain Products); apply a bandpass filter between 8 and 12 Hz such
371 as alpha band by EEGLAB (Delorme and Makeig, 2004). We decided to reduce 63, 20,
372 and 20 dimensions for the first layer (sources), second layer (modules), and third layer
373 (top-level ICs), respectively. We estimated EEG network features during N-back task
374 through the SPLICE model, and then downsampled at 5 Hz. The EEG network features
375 were convoluted with the hemodynamic response function (HRF) and decimated to
376 match the time instant of fMRI signal. To evaluate correlation analysis with fMRI signal
377 in the general linear model (GLM, SPM12), the design matrix consisted of EEG network
378 features, such as the output of twenty modules to identify whether a statistical cluster
379 was existed in the DLPFC/mFC or PCC/Precuneus thresholding p-value $< 0.001$,
380 uncorrected and cluster level $p < 0.05$ corrected by Family-Wised Error (FWE). Five
381 participants whom a cluster was found in the DLPFC/mFC performed to upregulate the
382 EEG network features, and three participants whom a cluster was identified in the
383 PCC/precuneus performed to downregulate the EEG network feature in the NF training.
384 Of course, the participants did not know upregulation/downregulation in the NF training.

385

386 **2.5 Neurofeedback protocol**

387 A real-time neurofeedback system consisted of a software, OpenViBE (v1.0.0, Renard

et al., 2010) for the real-time EEG signal acquisition, feedback module programed by MATLAB, and Brain Amp MR (Brain Products) with 63-ch EEG electrode cap. EEG raw signals were acquired at 500 Hz through Brain Amp MR and OpenViBE Acquisition Server. In the interactive GUI, OpenViBE Designer, the EEG data was applied a bandpass filter between 8 and 12 Hz and electrooculography artifact removal (e.g., blinking, eye movement) using online recursive ICA (ORICA; Hsu et al., 2016). Afterward, the signals were put into the MATLAB wrapper that predicted EEG network features by the optimized SPLICE model. According to the selected module of SPLICE in advance, an L2-norm of the module was estimated as a targeting EEG network feature, and then displayed a feedback score depending on a task design. In this process, an output of the selected modules was downsampled at 5 Hz.

In the NF training, all participants completed a 5-min resting state EEG recording and five neurofeedback sessions. We collected twenty trials in a single session. A single trial consisted of 'rest' for 6 s, 'induction' for 10 s, and 'feedback' for 4 s. Feedback was calculated as a score (-1 to 1) which was presented as a circle on monitor (Figure 4). For the NF group, the score was a difference of the mean power of between 'induction' and 'rest' period in logarithm scale into -1 to 1 in eleven steps (resolution = 0.2; if the score was less than -1 or more than 1, a circle size was minimum or maximum). On the other hand, for the sham group, the score was randomly generated between -1 to 1. On each trial, during the 'induction' period, participants had been instructed to try their best to "do something with their brain" to get the best feedback possible. During the instructions, a list of example strategies had been provided (e.g., mental calculation, recalling words), so that participants understood what was meant by "do something with your brain". It is

possible that this may have influenced what the participants did. Nonetheless, participants were not told to use any explicit strategies on any given trial or session, meaning that they had to learn how to get favorable feedback via trial and error during the task. None of these related to depression or the N-back task.

To randomly divide participants into two training groups (NF and sham), we confirmed no significant differences of depressive symptoms with two-sample t-test (BDI-II: $p = 0.257$; RRS: $p = 0.317$), but also other scores (OCI: $p = 0.264$; AQ: $p = 0.815$; STAI2(trait): $p = 0.512$) as control shown in Table 2.

(Figure 4 and Table 2 are inserted around here)

# 3. Results

## 3.1 The effect of NF training on scores of psychological questionnaires

In comparison between the NF group and the sham group, we calculated the effect sizes (Cohen's d) from the differences of scores between post- and pre-NF for BDI-II and RRS. In Table 3, we summarized the average (and standard deviation) scores on the questionnaires across two groups. About the scores associated with depressive symptoms, we found large effect sizes in RRS (total: -0.722, brooding: -0.626, reflection: -1.050), but did not find large effect sizes in the BDI-II (total: 0.184). In other scores not directly related with depressive symptoms, we also found large effect sizes in AQ (total, attention switching), OCI (total, obsessing) in Supplemental Table1. As another effect, we found an increasing of STAI2 (trait) in the NF group compared to the sham group, but not in the STAI2 (state). Our results suggest that our NF training may affect to reduce

436 rumination in particular, subscales such as “brooding” and “reflection”.

437

438 (Table3 is inserted around here)

439

440 **3.2 The effect of the NF training on the fMRI signal**

441 We evaluated effects on neural representation due to the NF training in a comparison of
442 fMRI data during N-back task recorded on Visit 2 & 3 and Visit 7 (Day4). In fact, we did
443 not collect the EEG-fMRI data on Visit 4 (Day1), because of practical reasons (scheduling
444 of experiment and less burden on participants), therefore, we decided to compare the
445 fMRI data on Visit 2 & 3 as pre-NF training and Visit 7 as post-NF training in this analysis.
446 In order to statistically evaluate the interactions of groups (NF, sham) and time (Visit 2&3,
447 Visit 7), we applied a full factorial analysis for fMRI data during N-back task using SPM
448 12. Two clusters with a contrast of 2-back > 0-back were identified at the right angular
449 gyrus and superior frontal gyrus 2 (labelled by AAL toolbox; thresholding by p-value <
450 0.001, uncorrected and cluster level $p < 0.05$ corrected by Family-Wised Error (FWE,
451 cluster size > 244 voxels; Figure 5 and Table 4). About behavioral performance of N-
452 back task during the EEG-fMRI simultaneous recording, we did not find significant effects
453 observed in d’ or hit rates due to the NF training because of ceiling effect (Table 5). It is
454 suggested that the neural basis of high cognitive load (2-back) could be processed with
455 less effort in these regions due to the NF training.

456

457 (Table 4, Table 5, and Figure 5 are inserted around here)

458

### 3.3 The effect of the NF training on EEG parameters

Participants in the NF group were given feedback scores calculated by online EEG NF system on each trial end, therefore, we illustrated changes of the feedback scores across three days (Day1 to Day3, Figure 6). Four subjects showed an increased mean of the feedback scores at Day3 higher than the one at Day1, but other subject did not show training effect in the feedback score. From this result, we did not find an obvious training effect represented in the feedback score.

(Figure 6 is inserted around here)

## 4. Discussion

In this study, we investigated the feasibility of an advanced EEG NF training targeting depressive symptoms, in particular, rumination related brain regions in subclinical participants. Before the NF training, EEG network features estimated by SPLICE were evaluated with fMRI data correlated with upregulation in the DLPFC/mFG or downregulation in the PCC/precuneus according to the previous biomarker studies (Ichikawa et al., 2020; Yamashita et al., 2020). To this end, we found large effect sizes in RRS (total, brooding and reflection) in the comparison between NF and sham groups. Here we start to discuss effect of our EEG NF protocol with SPLICE, its effect to neural basis associated with cognitive function, and methodological limitations of this study.

### 4.1 Improvement of depressive symptoms due to EEG NF training

Our hypothesis is that functional regularization in DMN or ECN may contribute the improvement of depressive symptoms, in particular, represented in BDI and RRS, we

then found large effect size on RRS (total, brooding and reflection) compared the NF group with the sham group, but not on BDI. The FCNef protocol of Taylor et al. (2021) for four days' training with fMRI showed significant reduction of BDI and brooding in RRS. In their protocol, targeting FC between the left DLPFC/mFG and the left PCC/precuneus was precisely localized by fMRI during N-back task in the first day of the training. Feedback scores were directly calculated from the targeting FC in real time under the training, therefore, it may help for participants to control the feedback scores. Tsuchiyagaito et al. (2021) also investigated the feasibility of FCNef targeting on FC between the left precuneus and right temporoparietal junction, and found large effect size in difference between the NF and sham groups at the post-NF training on state-rumination assessed by VAS, but not in RRS. Our result was consisted with the FCNef study of Taylor et al. (2021) that showed improvement of depressive symptoms associated with rumination, in particular, brooding. Brooding factor is specifically thought to be associated with depression and mistake concern (Treynor et al., 2003; Hasegawa 2013), which is considered a specific trait of melancholic depression. The EEG network feature estimated from SPLICE was not directly calculated from the target FC like FCNef (Yamada et al., 2017; Taylor et al., 2021). However, we confirmed the statistical equivalence between EEG network features and fMRI data in the selection process of SPLICE module. It may approximately predict the network activity associated with DMN or ECN from EEG data and help to normalize the targeting brain region.

Advantages of our NF protocol are fMRI-informed feature selection and not needed fMRI scanning in the NF training. One of the advanced approaches are simultaneous real-time fMRI and EEG NF (Zotev et al., 2020; Cury et al., 2020). A NF protocol of Zotev et

al. (2020) simultaneously displayed bars of fMRI signals in the left amygdala and frontal EEG asymmetry in the alpha band and required to upregulate these bars involved with positive autographical memory. They conducted their NF protocols for major depressive disorder (MDD) patients and demonstrated through proof-of-concept to reduce depressive symptoms, the profile of mood state (POMS) and STAI. This method can directly modulate the target regions precisely, however, the costs of these treatments are very expensive, and people who has claustrophobia cannot be applied. Instead of the simultaneous real-time fMRI and EEG NF, our approach of EEG NF training using SPLICE that extract EEG network features using SPLICE showed the feasibility of an advanced treatment for depressive symptoms with low cost, and applicable in the realistic environment (not required MRI equipment).

### 4.2 Training effect represented in fMRI activation during N-back task

In a comparison between pre- and post-NF, we found an interaction of brain activity with group (NF, sham) and time (pre, post) in the right angular gyrus and superior frontal gyrus with a contrast of 2-back > 0-back. These brain regions were part of ECN and activated greater in MDD patients than in healthy individuals (Fitzgerald et al., 2008). Though the NF training, participants in the NF group regularized their own brain activity in the close-loop process, therefore, the calculation cost in the right frontal and parietal regions may be decreased even though the same cognitive load. In contrast, participants in the sham group could not succeed to make efforts to get higher scores in the NF training, and may be required to do something for increasing size of the green disc in the induction period.

About behavioral performance of N-back task (0, 1, 2-back) on Day 1 and Day 4, we

expected that the NF group should be improved due to the NF training. In the case of 0, 1, 2-back condition, however, we did no find clear differences because of the ceil effect. One of the reasons is that we did not record EEG-fMRI of 3-back task in our protocol, and we could not accurately evaluate their performance of working memory with higher load. For healthy participants, we designed short duration numeral presentation in our N-back task paradigm such as 200 ms, but this was too easy in the case of 1 and 2-back condition to see effects of the NF training. For future studies, we need to improve our N-back task protocol to correctly evaluate the work memory performance.

### 4.3 Methodological limitation

In this study, subclinical healthy participants completed the NF training, therefore, the scores of depressive symptoms were lower than previous patient studies (Wang et al., 2019; Young et al., 2017; Zotev et al., 2020). Some participants were low scores from the beginning of the NF (Day 1) and did not show enough improvement due to the NF training because of the floor effect. Taylor et al. (2021) and Tsuchiyagaito et al. (2021) have conducted the NF training with healthy participants and is comparable to our study. One of difficulty is recruiting appropriate participants for this experiment, in fact, we recruited about a hundred participants as screening for subclinical participants with depressive symptoms. To this end, we completed the NF training for seventeen participants, however, it was not enough number of participants to show stable statistics like Wang et al. (2019). To improve the sensitivity of pre-screening, we should conduct questionnaires not only BDI, but also Self-rating Depression Scale (SDS; Zung, 1965). For depressive symptoms, the EEG NF protocol of Wang et al. (2019) have demonstrated a proof of concept for patients, with comorbid MDD and anxiety symptoms,

555 comparing three groups, frontal alpha asymmetry (ALAY group), high-beta
556 downregulation training (Beta group), and control. Eighty-seven patients completed 10
557 sessions NF training, twice a week, for five weeks. To this end, ALAY and Beta group
558 showed a reduction of BDI-II about 10 points, and a reduction of Beck Anxiety Inventory
559 (BAI) score about eight points. Both of psychological questionnaires showed significant
560 interaction comparing to the control group. This intensive and long-term NF training may
561 be enough to induce large effects on depressive symptoms. In our protocol, NF training
562 was set for three days, and healthy individuals were recruited. To consider a practical
563 application, there is a problem of scheduling and long-term treatment to achieve this NF
564 training. For example, a case of Taylor et al. (2021) is four-days training, Tsuchiyagaito
565 et al. (2021) and Zotev et al. (2020) are just one-day training. In making a NF protocol,
566 balance of training intensity and efficacy is a crucial issue, but critical point in the realistic
567 treatment, therefore, we have to carefully design the NF experiment.
568
569 To extract EEG network features, we proposed SPLICE that extend ISA assumed
570 hierarchical network structure in the EEG data. In the optimization process, SPLICE
571 automatically pools several components which are highly correlated, and estimates time
572 course of network feature including several components. For integration analysis of
573 EEG-fMRI data, for instance, HRF-convoluted frequency powers were calculated to
574 match time instants between EEG and fMRI data, then were applied GLM to identified
575 brain regions or networks (Mantini et al., 2007). Another approach is combined with EEG-
576 microstate analysis which divides four to six states based on the clustering algorithm,
577 then identifies brain regions using correlation analysis (Brechet et al., 2019; Al Zoubi et
578 al., 2020). These methods may achieve dimension reductions and simply describe the

579 relationship between fMRI and EEG data. However, each extracted component was not
580 considered a network structure which already known. To quantify the accuracy of
581 SPLICE model, we have to evaluate several conditions of EEG-fMRI data for future
582 studies.
583
584 In summary, normalization of the targeted brain regions in the DLPFC/mFC and
585 PCC/precuneus in the three days NF training have demonstrated reduction of depressive
586 symptoms such as rumination in a single-blind test. EEG network features extracted by
587 SPLICE demonstrated a feasibility of EEG NF treatment for depressive symptoms for
588 sub-clinical healthy individuals. Our results have showed a possibility of transferable NF
589 training from fMRI to EEG with sufficient accuracy.
590
591

592 # References

593 Beck, A. T., Steer, R. A., Brown, G. K. (1996). Beck Depression Inventory-II (BDI-II) The
594 Psychological Corporation, Harcourt Brace; Toronto.

595 Cheon, E. J., Koo, B. H., Choi, J. H. (2016). The Efficacy of Neurofeedback in Patients
596 with Major Depressive Disorder: An Open Labeled Prospective Study. Appl
597 Psychophysiol Biofeedback, 41 (1), 103-10.

598 Cury, C., Maurel, P., Gribonval, R., Barillot, C. (2020). A Sparse EEG-Informed fMRI
599 model for hybrid EEG-fMRI neurofeedback prediction. Front Neurosci, 13.
600 https://doi.org/10.3389/fnins.2019.01451.

601 Delorme, A., Makeig, S. (2004). EEGLAB: an open-source toolbox for analysis of single-
602 trial EEG dynamics, J Neurosci Methods, 134, 9-21.

603 Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y. et al. (2017).

604 Resting-state connectivity biomarkers define neurophysiological subtypes of
605 depression. Nat Med. 23, 28–38.

606 Fitzgerald, P. B., Srithiran, A., Benitez, J., Daskalakis, Z. Z., Oxley, T. J., Kulkarni, J. et
607 al. (2008). An fMRI study of prefrontal brain activation during multiple tasks in patients
608 with major depressive disorder. Hum Brain Mapp. 29 (4), 490–501.

609 Hasegawa, A. (2013). Translation and initial validation of the Japanese version of the
610 Ruminative Responses Scale. Psychol. Rep. 112, 716-726.

611 Hirayama, J., Hyvärinen, A., Kawanabe, M. (2017). SPLICE: fully tractable hierarchical
612 extension of ICA with pooling. PMLR, 70, 1491-1500.

613 Hsu, S. H., Mullen, T. R., Jung T. P. (2016). Real-Time Adaptive EEG source separation
614 using online recursive independent component analysis. IEEE Trans on Neural Syst
615 Rehabil Eng. 24(3), 309-19.

616 Hyvärinen, A. and Hoyer, P. O. (2000). Emergence of phase and shift invariant features
617 by decomposition of natural images into independent feature subspaces. Neural
618 Comput, 12(7), 1705–1720.

619 Ichikawa, N., Lisi, G., Yahata, N., Okada, G., Takamura, M., Hashimoto, R. et al. (2020).
620 Primary functional brain connections associated with melancholic major depressive
621 disorder and modulation by antidepressants. Sci Rep. 10:3542, doi: 10.1038/s41598-
622 020-60527-z.

623 Mantini, D., Perrucci, M. G., Cugini, S., Ferretti, A., Romani, G. L., Del Gratta, C. (2007).
624 Complete artifact removal for EEG recorded during continuous fMRI using independent
625 component analysis. Neuroimage, 34 82), 598-607.

626 Misaki, M., Tsuchiyagaito, A., Al Zoubi, O., Paulus, M., Bodurka, J., Tulsa 1000
627 Investigators. (2020). Connectome-wide search for functional connectivity locus
628 associated with pathological rumination as a target for real-time fMRI neurofeedback
629 intervention, Neuroimage Clin., 26(102244).

630 Northoff, G. (2016). Spatiotemporal psychopathology I: no rest for the brain's resting

631 state activity in depression? Spatiotemporal psychopathology of depressive symptoms.
632 J Affect Disord. 190, 854–66.

633 Ogawa, T., Aihara, T., Shimokawa, T., Yamashita, O (2018). Large-scale brain network
634 associated with creative insight: combined voxel-based morphometry and resting-state
635 functional connectivity analyses. Sci. Rep. 8. doi: 10.1038/s41598-018-24981-0.

636 Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V. et al. (2010).
637 OpenViBE: an open-source software platform to design, test and use brain-computer
638 interfaces in real and virtual environments, Presence, 19(1), 35-53.

639 Spielberger, C.D. (1983). State-trait anxiety inventory for adults. Palo Alto, CA: Mind
640 Garden.

641 Taylor, J. E., Yamada, T., Kawashima, T., Kobayashi, Y., Yoshihara, Y., Miyata, J. et al.
642 (2021). Reduction of brooding and more general depressive symptoms after fMRI
643 neurofeedback targeting a melancholic functional-connectivity biomarker. medRxiv,
644 https://doi.org/10.1101/2021.01.21.20248810.

645 Treynor, W., Gonzalez, R., Nolen-Hoeksema, S. (2003). Rumination reconsidered: A
646 psychometric analysis. Cognit Ther Res. 27, 248−259.

647 Tsuchiyagaito, A., Misaki, M., Al Zoubi, O., Tulsa 1000 Investigators, Paulus, M., Bodurka,
648 J. (2021). Prevent breaking bad: A proof of concept study of rebalancing the brain's
649 rumination circuit with real-time fMRI functional connectivity neurofeedback, Hum Brain
650 Mapp., 42, 922-940.

651 Wang, S. Y., Lin, I. M., Fan, S. Y., Tsai, Y. C., Yen, C. F., Yeh, Y. C. et al. (2019). The
652 effects of alpha asymmetry and high-beta down-training neurofeedback for patients with
653 the major depressive disorder and anxiety symptoms. J Affect Disord 257, 287–96.

654 Yamashita, A., Sakai, Y., Yamada, T., Yahata, N., Kunimatsu, A., Okada, N. et al. (2020).
655 Generalizable brain network markers of major depressive disorder across multiple
656 imaging sites. PLoS Biol. 18(12), e3000966.

657 Yamashita, A., Sakai, Y., Yamada, T., Yahata, N, Kunimatsu, A., Okada, N. et al., (2021).

658 Common brain networks between major depressive-disorder diagnosis and symptoms
659 of depression that are Validated for Independent Cohorts. Front. Psychiatry, 10 June
660 2021, https://doi.org/10.3389/fpsyt.2021.667881

661 Yahata, N., Kasai, K., Kawato, M. Computational neuroscience approach to biomarkers
662 and treatments for mental disorders. Psychiatry Clin Neurosci. 71, 215–237 (2017).

663 Yamada, T., Hashimoto, R., Yahata, N., Ichikawa, N., Yoshihara, Y., Okamoto, Y. et al.
664 (2017) Resting-state functional connectivity-based biomarkers and functional MRI-
665 based neurofeedback for psychiatric disorders: a challenge for developing theranostic
666 biomarkers. Int. J. Neuropsychopharmacol. 20, 769-781.

667 Young K.D., Zotev, V., Phillips, R., Misaki, M., Yuan, H., Drevets, W. C. et al. (2014).
668 Real-time FMRI neurofeedback training of amygdala activity in patients with major
669 depressive disorder. PloS One, 9, e88785.

670 Young, K.D., Siegle, G., J., Zotev, V., Phillips, R., Misaki, M., Yuan, H. et al. (2017).
671 Randomized clinical trial of real-time fMRI amygdala neurofeedback for major
672 depressive disorder: effects on symptoms and autobiographical memory recall. Am. J.
673 Psychiatry, 174, 748-755.

674 Young, K.D., Zotev, V., Phillips, R., Misaki, M., Drevets, W.C., Bodurka, J. (2018).
675 Amygdala real-time functional magnetic resonance imaging neurofeedback for major
676 depressive disorder: A review. Psychiatry Clin Neurosci. 72(7), 466–481.

677 Yu, S. H., Tseng, C. Y., Lin, W. L. (2020). A neurofeedback protocol for executive function
678 to reduce depression and rumination: a controlled study. Clin Psychopharmacol
679 Neurosci, 18 (3), 375-385.

680 Zotev, V., Phillips, R., Yuan, H., Misaki, M., Bodurka, J. (2014). Self-regulation of human
681 brain activity using simultaneous real-time fMRI and EEG neurofeedback. Neuroimage,
682 85, 985–95.

683 Zotev, V., Yuan, H., Misaki, M., Phillips, R., Young, K. D., Feldner, M. T. et al. (2016).
684 Correlation between amygdala BOLD activity and frontal EEG asymmetry during real-
685 time fMRI neurofeedback Training in Patients with Depression. Neuroimage Clin, 11,

224–38.

Zotev, V., Mayeli, A., Misaki, M., Bodurka, J. (2020). Emotion self-regulation training in major depressive disorder using simultaneous real-time fMRI and EEG neurofeedback. Neuroimage. Clin 27, 102331.

Zung, W. W. K. (1965). A self-rating depression scale. Arch Gen Psychiatry, 12, 63-70. doi: 10.1001/archpsyc.1965.01720310065008.

## Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Author Contributions

Takeshi Ogawa and Hiroki Moriya designed this study. Takeshi Ogawa and Nobuo Hiroe performed data acquisition. Takeshi Ogawa and Jun-ichiro Hirayama analyzed the data. Takeshi Ogawa, Jun-ichiro Hirayama, and Motoaki Kawanabe discussed data analysis and results. Takeshi Ogawa and Jun-ichiro Hirayama wrote the original draft.

## Funding

This study was supported by JSPS KAKENHI Grant Number JP17H06041, JP18H05395, JP18KK0284, JP21K12620, JP21K12055, JP21H03516, the ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), CREST, JST, the MIC/SCOPE/ #192107002.

## Acknowledgments

We thank Yoko Matsumoto for her help in recruiting participants, scheduling and conducting the experiment; Kana Inoue and Tachi Kaori for their help in screening participants in the Screening experiment; Mihoko Sato, Ayumi Yuki and Yoshiko Itakura for their supporting preparation of the EEG-fMRI and NF training experiments; Takashi Yamada for discussion to design the neurofeedback experiment. Furthermore, we would like to thank the team from the ATR Brain Activity Imaging Center (BAIC), in particular, Akikazu Nishikido and Ichiro Fujimoto for their technical support.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author, Takeshi Ogawa, upon reasonable request.

## Contribution to the Field Statement

Advanced treatments for depressive symptoms, such as a real-time fMRI neurofeedback have been proven by several studies. Regularization of functional connectivity (FC) between executive control network (ECN) and default mode network (DMN) during fMRI neurofeedback have been proposed to reduce depressive symptoms. However, it is difficult to install this system in practice because the cost is high and no practical signal processing techniques have existed to extract FC-related features from EEG. In this regard, stacked pooling and linear components estimation (SPLICE), recently proposed as a multilayer extension of independent component analysis and related independent

731 subspace analysis, can be a promising alternative, although its application to the resting-
732 state EEG have never been investigated. This method may help to modulate the target
733 FC in the EEG-based neurofeedback training. Here, we describe an EEG neurofeedback
734 paradigm for depressive symptoms using an EEG network features estimated by
735 SPLICE. We hypothesized upregulation of ECN or downregulation of DMN should
736 specifically predict decreases in depressive symptoms. Finally, we found large effect size
737 in the rumination response scale in the comparison between neurofeedback and sham
738 groups. We demonstrated a feasibility of EEG neurofeedback treatment for depressive
739 symptoms using EEG network features in the subclinical trials.
740

# Figure legends

## Figure 1.

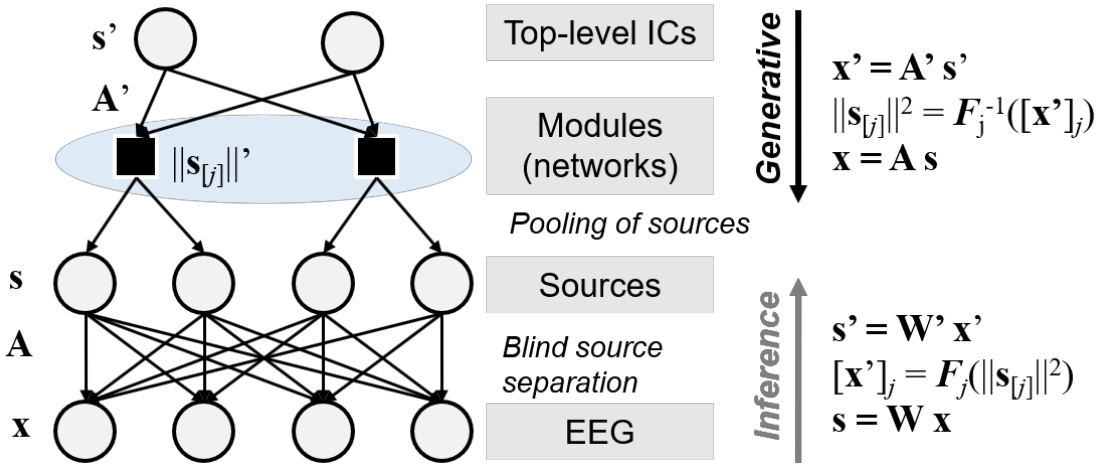
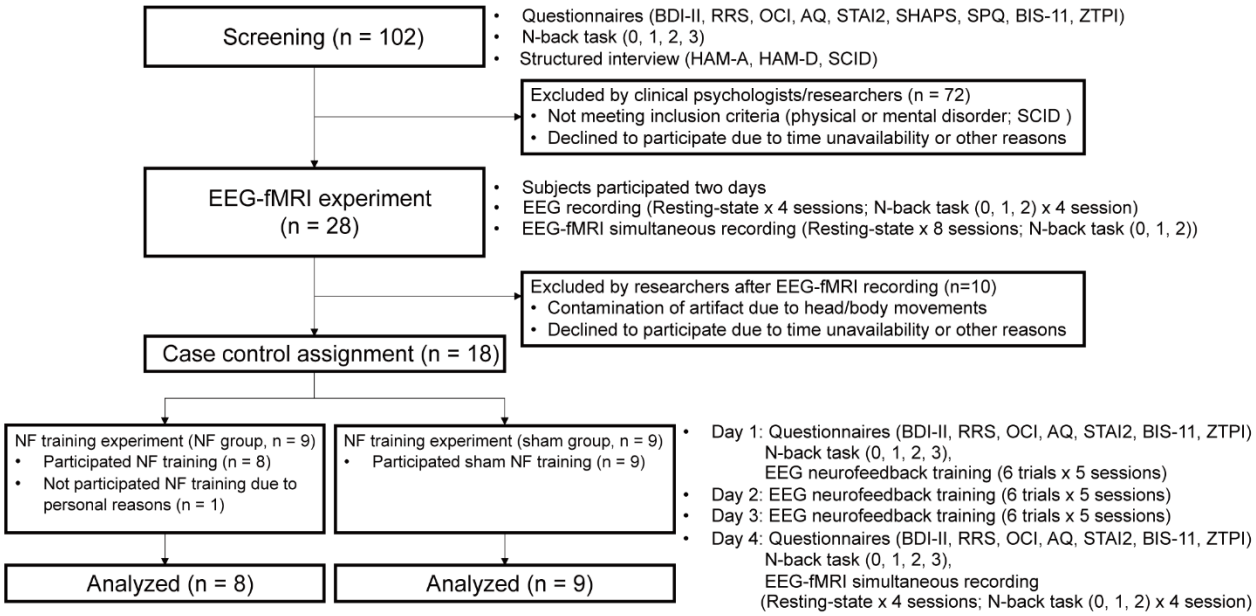



**Figure1. Structure of SPLICE.**

**s**: unobserved source signal; **A**: mixing matrix; **W**: demixing matrix of **A**; **x**: EEG measurements; **A**': invertible mixing matrix; **s**': source vector; $F_j$: link function; W': demixing matrix of **A**'.

749 **Figure 2.**



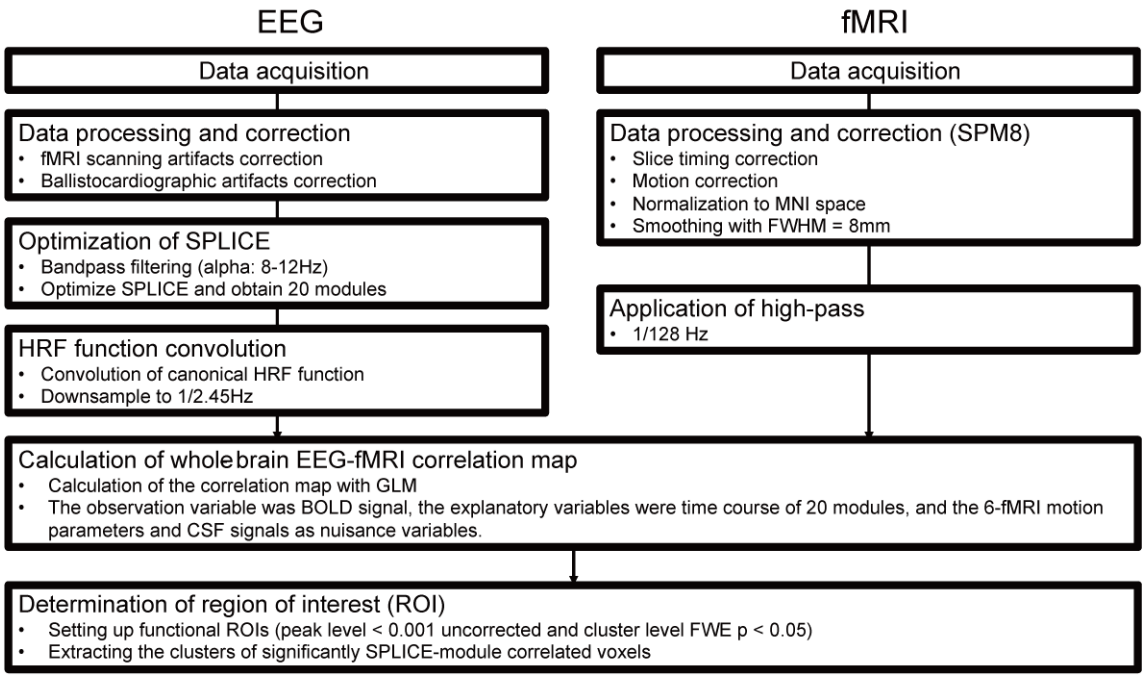

750

751 **Figure2. A flow of the experiment**

752 BDI-II: Beck Depression Inventory II; RRS: Rumination Response Scale; OCI:
753 Obsessive-Compulsive Inventory; AQ: Autism-Spectrum Quotient; STAI2: State-Trait
754 Anxiety Inventory, SHAPS: Snaith-Hamilton Pleasure Scale; SPQ: Schizotypal
755 Personality Questionnaire; BIS-11: Barratt Impulsiveness Scale 11; ZTPI: Zimbardo Time
756 Perspective Inventory; HAM-D: 21-item Hamilton Depression Rating Scale ; HAM-A:
757 Hamilton Anxiety Rating Scale; SCID: Structured Clinical Interview for DSM-IV-TR Axis
758 disorders (SCID); NF: Neurofeedback.

759

760 **Figure 3.**




761

762 **Figure3. A flow of preprocess of EEG-fMRI data and its integration analysis in this**

763 **study.**

764 SPLICE: stacked pooling and linear components estimation; HRF: hemodynamic

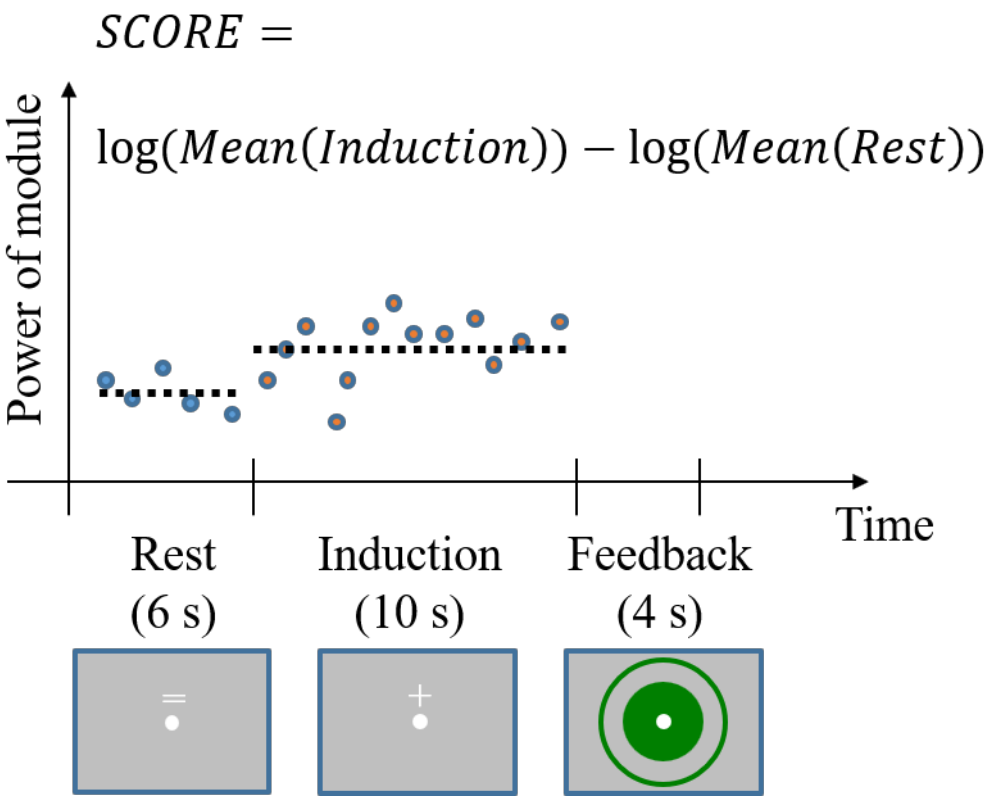
765 response function; MNI, Montreal Neurological Institute; FWHM, full width at half

766 maximum; GLM, General Linear Model; BOLD, blood-oxygen-level dependent; CSF,

767 cerebrospinal fluid; FWE, family-wise error.

768

769 **Figure 4.**

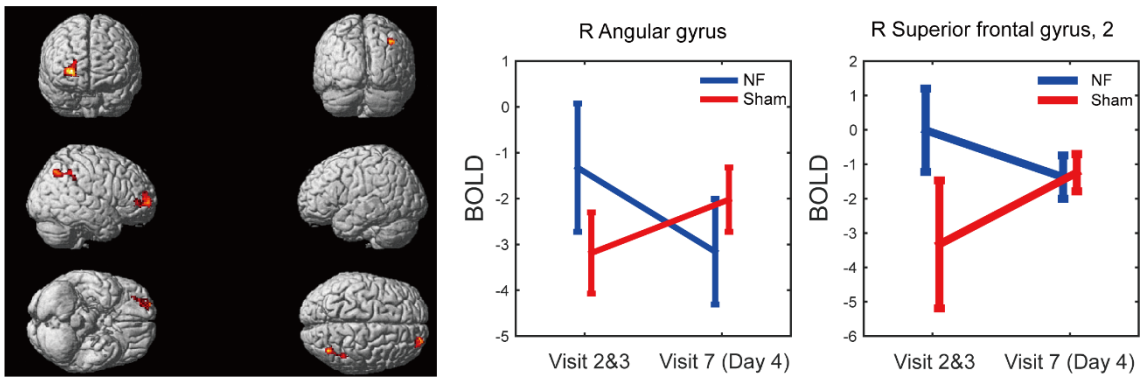



770

771 **Figure4. A protocol of EEG NF training in a single trial.**

772

773 **Figure 5.**




774

775

776 **Figure 5. Identification of clusters illustrated the interaction of groups and times.**

777 Clusters showed significant interaction of group (NF, sham) and time (pre: Visit 2&3, post:

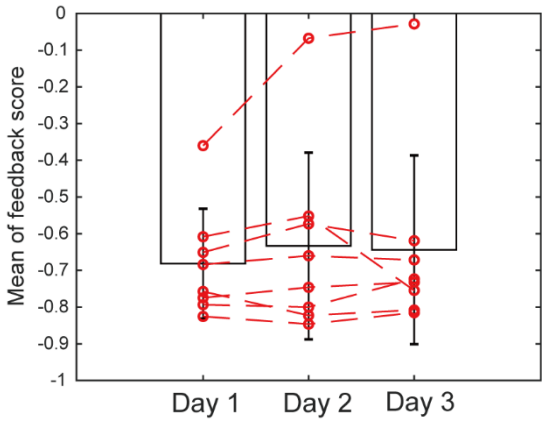
778 Visit 7 (Day4)). A contrast was 2-back > 0-back, and labelled by AAL toolbox;

779 thresholding by p-value < 0.001, uncorrected and cluster level p < 0.05 corrected by

780 Family-Wised Error (FWE, cluster size > 244 voxels).

781

782 **Figure 6.**




783

784 **Figure 6. Changes of feedback scores across three days NF training.**

785 Red dotted line showed mean of feedback scores for each participant.

786

# Tables

## Table 1. Experimental Outline

| | Screening | EEG-fMRI | | NF training | | | |
|---|---|---|---|---|---|---|---|
| | Visit 1 | Visit 2 | Visit 3 | Visit 4 (Day 1) | Visit 5 (Day 2) | Visit 6 (Day 3) | Visit 7 (Day 4) |
| **Questionnaires** | BDI-II, RRS, OCI, AQ, STAI2, SHAPS, SPQ, BIS_11, ZTPI | – | – | **BDI-II, RRS**, OCI, AQ, STAI2, BIS-11, ZTPI | – | – | **BDI-II, RRS,** OCI, AQ, STAI2, BIS-11, ZTPI |
| **Structured interview** | HAM-A, HAM-D, SCID | – | – | – | – | – | – |
| **N-back task** | 0, 1, 2, 3 | – | – | 0, 1, 2, 3 | – | – | 0, 1, 2, 3 |
| **EEG data** | – | -5-min resting state×4 -N-back task × 4 (0, 1, 2) | | -5-min resting-state -NF training ×5 | -5-min resting-state -NF training× 5 | -5-min resting-state -NF training ×5 | – |

| | | | | | | |
|---|---|---|---|---|---|---|
| **EEG-fMRI data** | – | -5-min resting state×8 -N-back task×4 (0, 1, 2) -T1-weighted image | – | – | – | 5-min resting state×4 -N-back task×2 (0, 1, 2) -T1-weighted image |

NOTE: BDI-II: Beck Depression Inventory II; RRS: Rumination Response Scale; OCI: Obsessive-Compulsive Inventory; AQ: Autism-Spectrum Quotient; STAI2: State-Trait Anxiety Inventory, SHAPS: Snaith-Hamilton Pleasure Scale; SPQ: Schizotypal Personality Questionnaire; BIS-11: Barratt Impulsiveness Scale 11; ZTPI: Zimbardo Time Perspective Inventory; HAM-D: 21-item Hamilton Depression Rating Scale ; HAM-A: Hamilton Anxiety Rating Scale; SCID: Structured Clinical Interview for DSM-IV-TR Axis disorders (SCID); NF: neurofeedback.

797 **Table 2. The demographics of participants in the NF training**

798 NOTE: NF: neurofeedback; BDI-II:Beck Depression Inventory II; RRS: Rumination

799 Response Scale

| **Variables** | | **NF group (n = 9)** | **Sham group (n = 9)** | **p-value of t-test** |
|---|---|---|---|---|
| **Age (years)** | | 28.62 ± 10.84 | 25.97 ± 7.22 | 0.606 |
| **Gender** | Female | 2 | 2 | |
| | Male | 7 | 7 | |
| **BDI-II** | | 7.89 ± 3.26 | 10.44 ± 6.50 | 0.257 |
| **RRS** | | 40.67 ± 9.18 | 35.44 ± 12.91 | 0.317 |

800

801 **Table 3. Comparison of statistical effect size associated with depressive**

802 **symptom scores between NF and sham group**

| **Variables** | **NF group (n = 8)** | | **Sham group (n = 9)** | | **Cohen's d** |
|---|---|---|---|---|---|
| | **Pre** | **Post** | **Pre** | **Post** | |
| **BDI-II (total)** | 5.25 (5.52) | 3.75 (4.43) | 7.22 (7.72) | 5.33 (6.34) | 0.184 |
| **-Cognitive depression** | 2.25 (3.37) | 1.75 (2.76) | 2.56 (3.47) | 2.44 (3.05) | -0.390 |

| | | | | | |
|---|---|---|---|---|---|
| **-Somatic depression** | 3.00 (3.46) | 2.00 (2.61) | 4.67 (4.66) | 2.89 (3.76) | 0.479 |
| **RRS (total)** | 31.00 (5.68) | 28.25 (4.40) | 31.33 (8.00) | 31.67 (8.70) | -0.722 |
| **-Brooding** | 7.50 (5.68) | 6.63 (1.92) | 7.78 (2.11) | 7.89 (2.93) | -0.626 |
| **-Depression** | 16.25 (2.96) | 15.88 (2.95) | 17.67 (5.48) | 17.56 (5.00) | -0.101 |
| **-Reflection** | 7.25 (1.98) | 5.75 (0.89) | 5.89 (1.45) | 6.22 (1.64) | -1.050 |

803

804

805 **Table 4. Comparisons of Pre (Visit 2 & 3)/Post (Day 4) fMRI data during N-**

806 **back task.**

| Brain region | Peak coordinate (MNI) | | | Cluster size | | Details |
|---|---|---|---|---|---|---|
| | X | Y | Z | Voxels | Peak of T score | |
| 0-back task | | | | | | |
| R Angular gyrus | 38 | -66 | 44 | 244 | 5.38 | R Angular<br>R Inferior parietal gyrus<br>R Supramarginal gyrus |
| R Superior frontal gyrus, 2 | 34 | 44 | 2 | 330 | 4.17 | R Superior frontal gyrus, 2<br>R Middle frontal gyrus, 2 |

807

808

809 **Table 5. Comparison of N-back task performance during EEG-fMRI**

810 **simultaneous recording**

| | | NF group | | Sham group | | Cohen's d |
|---|---|---|---|---|---|---|
| Measure | Level | Visit 2 & 3 (pre) | Visit 7 (post) | Visit 2 & 3 (pre) | Visit 7 (post) | |
| D' | 0-back | 4.12 ± 0.1 | 4.32 ± 0.1 | 4.32 ± 0.1 | 4.28 ± 0.1 | 0.48 |
| | 1-back | 3.85 ± 0.6 | 4.11 ± 0.4 | 4.20 ± 0.4 | 4.28 ± 0.2 | 0.33 |
| | 2-back | 3.18 ± 0.9 | 3.67 ± 0.7 | 3.52 ± 0.7 | 3.81 ± 0.6 | 0.30 |
| p (hit) | 0-back | 0.93 ± 0.2 | 0.99 ± 0.0 | 0.99 ± 0.0 | 0.99 ± 0.0 | 0.50 |
| | 1-back | 0.90 ± 0.2 | 0.96 ± 0.1 | 0.98 ± 0.1 | 0.99 ± 0.0 | 0.36 |
| | 2-back | 0.81 ± 0.2 | 0.91 ± 0.1 | 0.88 ± 0.1 | 0.94 ± 0.1 | 0.22 |

811

812

813