scientific reports

OPEN

Stratification of responses to tDCS intervention in a healthy pediatric population based on resting-state EEG profiles

Paulina Clara Dagnino[1,7], Claire Braboszcz[1], Eleni Kroupi[1], Maike Splittgerber[2], Hannah Brauer[3], Astrid Dempfle[4], Carolin Breitling-Ziegler[5], Alexander Prehn-Kristensen[3], Kerstin Krauel[5], Michael Siniatchkin[6], Vera Moliadze[2] & Aureli Soria-Frisch[1,7]✉

Transcranial Direct Current Stimulation (tDCS) is a non-invasive neuromodulation technique with a wide variety of clinical and research applications. As increasingly acknowledged, its effectiveness is subject dependent, which may lead to time consuming and cost ineffective treatment development phases. We propose the combination of electroencephalography (EEG) and unsupervised learning for the stratification and prediction of individual responses to tDCS. A randomized, sham-controlled, double-blind crossover study design was conducted within a clinical trial for the development of pediatric treatments based on tDCS. The tDCS stimulation (sham and active) was applied either in the left dorsolateral prefrontal cortex or in the right inferior frontal gyrus. Following the stimulation session, participants performed 3 cognitive tasks to assess the response to the intervention: the Flanker Task, N-Back Task and Continuous Performance Test (CPT). We used data from 56 healthy children and adolescents to implement an unsupervised clustering approach that stratify participants based on their resting-state EEG spectral features before the tDCS intervention. We then applied a correlational analysis to characterize the clusters of EEG profiles in terms of participant's difference in the behavioral outcome (accuracy and response time) of the cognitive tasks when performed after a tDCS-sham or a tDCS-active session. Better behavioral performance following the active tDCS session compared to the sham tDCS session is considered a positive intervention response, whilst the reverse is considered a negative one. Optimal results in terms of validity measures was obtained for 4 clusters. These results show that specific EEG-based digital phenotypes can be associated to particular responses. While one cluster presents neurotypical EEG activity, the remaining clusters present non-typical EEG characteristics, which seem to be associated with a positive response. Findings suggest that unsupervised machine learning can be successfully used to stratify and eventually predict responses of individuals to a tDCS treatment.

In recent years, Non Invasive Brain Stimulation (NIBS) techniques and, in particular, transcranial Direct Current Stimulation (tDCS)[1,2], have been successfully used to study brain function, treat mental disorders and enhance cognitive functions in adults. TDCS therapy can be applied as an effective treatment option[3,4], which is particularly interesting in the absence of alternative interventions. Its application in paediatric populations however has been limited due to initial concerns[5]. In children and adolescents the stimulation parameters cannot be directly translated from studies made in adult populations since the electric fields generated in the brain during stimulation are affected by anatomical differences such as brain volume, skull thickness, white matter - gray matter ratio and cerebrospinal fluid (CSF) volume[6].

[1]Neuroscience BU, Starlab Barcelona SL, Av Tibidabo 47 bis, Barcelona, Spain. [2]Institute of Medical Psychology and Medical Sociology, University Medical Center Schleswig-Holstein, Kiel University, Kiel, Germany. [3]Department of Child and Adolescent Psychiatry, Center for Integrative Psychiatry Kiel, University Medical Center Schleswig-Holstein, Kiel, Germany. [4]Institute of Medical Informatics and Statistics, University Hospital Schleswig Holstein, Kiel University, Kiel, Germany. [5]Department of Child and Adolescent Psychiatry and Psychotherapy, University of Magdeburg, Magdeburg, Germany. [6]Clinic for Child and Adolescent Psychiatry and Psychotherapy, Protestant Hospital Bethel, University of Bielefeld, Campus Bielefeld Bethel, Bielefeld, Germany. [7]These authors contributed equally: Paulina Clara Dagnino and Aureli Soria-Frisch. ✉email: aureli.soria-frisch@starlab.es

Attention-Deficit Hyperactivity Disorder (ADHD) is one of the most prevalent neurodevelopmental conditions, which affects 5.9% of youth and 2.5% of adults (in America, Europe, Australia, and Asia)[7]. Electroencephalography (EEG) has been used to characterize the clinical population with an ADHD diagnosis. The use of EEG-based markers overcomes some limitations of classical subjective assessments based on interviews and physical examinations. Existing EEG biomarkers include an increased theta band power and decreased beta band power in the frontocentral brain region[8,9], which is generally known as slowing, phase-amplitude coupling (beta-gamma) deficits in the frontal-left hemisphere[10], and modified peak alpha frequency[11] among others.

Given the concerns that drug-based treatments of ADHD have raised in the past, there is an ongoing search for treatments with reduced or no side effects[12]. Current drug-based treatments and behavioral therapy in ADHD are also challenged by the existence of different phenotypes and related symptoms of the pathology, as well as the complexity associated with the developing brain[13]. TDCS interventions for ADHD constitute an interesting alternative to drug treatments[14]. However, an important factor to take into account in the development of efficient tDCS interventions is the inter-subject variability of the effects of tDCS[15,16]. Here, machine learning techniques have been proposed to identify subgroups with differential tDCS responses based on the clustering of behavioral responses during tDCS interventions[15]. Moreover, other works have implemented unsupervised methods to characterize responses to different treatments based on the clustering of genetic[17], clinical[18] and functional Magnetic Resonance Imaging (fMRI)[19] data. These works report the presence of patient subgroups within an a priori uniform group of patients. We propose to perform patient stratification based on EEG data. Previous work using EEG data attains the prediction of response to treatment through supervised Machine Learning (ML) approaches[20,21]. Variability in response to TMS and tDCS interventions in patients suffering from depression or anxiety can be also explained by individual differences in baseline EEG activity[22,23]. EEG is an attractive modality for clinical application due to its lower cost and faster application time than other neuroimaging modalities. In the present study, we have developed an unsupervised clustering method applied to EEG data to extract so-called digital phenotypes[24], making use of EEG for the first time to stratify tDCS response in a healthy pediatric population.

As an exploratory study in preparation for the main clinical trial, a tDCS intervention in a healthy pediatric population has been designed to target 2 areas of interest: the left Dorsolateral Prefrontal Cortex (lDLPFC) and the right Inferior Frontal Gyrus (rIFG). TDCS is applied to each area of interest both while the participants are performing a task during the stimulation (concurrent task condition) and while the participants are at rest during the stimulation (non-concurrent task condition). Previous studies on tDCS were not conclusive enough about the need to have a concurrent task to increase the response effect of tDCS[25,26]. Therefore, one of the objectives of the exploratory study was to find out whether the response was better with or without a concurrent task. Additionally, three tasks were always performed independently of the stimulation condition (active vs sham, concurrent vs non-concurrent task) and always once the stimulation session has concluded. They were conducted after the stimulation for the sake of evaluating the effect of the stimulation condition in terms of behavioral parameters. Resting state EEG was acquired before and after the tDCS intervention.

We present a methodology for healthy participants' stratification based on their profile of resting state EEG activity before the tDCS stimulation. The effect of the tDCS intervention on their performance in behavioral tasks performed after the tDCS stimulation is taken into account to label the resulting groups. The considered behavioral metrics correspond to Accuracy and Reaction Times of the Flanker, N-Back, and Continuous Performance Task (CPT) tasks. The performance effect is assessed as the difference in these metrics between the active and the sham conditions. In order to determine the baseline EEG profiles, we extracted spectral features that were found relevant in previous studies in pediatric EEG studies[27] and in particular ADHD[8,28]. We then computed the relative band power—this measure is preferred over the absolute band power to normalize absolute differences among participants who may present abnormalities in their anatomy[29]. Participant stratification was performed using the algorithms denoted as Spectral Clustering, which has good results for non-convex data[30], and Fuzzy Clustering, which is robust when boundaries are ambiguous and data contain outliers[31]. A correlational analysis was then applied between the membership of participants to each of the electrophysiologically homogeneous groups and the behavioral assessment metrics after stimulation. Significant correlations of a particular cluster group are tested against the rest of the groups in order to confirm responders. The methodology and results we describe in this article contribute to the understanding of the variability of response to tDCS treatment, with a particular focus on the specific conditions in which the treatment is effective. We support the relationship between the tDCS response and the EEG profiles on the fact that both depend on bioelectrical features of the brain. Hence the EEG features act as a proxy of the biophysical characteristics in each individual's brain.

Interestingly, the positive responder groups correspond to participants with a non-typical profile of EEG activity, presenting either a slowing of the EEG or a spread of increased alpha rhythm over frontal areas. Our methodology could be applied in the future to stratify patients for tDCS treatment, enabling to focus on those specific digital phenotypes in which tDCS is more likely to become effective. Personalization of treatment, i.e. applying tDCS only to groups of patients predicted to have positive results, would save costs, and time, and improve the efficacy of the treatment at lower sample sizes.

Materials and methods

Participants.
Data were obtained from a healthy population taking part in a clinical trial for the development of treatments for Attention Deficit Hyperactivity Disorder (ADHD) and Autism Spectrum Disorder (ASD) based on tDCS. The study was approved by the local ethics committee of the Medical Faculty at Kiel University, Germany, and was carried out in accordance with the latest revision of the Declaration of Helsinki and registered in the German Trial Register with number DRKS00008207.

A total of N=56 children and adolescents aged 10–17 years (32 females, mean age: 14.09 years, SD: 2.1) were recruited in the Medical Faculty at Kiel University. Informed consent from the participants and their parents was obtained before the experiment. The exclusion criteria consisted of substance or medication consumption, implants and devices in the body, pregnancy, birth before pregnancy week 37, birth weight lower than 2500 grams, IQ score lower than 80 assessed with CFT 20-R Test (Culture Fair Intelligence Test)[32], neurological and psychiatric disorders (past or present), brain surgery, presence of epilepsy in the participant or family history, social and health impairments as assessed with the Child Behavior Checklist (CBCL)[33], the German ADHD Rating scale FBB ADHS[34] and the Social Responsiveness Scale (SRS)[35]. Participants' age, sex, IQ and Edinburgh Handedness Inventory (EHI) are given in Tables 2 and 3 of the Supplementary Information.

Experimental protocol. *Experimental design.* A randomized, sham-controlled, double-blind, crossover study design was implemented. Participants came for 6 experimental sessions (see Fig. 1). The first 2 sessions (T1-T2) correspond to participants screening and an optional fMRI scan. Participants then came back 4 times to receive the tDCS interventions (T3-T6) which were of 4 different types, administered in a randomized order: active tDCS with a concurrent task, active tDCS with a non-concurrent task, sham tDCS with a concurrent task and sham tDCS with a non-concurrent task. Additionally to this experimental condition, which includes the concurrent task in the concurrent-task conditions and no task in the non-concurrent task conditions, 3 tasks were included in the experimental design for behavioral effect evaluation. These 3 tasks were conducted after the stimulation finished for the sake of evaluating the effect of active or sham stimulation in terms of behavioral parameters. These 3 tasks were always performed independently of the stimulation condition (active vs. sham, concurrent vs. non-concurrent task) and always once the stimulation session had concluded. There was a minimum washout period of 7 days in-between sessions. At the start of each stimulation session, participants were asked to complete an assessment of their current mood, motivation, and any adverse event since the last session. Following the stimulation session, participants completed a questionnaire on the safety, tolerability, and blinding of the stimulation.

Participants were randomly divided into 2 groups: participants in Group A were stimulated in the left Dorsolateral Prefrontal Cortex (lDLPFC) and performed the N-Back Task in the concurrent task condition, whereas participants in Group B were stimulated in the right Inferior Frontal Gyrus (rIFG) and performed the Flanker task in the concurrent task condition. In each stimulation session, both groups randomly received either a sham or an active tDCS stimulation. Each tDCS stimulation session unfolded as follows: once participants were equipped with the stimulation headset, resting state EEG was recorded during 2 min eyes open (EO) and 2 min eyes closed (EC), followed by 20 min of tDCS stimulation. In the concurrent task condition, the task started after 2.5 min of tDCS stimulation and ended 2.5 min before the end of the stimulation. In the non-concurrent task condition, participants were instructed to sit and relax with eyes opened while receiving tDCS. For all participants, after stimulation, EEG data were recorded during a second resting state period (2 min EO, 2 min EC) and then during the execution of the following cognitive tasks: Flanker Task, N-Back Task and Continuous Performance Task (Fig. 2).

Transcranial direct current stimulation. Transcranial direct current stimulation (tDCS) was applied using a Starstim 32 multi-channel stimulation device (Neuroelectrics - Barcelona, Spain). The montage consisted of 5 circular electrodes (3.14 cm²) filled with gel. The total stimulation current consisted of anodal 2mA during 20 min, with a 30 s of ramp up and down at the beginning and end of the stimulation (see Fig. 3). The interested reader can find details on the DLPFC[36] and the IFG[37] montages in the referred works.

EEG recording. EEG data were recorded using the same device used for the stimulation at a sampling rate of 500 Hz with 32 channels, following the 10-20 system: P8, T8, AF7, AF8, F8, F4, C4, P4, FC5, AF4, Fp2, Fp1, AF3, Fz, C6, Cz, C5, PO3, O1, Oz, O2, PO4, Pz, Fpz, FC6, P3, C3, F3, F7, FCz, T7, P7.

Behavioral tasks. Behavioral tasks included the Flanker Task, N-Back Task and Continuous Performance Test (Fig. 4). These were programmed using Presentation® Software (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, USA). Behavioral performance metrics were computed in R (version 3.6.1, R Core Team). The primary endpoints are the N-Back Accuracy for participants in Group A and the Flanker Accuracy for participants in Group B. For more details please see the associated experimental work[36].

EEG signal processing and feature extraction. The EEG data analysed here correspond to the 2-min resting state EEG in eyes closed condition recorded before the start of tDCS stimulation in each T3 to T6 trial for
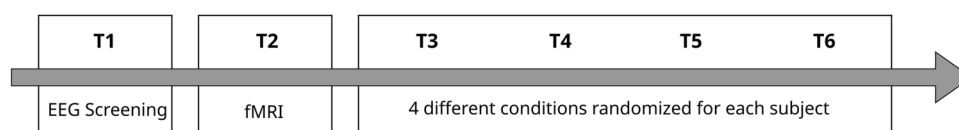

Figure 1. Experimental design. Each participant is assessed at 6 time points, T1 corresponds to an EEG screening, a screening questionnaire and informed consent, T2 to the optional fMRI, and T3-T6 to the randomized interventions (offline/online, tDCS/sham).
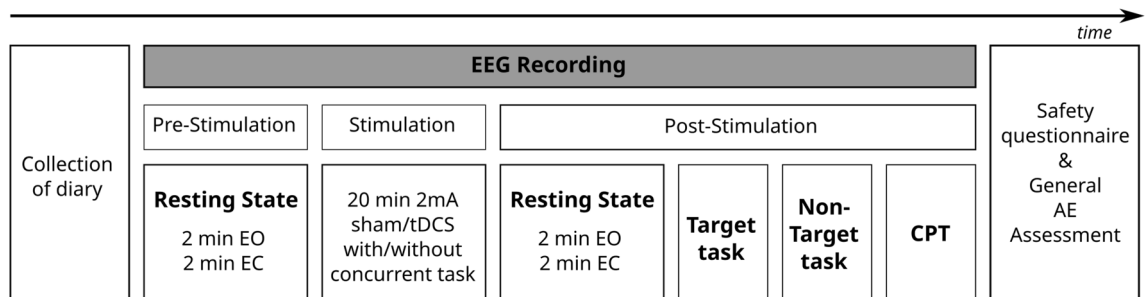
Figure 2. T3 to T6 trial design. During each trial, 2 min eyes open and 2 min eyes closed resting state EEG are being recorded before and after the tDCS stimulation. Participants then came back 4 times to receive 4 different tDCS interventions (T3–T6) in a randomized order: active tDCS with a concurrent task, active tDCS with a non-concurrent task, sham tDCS with a concurrent task and sham tDCS with a non-concurrent task. After each tDCS intervention, 3 tasks were included in the experimental design for behavioral effect evaluation. These 3 tasks were always performed after the stimulation session had ended, and independently of the stimulation condition (active vs. sham, concurrent vs. non-concurrent task). Each stimulation session lasted 20 min. For the concurrent type of intervention, the task started and ended 2.5 min after and before the stimulation respectively (in order to avoid current ramping distractions). The tasks performed after the stimulation had a duration of approximately 16 min for Flanker Task and N-Back N-Back Task, and approximately 18 min for CPT.
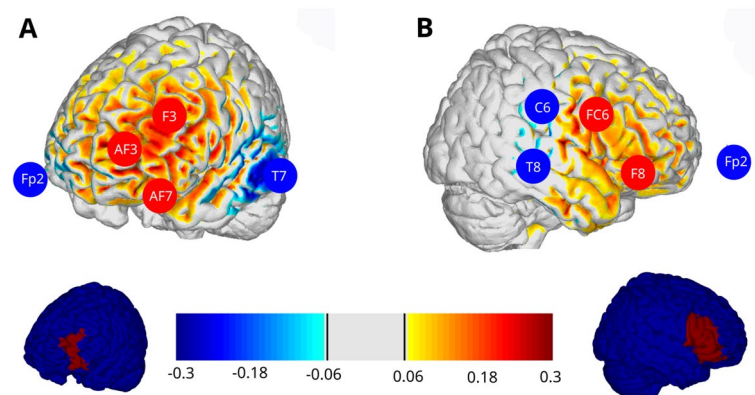

Figure 3. Stimulation areas. Brain renderings showing the component of the E-field normal to the cortical surface (En), positive/negative values indicate the into/out-of the cortical surface direction of the En, leading to an increase/decrease of excitability in cortical pyramidal cells. Group A of participants received stimulation in (**A**) left Dorsolateral Prefrontal cortex (see target area in the bottom row A). Group B of participants received stimulation in (**B**) right Inferior frontal gyrus, the montage was optimized defining a target of Broadmann-areas 44, 45 and 47 on the right hemisphere (see target area in the bottom row B).

each participant. Data were first band pass filtered using a 1000 coefficients FIR filter with cut-off frequencies at 2Hz and 45 Hz, demeaned and detrended. Frontal channels AF4, Fp2, Fp1 and AF3 were removed from further processing and analysis due to the presence of numerous eye-movement related artifacts. High amplitude artifacts were removed using an amplitude threshold of 100 μV. Finally, the data were re-referenced to electrode Cz, which was then excluded from further processing. The relative band power was computed in 4-s time-windows epochs with 50% overlap in bands delta (2–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma (30–45 Hz). The final relative band power values for each subject and session were taken as the median of the epochs. The EEG data pre-processing was done in Python 3.9 using custom scripts and Python packages (i.e. MNE).

Overview of the analytical approach. Figure 5 gives a schematic of the methods used in the core analytical approach of our work. Details of each step are given in the following Sections. All analysis stages were implemented in Python 3.9 using Scikit-learn and SciPy packages.

Unsupervised clustering analysis. We applied 2 different unsupervised clustering algorithms to the spectral features extracted from the EEG data: Spectral Clustering[30] (implementation from Python scikit-learn package[41]) and Fuzzy C-Means[42] (implementation from Holt Skinner 2018[43]). The definition of parameters for the clustering pipeline consisted of selecting the optimal number of clusters k ranging from 2 to 6, and the specific parameters for each algorithm (number of components n for Spectral Clustering and fuzzification param-
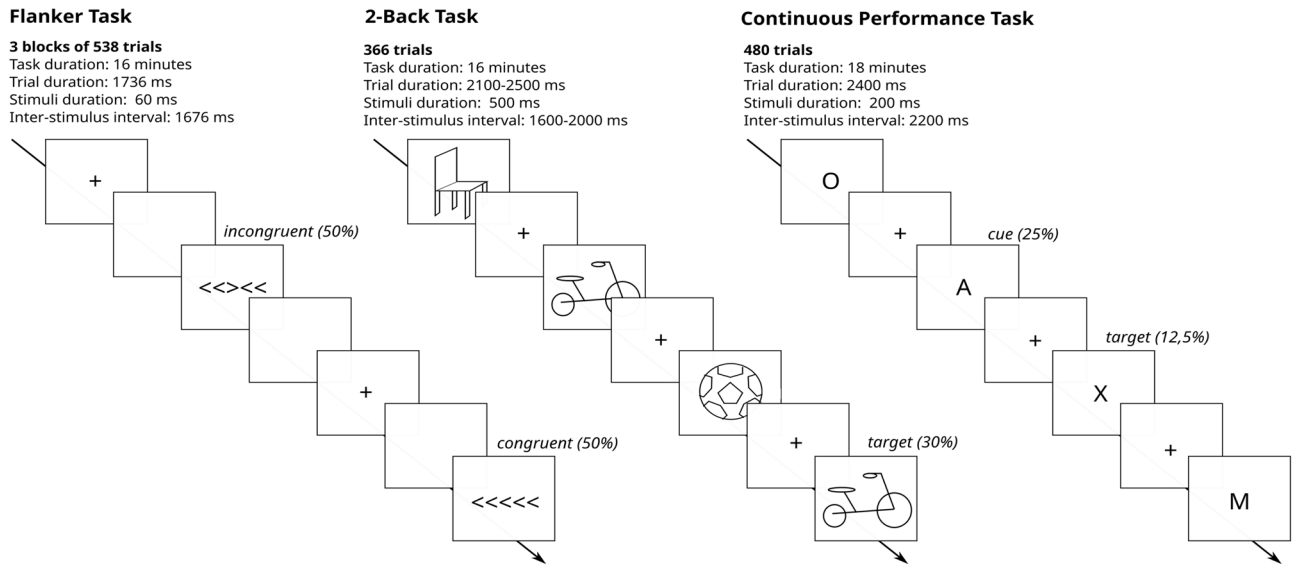
Figure 4. Behavioral Tasks. Examples for Flanker Task (left), 2-Back Task (middle) and Continuous Performance Task (right). In the Flanker Task[38], 5 arrows in a row were shown in which the middle one (target) points to the same or opposite direction as the rest (distractors). Participants had to indicate if the target pointed left or right. During the 2-Back Task[39] participants had to detect if a presented picture is the same as a picture shown 2 steps back. In the Continuous Performance Task[40], participants were presented with a series of letters. They were shown pairs of letters in which a cue is followed by a target. They had to respond when the target appears, only when it is preceded by the cue (which is only one of the 4 existent possibilities).
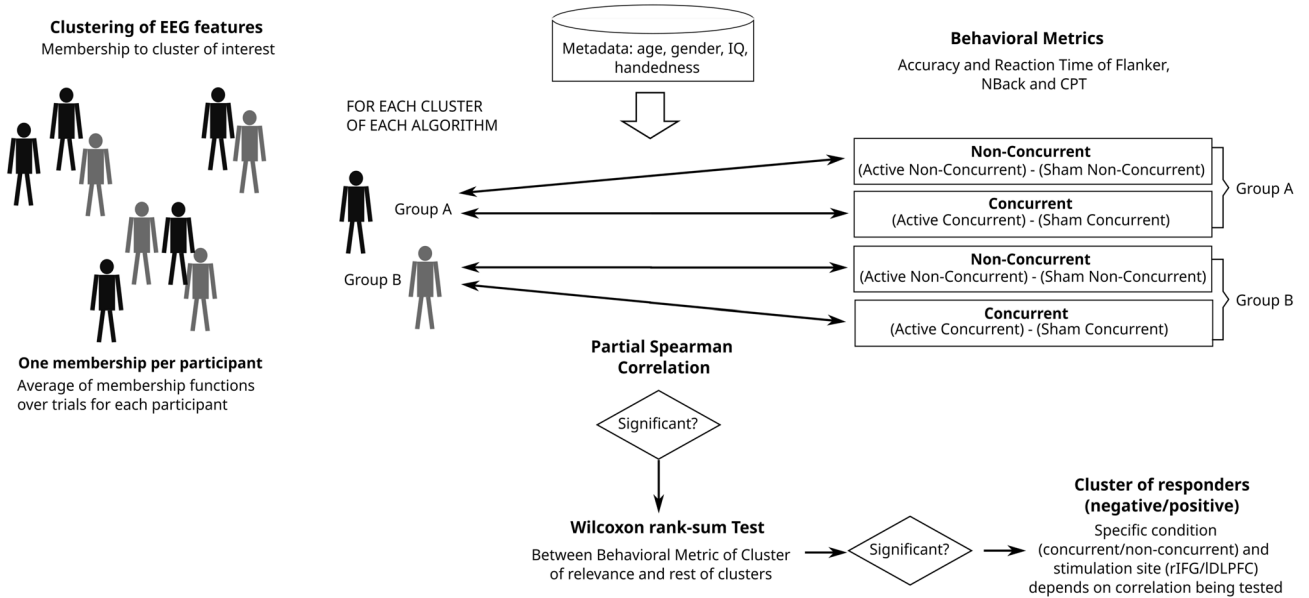
Figure 5. Schematic of clustering and Wilcoxon test pipeline. EEG features are clustered and membership values are correlated with different behavioral metrics. If significance is reached (correlation p-value is less than 0.05), a Wilcoxon Test is done between the behavioral metric of that cluster and the rest of the points.

eter value m for Fuzzy C-Means). The parameters were chosen based on several internal validity measures and the visual inspection of the feature space in a 2-dimensional projection based on the t-distributed stochastic neighbour embedding (t-SNE). The used internal validity metrics are: Silhouette, Davies Bouldin Score (DB), Calinski-Harabasz (CH), Bayesian Information Criteria (BIC) and Inertia (or Within-Cluster Sums of Squares). Homogeneous clustering is represented by low values of DB, Inertia and BIC, and high values of Silhouette and CH. We experienced a relative lack robustness of these metrics in our large dimensional space (5 different band power values for each of the 27 EEG channels). Therefore we added a second criteria for the selection of the parameters, namely the t-distributed stochastic neighbor embedding (tSNE) visualization. TSNE is appropriate for maintaining local interactions, addressing non-linearity and robust to the presence of outliers[44], and showed

to be flexible enough to visually depict in a very clear manner the clusters determined by the employed algorithms. We have also used PCA for visualization and found tSNE was offering a better visualization in terms of interpretability. This may be due to PCA considering linear interactions, and not being robust to the presence of outliers that spuriously increase the variance of some dimensions.

Cluster membership definition and computation. For each participant we used 4 independent recordings (corresponding to each pre-stimulation EEG resting state recorded in sessions T3–T6). We thus obtained 4 different membership functions for each subject, which may present small variability due to the time between recording sessions, i.e. 1 week. The FCM algorithm directly provides cluster prototypes and a fuzzy membership function of each particular datapoint to the different clusters. In Spectral Clustering, which does not provide these as direct outputs, cluster prototypes are first calculated as the average of the points belonging to each specific cluster. The membership is then calculated as a probability function of the Euclidean distance from each datapoint to the aforementioned prototype. As we were trying to establish a participant profile, we averaged the cluster membership functions of each subject over the 4 sessions. This average membership function is expected to be more robust than the individual session membership and to get rid of the temporal variability of subject profiles.

It is worth pointing out that the membership of each data sample to each of the clusters is defined in a real-valued domain in order to avoid loosing information due to the binarization of the data. By defining memberships in a real-valued domain we are able to treat differently the points that are close to the cluster centres and the points that are far from them, i.e. since they present very different membership values to the cluster under consideration. This is expected to increase the robustness of the analysis pipeline, given the fact that early binarization thresholds largely influence final results.

Characterizing digital EEG phenotypes through statistical analysis.
Correlation between clusters and behavioral data. After clustering, we correlated the membership function and the behavioral data of the tasks after tDCS of each participant using a Partial Spearman Correlation. This type of correlation allows to take into account spurious factors that may affect the EEG features. The correlation was corrected for age, sex, IQ and handedness. The False Discovery Rate (FDR) multiple comparisons correction was applied in order to avoid random significant values of the correlation measure using the Benjamini and Hochberg step-up procedure (BH)[45].

The purpose of the correlational analysis is to label the clusters of homogeneous EEG digital phenotypes that we obtained either as positive, negative or non-responders to the tDCS intervention depending on their behavioral response. As behavioral response we used the accuracy and reaction times in the Flanker task, N-Back and CPT tasks performed after the tDCS stimulation. In order to obtain a unique behavioural measure to conduct the correlation, we computed the difference in performance between the active and sham tDCS conditions for each task and behavioural metric (accuracy and reaction time). This performance difference was computed independently for trials with a concurrent task during tDCS and trials without a concurrent task.

From combining the 2 stimulation sites (Groups A/B) with the 2 task concurrency conditions (concurrent/non-concurrent) we obtained 4 different experimental conditions and thus for each clustering algorithm 4 correlations were computed. Hence we correlated the performance difference between active and sham tDCS of each experimental condition and the average membership to the cluster of each subject. The rationale for computing these 4 correlations is to identify the concrete condition, i.e. stimulation site A/B and task concurrence, in which significant results are found and predict later the response of the subjects who fall into the same cluster. If there is a significant correlation (results are considered statistically significant if the corrected correlation p-value is less than 0.05), we can establish a relationship between the digital EEG profile and the behavioral performance difference between the active and sham conditions. Clusters of positive responders are defined as showing positive accuracy correlations and negative RT correlations, i.e. positive responders show higher accuracy and smaller RTs following active tDCS than following sham tDCS. On the contrary, clusters of negative responders, are defined as presenting a negative correlation with Accuracy and positive correlation with RT measures.

Statistical comparison of behavioral measures between clusters of positive and negative responders using the Wilcoxon Test. This step aims at testing the significance of the difference in behavioral response between each specific clusters of positive or negative responders. If any of the 4 correlations computed for each cluster showed significance (positive or negative), we then analysed the difference in behavioral response between the specific cluster members and all non-members of the cluster. Hence, a rank-sum Wilcoxon Test was calculated on the behavioral metrics within each cluster samples and the samples belonging to the remaining clusters. The FDR Benjamini and Hochberg step-up procedure[45] was used here again to correct for multiple comparisons.

In this part of the analysis we want to assess if there are significant differences between the behavioral outcomes for each specific cluster, once the cluster has shown a correlation between the membership to that cluster and the corresponding behavioral metric. We illustrate this operation with an example: If a significant correlation is found for Cluster 1 in behavioral metric Accuracy of the Flanker task—i.e. considering the membership values of all data-points to that specific Cluster 1—then a Wilcoxon test is performed on the Flanker Accuracy active-sham difference. In this test we compare the performance of the subjects who belong to Cluster 1 versus the rest of the subjects. The objective of this step is to have an estimation of the improvement in the treatment response in case the clustering analysis is used for recruiting subjects in a future clinical trial.

At this point, since we have been calculating the significant difference between behavioral response of the subject samples belonging to a cluster and the rest of the subject samples, we need to determine to which final cluster a given subject belongs to. This was done using the arg max function over the membership values to the different clusters, i.e. determining the cluster with a highest membership value (e.g., if point X has a membership value of 0.2, 0.3, 0.4 and 0.1 for Clusters 1, 2, 3 and 4 respectively, then it would belong to Cluster 3).

Results

Feature space analysis. We computed the relative band power in the delta, theta, alpha, beta and gamma frequency bands for each electrode, resulting in a feature space of 135 components. After removal of any artifacts and due to missing trials or dropout subjects, 206 data points were used in the analysis. Figure 6 shows the feature space visualization with t-SNE in 2 dimensions, given by an optimal perplexity index value of 30[44]. One interesting aspect here is to observe how reproducible is our approach given that for each subject there was a 1-week delay between the recording of each of the 4 samples. We can observe that the samples of all 4 trials for each participant are very close to each other. Therefore we can state that the intra-subject variability is small enough to consider the feature representation as a robust one, which gives grounds for averaging the membership functions of a participant's 4 trials in order to obtain a unique membership per subject for the correlational analysis.

Clustering results. We implemented the different clustering algorithms to establish homogeneous subgroups among subjects based on their resting-state EEG spectral features before intervention. The evolution of the validity measures over the different parameters does not offer a clear criterion for parameter selection. We believe this may be due to the high dimensionality of the feature space, i.e. the "curse of dimensionality"[46], which seems to be not clearly tackled by the most common validity measures. Nevertheless, we observe some singular points in the evolution of the selected validity measures over clustering parameters (see Fig. 7B). Given the shortcom-
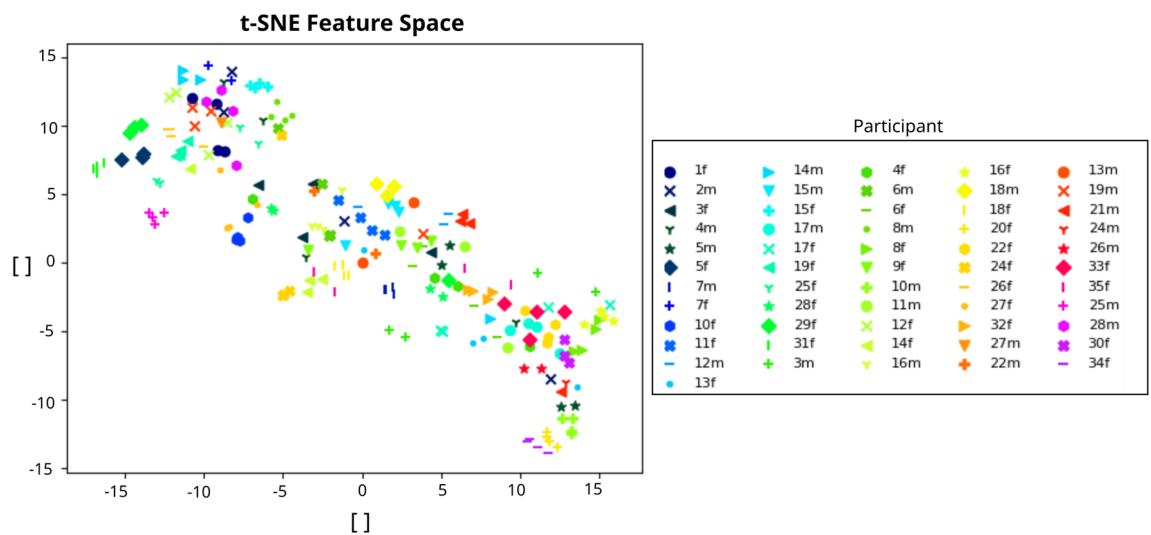

Figure 6. Feature space. Visualization of EEG data-points with t-SNE in 2 dimensions. Each point corresponds to a specific trial for each participant, the features of each data-point are the RBP in each channel.
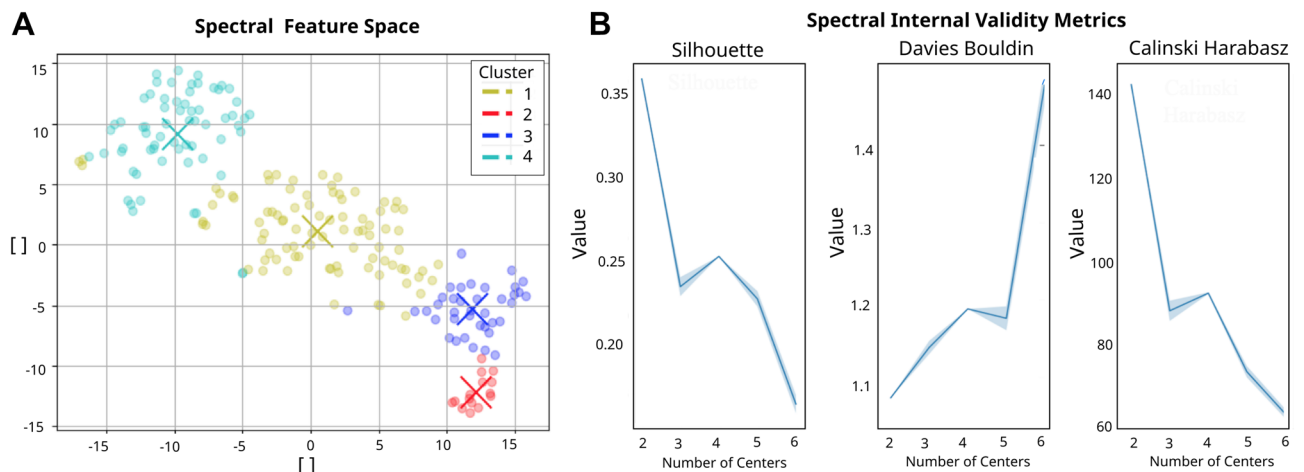

Figure 7. Clustering results for spectral clustering. The left plot shows the feature Space, the colour code represents each cluster. From all data-points, 42,7% lie in Cluster 1, 7.3% in Cluster 2, 18.4% in Cluster 3 and 31.6% in Cluster 4. The right plots show the evolution of the internal validity metrics Silhouette and Davies Bouldin when changing the number of centres. An elbow at k=4 can be seen.

ing of the validity measures to offer clear objective criteria for parameter selection, we have taken into account an additional qualitative evaluation criteria based on the visualization of the t-SNE projection of the clustered feature space. Figure 7 A shows well separated and clear cluster clouds. For the Spectral Clustering, a number of components n equal to the number of centres was used, as it generally corresponds to the default parameter. This gives better results than a number of components n equal to 100 applied as an example for extreme values, and gives similar results than small values such as 2, 6 and 8. For FCM the chosen parameter was a fuzzification parameter m of 1.7 since it differs sufficiently from 1 (so as to be more dissimilar to K-means) and presenting optimal values for the validity metrics and visualization in the feature space.

Concretely we can observe some singular points, i.e. "elbows", for $k = 4$ clusters for the Spectral Clustering algorithm that correspond to the optimal values. For the case of the FCM algorithm it is optimal in $k = 5$. In order to have the largest amount of data-points in each cluster cloud, and since it shows clearly defined clusters in the visualization from both algorithms, we eventually defined the number of cluster centres as $k = 4$. Since the clustering output visualized in the feature space shows similar results for both algorithms, a possible hypothesis for the light differences in point cluster assignment between Spectral and FCM is the difference in the way the membership functions are computed in each algorithm (FCM results for $k = 4$ can be found in Supplementary Information for the sake of comparison). Given the similarity of the clustering results the results in the following sections are shown only for the Spectral Clustering.

Clustering validation. The robustness of the parameter search in the clustering methodology was validated through an M-hold-out using 80% of the data for parameter search. The chosen algorithms with their corresponding parameters (i.e. fuzzification parameter m of 1.7 for FCM and a number of components n equal to the number of centers for Spectral Clustering) and number of centers k were found out in a subset of the data points and then tested in the remaining points, checking whether the clusters assigned in the test set generalize from the ones obtained on the training set. This was done by first clustering the train data samples and plotting them in the feature space. The remaining test data sample labels were then assigned according to their nearest distance to the final center obtained previously.

The validation is successful based on the visual overlap between the clusters assigned to the test set, which are not used in the parameter search (see Figure 6A in Supplementary Information). Moreover, the internal validity metrics follow the same trend in both data subsets (see Figure 6B in Supplementary Information).

Digital electroencephalography phenotypes. In order to characterize each of the clusters in terms of the associated digital EEG profiles, the prototypes generated for each cluster of each algorithm were used. The electrophysiological profiles are set up as a combination of a representation in the frequency and spatial domains. First, the power spectral density (PSD) for each cluster prototype is plotted (Fig. 8) using the mean of occipital and parietal electrodes (i.e. O1, O2, Oz, P8, P4, P3, P7, PO3, PO4, Pz). Taking the alpha peak during eyes closed as a reference, differences can be observed between the different clusters. These can be further analyzed on hand of the scalp power distributions in the spatial domain. Topoplots are used for the representation of prototypes in the spatial domain (Fig. 9). They are visualized by calculating the mean of the relative band power of all of the data points belonging to each cluster.

Cluster 1 presents the neurotypical scalp power distribution for eyes closed, with alpha power localized over posterior regions. This could be observed as well in the PSDs (see Fig. 8). Cluster 2 presents a very clear slowing in both the PSD figure and the topographical maps, from the alpha band towards the theta band. Cluster 3 presents low amplitude at 10 Hz and large amplitudes at higher power values in the PSD. Cluster 4 presents the largest amplitude at 10 Hz and a peak at 20 Hz in the PSD, and the difference between frontal and parietal areas for the alpha band in the topoplot is smaller compared to the rest of the clusters. As a remark, Cluster 1 topographical map of activation is the most similar to the ones of the overall study population (see Fig. 10).
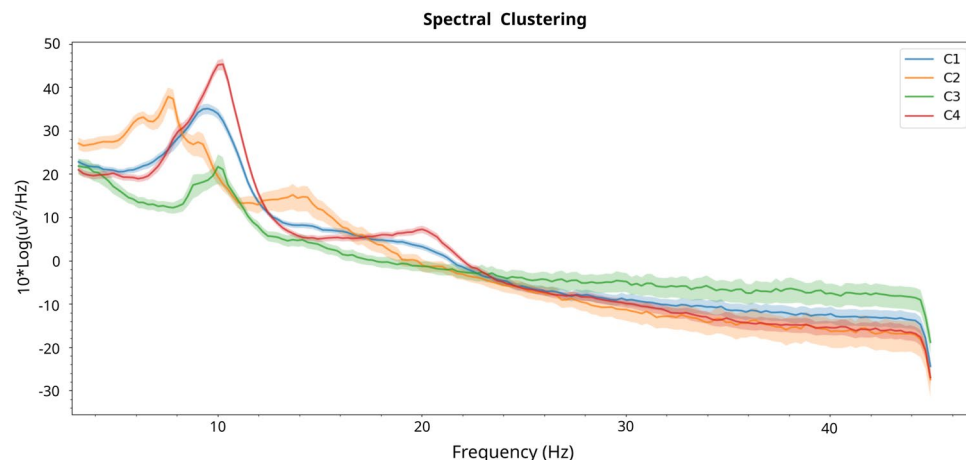

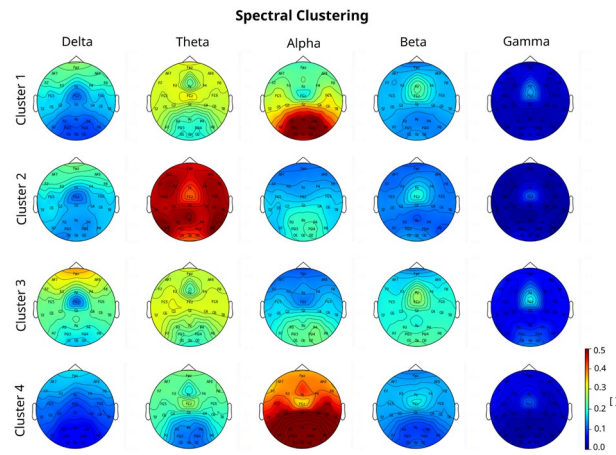**Figure 8.** Power spectrum density of each cluster for spectral clustering results.

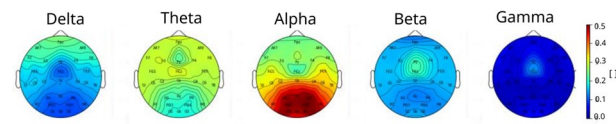**Figure 9.** Power spectrum plot of each cluster for spectral algorithm results.

Figure 10. Overall population topographical map.

Behavioral correlational analysis. We applied the described correlation analysis between participants' membership to each cluster and the post-tDCS behavioral results of Accuracy and RT for the Flanker, N-Back and CPT tasks (Fig. 6), the average of the membership function over all trials of each participant was computed as formerly mentioned. Eleven subjects were excluded due to incorrect randomization and/or dropouts during the trials, and behavioral outliers are removed for each metric using a threshold of $+/-2.5$ standard deviations. The correlational analysis can be used to label the clusters in terms of their positive or negative responding nature. A specific example of the positive response is that of cluster 2, whose members are positive responding to the rIFG stimulation with non-concurrent task (see Fig. 11). Negative responses are detected by the correlational analysis as well (see Figure 5A in Supplementary Information).
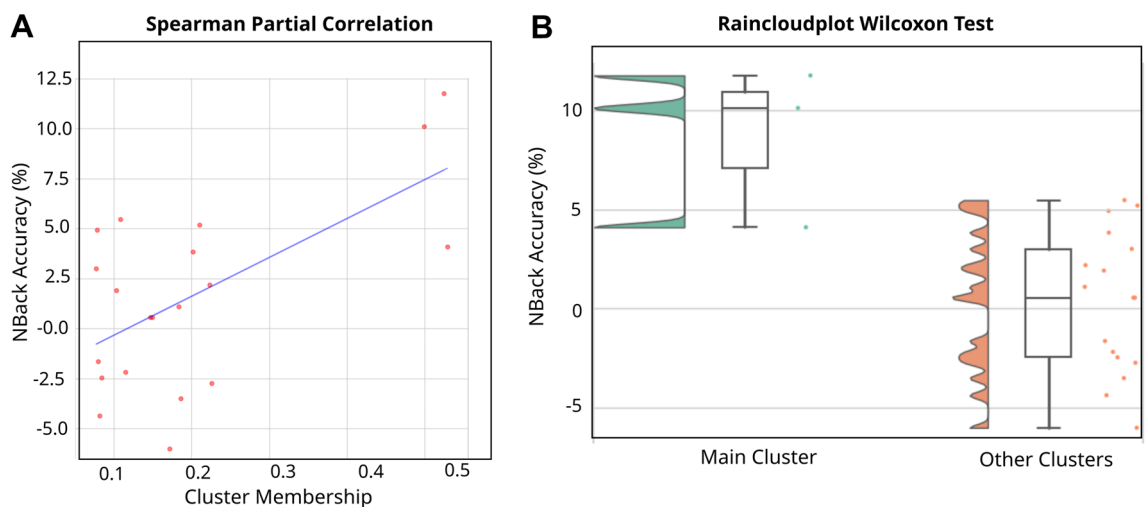

Figure 11. Example of significant positive responder group for Spectral Clustering (Cluster 2), with rIFG stimulation with no concurrent task, measured with N-Back Accuracy. (**A**) Correlation between cluster membership of pre treatment EEG spectral features of all participants and N-Back Accuracy at the end of the trials. Best-fit line showing least square polynomial fit. (**B**) Wilcoxon Test between N-Back Accuracy of Cluster 2 Spectral members, and the members of the rest of the Clusters. The raw data is visualized as points as well as the probability density, and key summary statistics in the boxplot (median, mean, and confidence intervals).

The following matrices (Fig. 12) show the strength of the partial Spearman correlation and the range of the *p*-values when significant (* − 10%,** −5% and *** − 1%) after correction for multiple comparison with FDR (24 corrections for Group B, 20 corrections for Group A). These matrices summarize the findings achieved by the correlational analysis.

Differences between cluster versus non-cluster members. We computed the rank-sum Wilcoxon test only in clusters with significant correlation results. Responder clusters (positive or negative) are analyzed with respect to the difference to non-responder ones. Responding clusters are confirmed by giving significant differences in this test as well (see Fig. 11 for an example of a positive responder cluster and Figure 5B in Supplementary Information for a negative one). Corrected results for multiple comparisons for the Wilcoxon analysis, as well as number of corrections, are presented in Table 1.

The robustness of the correlational analysis was tested through a 1000-hold-out validation with different percentages of the data set (specifically 80%, 60% and 40%). We find a trend for the coefficients and *p*-values of both correlations and Wilcoxon tests (see example in Supplementary Figure 7). Overall, the subsets have the same direction of correlation as the whole dataset, and the strength goes diminishing in the sequence 80%–60%–40%. With respect to the correlation *p*-value, it is almost always significant (correlation *p*-value is less than 0.05) for the 80% dataset. In the 60% dataset it is generally non-significant although very close to the boundary (correlation *p*-value near 0.05) and for the 40% it is never significant. In both cases, the standard deviation increases when the percentage of data-points decreases. These takeaways have similar trends for the analysis using Wilcoxon test. For the case of the distance to the original fit line (with the 100% of data-points), the mean is very similar in all percentages of the dataset, whereas the standard deviation increases for decreasing percentages. Overall,
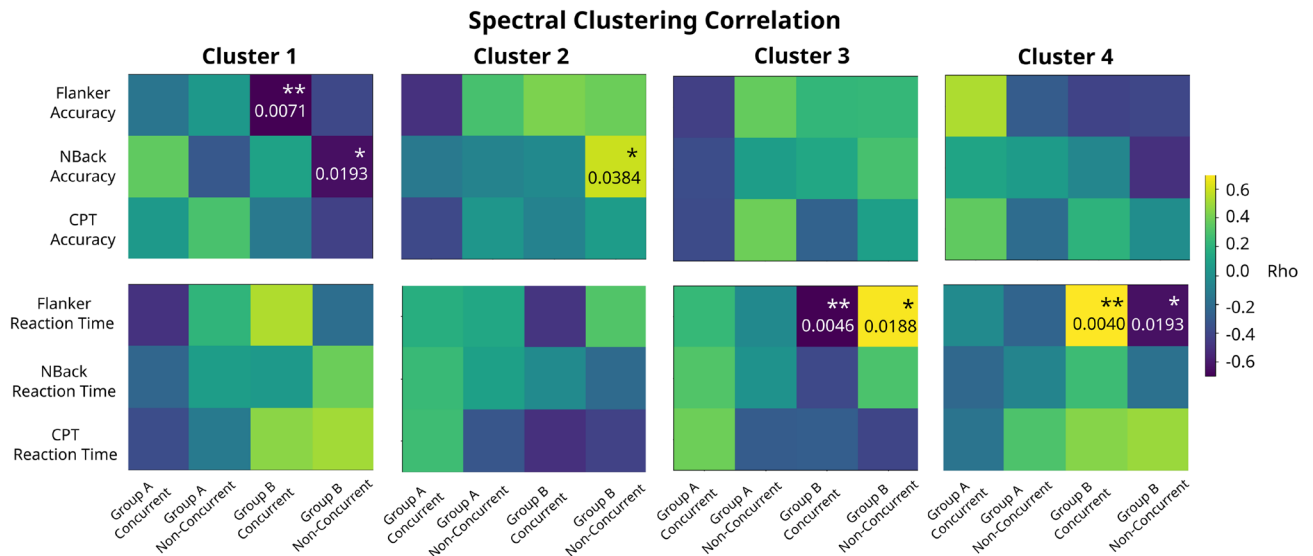

Figure 12. Correlation results for spectral clustering. Each matrix row corresponds to a specific behavioral task endpoint, and each matrix column corresponds to a particular treatment protocol. The colour bar represents the sign of the correlation (positive or negative) and *p*-values are written in the significant cells. For Group B, 24 corrections are done; for Group A, 20 corrections are applied.

Cluster	Group A, Concurrent	Group A, Non-concurrent	Group B, Concurrent	Group B, Non-concurrent
1			Flanker accuracy (−, 0.0150)	
2				N-back accuracy (+, 0.0345)
3			Flanker RT (+, 0.0150)	Flanker RT (−, 0.0433)
4			Flanker RT (−, 0.0402)	Flanker RT (+, 0.0345)

Table 1. Spectral significant responders. These results correspond to significant correlations and significant Wilcoxon tests. In parenthesis are: (1) type of response (positive or negative), (2) Wilcoxon *p*-values. There are significant correlations only in Group B, therefore for Wilcoxon Test 4 corrections are done for Non-concurrent Group and 3 for Concurrent Group. Stimulation protocol Group B (rIFG), offline mode (non-concurrent task) presents positive response in Clusters 2 and 4 measured with N-Back Accuracy and Flanker RT respectively, and negative response in Cluster 3 assessed with Flanker RT. Stimulation protocol B (rIFG), active mode (with Flanker concurrent task) presents negative response in Clusters 1 and 4 measured with Flanker Accuracy and Flanker RT respectively, and positive response in Cluster 3 assessed with Flanker RT.

results show that our correlation method is robust and that the sample size used is the minimum required for obtaining these results.

Since previous work state that subjects who better respond to treatment correspond to those who start in worst conditions[26,47,48], an extra analysis was done in order to discard any effect of this factor on the observed differential response. Hence we further studied whether there is a relationship between baseline behavioral performance (as measured in the first session at T3) and treatment response. The following Fig. 13 shows the behavioral response for the baseline (T3) of the main endpoints (Flanker Accuracy and N-Back Accuracy) as well as for the RT of those tasks. No cluster structure can be seen on the behavioral metric at baseline. Therefore we can state that the cluster structure is solely due to homogeneous digital electrophysiological profiles and not due to a differential baseline response to the tasks.

Discussion

Our work provides a novel approach to study the effects of multichannel anodal tDCS treatment in healthy children and adolescents with different specific brain profiles. This is of interest given the limitations of existent treatments in pediatric clinical populations, e.g. ADHD and ASD, for which new treatments with no side effects[49] are being developed. The unsupervised learning approach we propose constitutes a good alternative to the use of one-model-fit-all approaches based on supervised learning, whose issues to predict behavioral response have been recently analyzed[50]. Furthermore the prognostic value of unsupervised approaches can be exploited in the context of the development of reliable and well-defined personalized protocols. Given that a significant positive response to treatment is not always guaranteed, the proposed subject stratification approach can be of enormous value for the cost-effective development of new treatments.

By grouping experimental subjects in clusters of typical versus non-typical digital EEG phenotypes, a differential response can be found in subjects with a non-typical electrophysiological profile. EEG profiles (summarized in Fig. 14) are characterized only by homogeneous resting-state EEG spectral features before each intervention. The clustering does not take into account the behavioral response to tasks. This confirms our hypothesis that the electrophysiological profile can be used as a proxy of different biophysical characteristics of the brain in experimental subjects. These characteristics are almost identical at 4 different measurement time points each taken a week apart as shown in the intra-subject variability analysis. Therefore we can state that the electrophysiological profiles are subject-specific and do not correspond to a difference in brain state. This is an important question given that brain state is known to influence as well the response to Non-Invasive Brain Stimulation[51]. We propose the use of the stratified electrophysiological responses as shown in this article as digital EEG phenotypes for the characterization of tDCS response.

It can be observed that Group B (i.e. receiving the stimulation over the rIFG brain area) is the only one with responders (both positive and negative), and Group A does not have any significant correlations supported with significant Wilcoxon-test results. This suggests an advantage in terms of post-tDCS task performance in stimulating the rIFG over the more traditional approach of lDLPFC stimulation, even when participant stratification is applied. In almost all groups of participants, the behavioral task showing changes is the Flanker task, when the applied stimulation montage targets enhancing the activity of the rIFG. This result is consistent with the current understanding of the role of the rIFG, associated with response inhibition and behavioral control[52,53].

Future work could be done in order to establish a relationship between the electrophysiological profiles and functional outcomes. A hypothesis would be based on the transfer of stimulation effects from the stimulated working memory network components involved in the Flanker task to the working memory network components involved in the N-Back task[36,54–56]. There might be a response dependency due to the brain's stage of development since, for example, Clusters 3 and 4 present different neurophysiological characteristics and show opposite results[36,57].

In summary, we can state that the effect of brain stimulation (and its parameters including location and presence/absence of concurrent task during the protocol), measured by different behavioral measures, depends on the individual digital EEG phenotype of the subjects undergoing the stimulation intervention. Such a digital
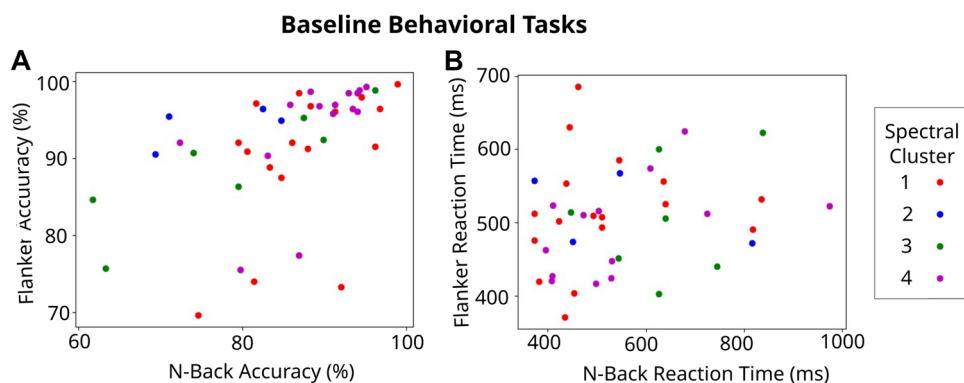

Figure 13. Baseline behavioral tasks distinguished by Spectral Algorithm clustering for (**A**) accuracy and (**B**) reaction time.
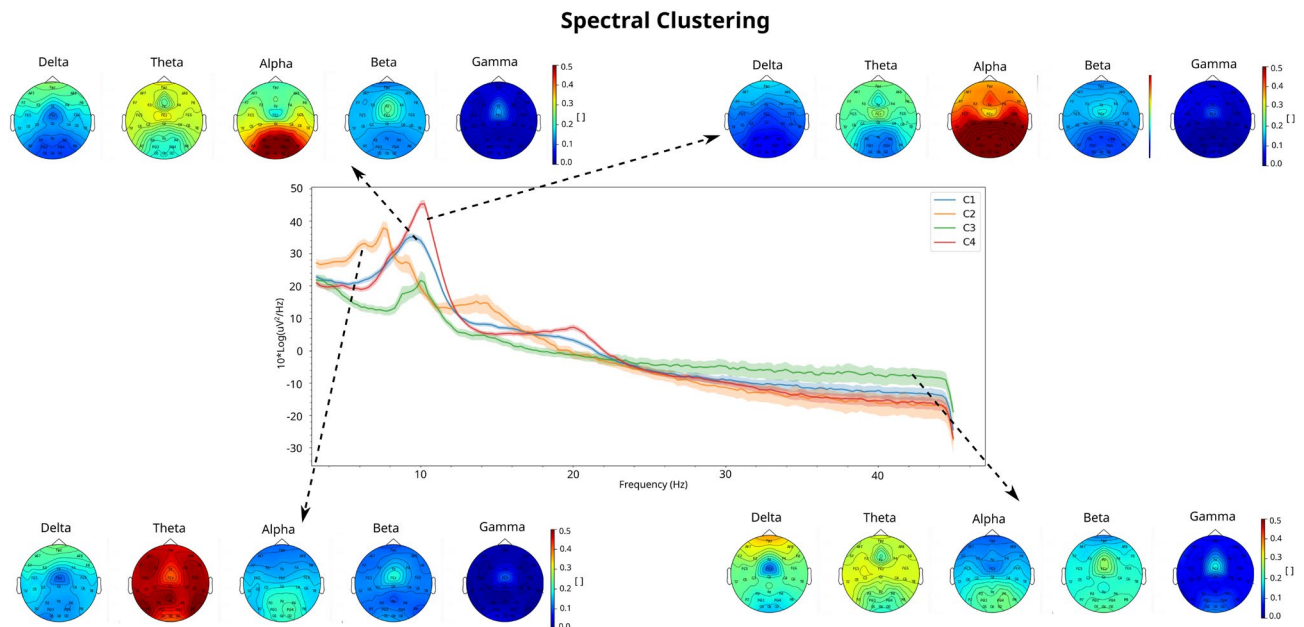
Figure 14. Summary of EEG prototypes. Topoplots and PSD for Clusters that present a significant response in Spectral Algorithm. Cluster 1 (blue PSD) has a negative response to rIFG stimulation with Flanker concurrent task, measured with Flanker Accuracy. Cluster 2 (orange PSD) has a positive response to rIFG stimulation with no concurrent task, measured with N-Back Accuracy. Cluster 3 (green PSD) has a positive response to rIFG stimulation with Flanker concurrent task, measured with Flanker RT; and a negative response to rIFG stimulation with no concurrent task, measured with Flanker RT. Cluster 4 (red PSD) has a positive response to rIFG stimulation with no concurrent task, measured with Flanker RT; and a negative response to rIFG stimulation with Flanker concurrent task, measured with Flanker RT.

phenotype can be derived by applying the unsupervised clustering approach that we have developed. The robustness of our overall methodology was validated in 3 independent tests. First, the feature space of all participants was visualized showing that the features of all trials of pre-treatment EEG for each subject lie close to each other. Therefore, their baseline neurophysiological activity was similar over the different trials, which were conducted one week apart. Second, the k-hold-out methodology applied to the clustering pipeline demonstrates the robustness of the clustering algorithms in our specific dataset. This validation let us ensure to a certain extent the generalization capability of the clustering for generating similar digital EEG phenotypes on unseen data. Finally, the M-hold-out validatioin with $M = 1000$ applied to the correlational and Wilcoxon test analysis shows a decreasing trend with decreasing sample sizes. Moreover it shows also that the correlational analysis and the behavioral response differences are robust enough to resist a reduction of 20% in the data set sample size.

Some limitations arise from the fact that the dataset is relatively small compared to the large amount of features analysed, resulting in a very high dimensional space for the employed clustering validity measures. This hinders to a certain extent the full automation of the methodology. In order to reduce the feature space dimensionality, Principal Feature Analysis[58], permutation analysis[20] and other feature selection algorithms[59,60] could be explored to reduce the number of features. However the application of such approaches would decrease the explainability of the resulting methodology, i.e. digital phenotypes could not be given in the form of a complete electrophysiological characterization.

Conclusions

The advanced clustering methodology we developed shows promising results in the context of broad and heterogeneous EEG signals[61]. Moreover the use of unsupervised learning approaches allows to obtain an actual data-driven stratification, since the learning procedure is freely run without the constraints imposed by the ground truth. This is an important factor given the large misdiagnosis rates in the target diseases[62,63], especially important in ADHD[63]. In case of having used a supervised approach, a false ground truth might have driven the class structure wrongly and hinder the accuracy of the stratification.

The stimulation target areas studied were the lDLPFC and rIFG, paired or not with concurrent tasks N-Back and Flanker Tasks respectively. Effectiveness of the treatment was assessed by behavioral reaction times and accuracy, when performing experimental tasks after the tDCS treatment. The digital phenotypes obtained from pre-treatment EEG, as computed by the clustering procedure described herein, can then be used to predict the response of individuals to particular stimulation protocols. We have successfully developed a procedure for labelling clusters in terms of their behavioral response in the different tasks. Moreover, a rank-sum Wilcoxon test is successfully applied for validation and consistency of results.

Interestingly, positive responder groups correspond to participants with a non-typical brain profile, presenting either a slowing of the EEG or a spread of increased alpha rhythm over frontal areas. This finding is extremely interesting from a neurophysiological point of view. However, we focus in this communication on the methodological novelty of our work. A larger cohort of participants would be needed for confirming the clinical utility of the proposed approach, which we aim to achieve in future communications.

Data availibility

References

1. Ruffini, G. et al. Transcranial current brain stimulation (tcs): Models and technologies. *IEEE Trans. Neural Syst. Rehabil. Eng.* **21**, 333–345. https://doi.org/10.1109/TNSRE.2012.2200046 (2013).
2. Knotkova, H., Nitsche, M., Bikson, W. & Woods, A. J. *Practical Guide to Transcranial Direct Current Stimulation: Principles, Procedures and Applications* (Springer, 2019).
3. Lefaucheur, J.-P. et al. Evidence-based guidelines on the therapeutic use of transcranial direct current stimulation (tdcs). *Clin. Neurophysiol.* **128**, 56–92. https://doi.org/10.1016/j.clinph.2016.10.087 (2016).
4. Antal, A. et al. Low intensity transcranial electric stimulation: Safety, ethical, legal regulatory and application guidelines. *Clin. Neurophysiol.* **128**, 1774–1809. https://doi.org/10.1016/j.clinph.2017.06.001 (2017).
5. Friel, K., Gordon, A., Carmel, J., Kirton, A. & Gillick, B. Pediatric issues in neuromodulation: safety, tolerability and ethical considerations. In *Pediatric Brain Stimulation*, 131–149, https://doi.org/10.1016/B978-0-12-802001-2.00007-2 (Elsevier, 2016).
6. Salehinejad, M. A., Vicario, C., Vila-Rodriguez, F. & Cohen Kadosh, R. *tDCS in Child and Adolescent Psychiatry* (Springer, 2021).
7. Faraone, S. V. et al. The world federation of adhd international consensus statement: 208 evidence-based conclusions about the disorder. *Elsevier* **128**, 789–818. https://doi.org/10.1016/j.neubiorev.2021.01.022 (2021).
8. Snyder, M. et al. Integration of an eeg biomarker with a clinician's adhd evaluation. *Brain Behav.* https://doi.org/10.1002/brb3.330 (2015).
9. Ogrim, G., Kropotov, J. & Hestad, K. The qeeg theta/beta ratio in adhd and normal controls: Sensitivity, specificity, and behavioral correlates. *Psychiatry Res.* **198**, 482–8. https://doi.org/10.1016/j.psychres.2011.12.041 (2012).
10. Ibañez Soria, D. et al. Resting eyes closed beta-phase high gamma-amplitude coupling deficits in children with attention deficit hyperactivity disorder. *Biorxiv* https://doi.org/10.1016/j.neubiorev.2021.01.022 (2019).
11. Krepel, N. et al. A multicenter effectiveness trial of qeeg-informed neurofeedback in adhd: Replication and treatment prediction. *Neuroimage Clin.* **28**, 102399. https://doi.org/10.1016/j.nicl.2020.102399 (2020).
12. Schneider, B. & Enenbach, M. Managing the risks of adhd treatments. *Curr. Psychiatry Rep.* **16**, 479. https://doi.org/10.1007/s11920-014-0479-3 (2014).
13. Eckerl, C., Spooren, W. & Murphy, D. G. Developing new pharmacotherapies for autism. *J. Intern. Med.* **274**, 308–20. https://doi.org/10.1111/joim.12113 (2013).
14. Salehinejad, M. A., Wischnewski, M., Nejati, V., Vicario, C. M. & Nitsche, M. A. Transcranial direct current stimulation in attention-deficit hyperactivity disorder: A meta-analysis of neuropsychological deficits. *PLoS ONE* **14**, e0215095 (2019).
15. Luque-Casado, A., Rodríguez-Freiría, R., Fogelson, N., Iglesias-Soler, E. & Fernández-del Olmo, M. An integrative clustering approach to tdcs individual response variability in cognitive performance: Beyond a null effect on working memory. *Neuroscience* **443**, 120–130 (2020).
16. Albizu, A. et al. Machine learning and individual variability in electric field characteristics predict tdcs treatment response. *Elsevier* **13**, 1753–1764. https://doi.org/10.1016/j.brs.2020.10.001 (2020).
17. Lopez, C., Tucker, S., Salameh, T. & Tucker, C. An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *J. Biomed. Inform.* **85**, 30–39. https://doi.org/10.1016/j.jbi.2018.07.004 (2018).
18. Kueffner, R. et al. Stratification of amyotrophic lateral sclerosis patients: A crowdsourcing approach. *Sci. Rep.* **9**, 690. https://doi.org/10.1038/s41598-018-36873-4 (2019).
19. Hawco, C., Dickie, E. W., Jacobs, G., Daskalakis, Z. J. & Voineskos, A. N. Moving beyond the mean: Subgroups and dimensions of brain activity and cognitive performance across domains. *Psychiatry Res.* https://doi.org/10.1016/j.neuroimage.2021.117823 (2021).
20. Kambeitz, J. et al. Clinical patterns differentially predict response to transcranial direct current stimulation (tdcs) and escitalopram in major depression: A machine learning analysis of the elect-tdcs study. *J. Affect. Disord.* **265**, 460–467. https://doi.org/10.1016/j.jad.2020.01.118 (2020).
21. Wu, W. et al. An electroencephalographic signature predicts antidepressant response in major depression. *Nat. Biotechnol.* **38**, 439–447. https://doi.org/10.1038/s41587-019-0397-3 (2020).
22. Roelofs, C. L. et al. Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the icon-db consortium. *Clin. Neurophysiol.* **132**, 643–649. https://doi.org/10.1016/j.clinph.2020.10.017 (2021).
23. Nord, C. L. et al. Neural predictors of treatment response to brain stimulation and psychological therapy in depression: a double-blind randomized controlled trial. *Neuropsychopharmacology* **44**, 1613–1622. https://doi.org/10.1038/s41386-019-0401-0 (2019).
24. Torous, J., Onnela, J. & Keshavan, M. New dimensions and new tools to realize the potential of rdoc: digital phenotyping via smartphones and connected devices. *Transl. Psychiatry* **7**, e1053–e1053 (2017).
25. Martin, D. M. et al. Can transcranial direct current stimulation enhance outcomes from cognitive training? a randomized controlled trial in healthy participants. *Int. J. Neuropsychopharmacol.* **16**, 1927–36. https://doi.org/10.1017/S1461145713000539 (2013).
26. Hill, A. T., Rogasch, N. C., Fitzgeralda, P. B. & Hoya, K. E. Impact of concurrent task performance on transcranial direct current stimulation (tdcs)-induced changes in cortical physiology and working memory. *Cortex* **12**, 447. https://doi.org/10.1016/j.cortex.2018.11.022 (2019).
27. Saby, J. N. & Marshall, P. J. The utility of eeg band power analysis in the study of infancy and early childhood. *Dev. Neuropsychol.* **37**, 253–73. https://doi.org/10.1080/87565641.2011.614663 (2012).
28. Clarke, A. R. et al. Coherence in children with attention-deficit/hyperactivity disorder and excess beta activity in their eeg. *Clin. Neurophysiol.* **118**, 1472–9. https://doi.org/10.1016/j.clinph.2007.04.006 (2007).
29. Jin, H.-K. & Cho, S.-H. Comparison of relative alpha power spectral electroencephalogram activity analysis according to electrical stimulation levels in normal adults. *J. Psychiatry* https://doi.org/10.4172/2378-5756.1000308 (2015).

30. Hess, S., Duivesteijn, W., Honysz, P. & Morik, K. The spectacl of nonconvex clustering: A spectral approach to density-based clustering. *Proc. AAAI* **465**, 3788–3795. https://doi.org/10.1609/aaai.v33i01.33013788 (2019).
31. Ross, T. J. *Fuzzy Logic with Engineering Applications* (Wiley, 2010).
32. Rudolf, H. *CFT 20 R: Grundintelligenztest Skala 2 - Revision* (Hogref, 2006).
33. Döpfner, M., Plück, J. & Kinnen, C. *Arbeitsgruppe Deutsche Child Behavior Checklist* (Hogrefe, Göttingen, 2014).
34. Katharina, P., Manfred, D. & Gerd, L. Diagnostik-system für psychische störungen im kindes- und jugendalter nach icd-10 und dsm-iv (disyps-kj). *Diagnostica* **46**, 165–166. https://doi.org/10.1026//0012-1924.46.3.165 (2000).
35. Constantino, J. N. *Social Responsiveness Scale* 2919–2929 (Springer, 2013).
36. Splittgerber, M. et al. Multichannel anodal tdcs over the left dorsolateral prefrontal cortex in a paediatric population. *Sci. Rep.* **11**, 1–15. https://doi.org/10.1038/s41598-021-00933-z (2021).
37. Merschformann, C. et al. P 16. multichannel anodal transcranial direct current stimulation over the right inferior frontal gyrus in pediatric population: preliminary behavioural results. *Clin. Neurophysiol.* **132**, e7–e8. https://doi.org/10.1016/j.clinph.2021.02.338 (2021).
38. Eriksen, B., Riksen, C. & Hoffman, J. Recognition memory and attentional selection: Serial scanning is not enough. *J. Exp. Psychol. Hum. Percept. Perform.* **12**, 476–483. https://doi.org/10.1037//0096-1523.12.4.476 (1986).
39. Kirchner, W. Age differences in short-term retention of rapidly changing information. *Exp. Psychol.* **55**, 352–8. https://doi.org/10.1037/h0043688 (1958).
40. Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D. & Beck, L. H. A continuous performance test of brain damage. *J. Consult. Psychol.* **20**, 343–50. https://doi.org/10.1037/h0043220 (1956).
41. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
42. Rui, X. & Donald, C. W. *Clustering* (Wiley, 2008).
43. Holt Skinner. Implementation of fuzzy c-means and possibilistic c-means. https://github.com/holtskinner/PossibilisticCMeans (2018). Accessed: 2022-12-25.
44. Nguyen, L. H. & Holmes, S. Ten. quick tips for effective dimensionality reduction. *Plos Comput. Biol.* https://doi.org/10.1371/journal.pcbi.1006907 *(2011)*.
45. Korthauer, K. et al. A practical guide to methods controlling false discoveries in computational biology. *Genome Biol.* https://doi.org/10.1186/s13059-019-1716-1 *(2019)*.
46. Liu, Y., Li, Z., Xiong, H., Gao, X. & Wu, J. Understanding of internal clustering validation measures. *IEEE International Conference on Data Mining* 911–916, https://doi.org/10.1109/ICDM.2010.35 (2010).
47. Ruf, S. P., Fallgatter, A. J. & Plewnia, C. Augmentation of working memory training by transcranial direct current stimulation (tdcs). *Sci. Rep.* https://doi.org/10.1038/s41598-017-01055-1 *(2017)*.
48. Gözenman, F. & Berryhil, M. E. Working memory capacity differentially influences responses to tdcs and hd-tdcs in a retro-cue task. *Neurosci. Lett.* **629**, 105–109. https://doi.org/10.1016/j.neulet.2016.06.056 (2016).
49. Graham, J. & Coghill, D. Adverse effects of pharmacotherapies for attention-deficit hyperactivity disorder. *CNS Drugs* **22**, 213–37. https://doi.org/10.2165/00023210-200822030-00003 (2008).
50. Greene, A. S. et al. Brain-phenotype models fail for individuals who defy sample stereotypes. *Nature* **609**, 109–118 (2022).
51. Silvanto, J., Muggleton, N. & Walsh, V. State-dependency in brain stimulation studies of perception and cognition. *Trends Cogn. Sci.* **12**, 447–454 (2008).
52. Schaum, M. et al. Right inferior frontal gyrus implements motor inhibitory control via beta-band oscillations in humans. *eLife* https://doi.org/10.1016/j.tics.2013.12.003 *(2021)*.
53. Aron, A., Robbins, T. & Poldrack, R. Inhibition and the right inferior frontal cortex: One decade on. *Trends Cogn. Sci.* **18**, 177–85. https://doi.org/10.1016/j.tics.2013.12.003 (2014).
54. Collette, F. & Van der Linden, M. Brain imaging of the central executive component of working memory. *Neurosci. Biobehav. Rev.* **26**, 105–25. https://doi.org/10.1016/s0149-7634(01)00063-x (2002).
55. Scharinger, C., Soutschek, A., Schubert, T. & Gerjets, P. When flanker meets the n-back: What eeg and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology* **52**, 1293–304. https://doi.org/10.1111/psyp.12500 (2015).
56. Szmalec, A., Verbruggen, F., Vandierendonck, A. & Kemps, E. Control of interference during working memory updating. *J. Exp. Psychol. Hum. Percept. Perform.* **37**, 137–51. https://doi.org/10.1037/a0020365 (2011).
57. Hsu, T.-Y., Chi-Heng, J. & Tseng, P. Individual differences and state-dependent responses in transcranial direct current stimulation. *Front. Hum. Neurosci.* https://doi.org/10.3389/fnhum.2016.00643 *(2016)*.
58. Lu, Y., Cohen, I., Sean Zhou, X. & Tian, Q. Feature selection using principal feature analysis. *Proceedings of the 15th International Conference on Multimedia 2007* 23–29, https://doi.org/10.1145/1291233.1291297 (2007).
59. Zhao, Z. & Liu, H. Spectral feature selection for supervised and unsupervised learning. *ICML* https://doi.org/10.1145/1273496.1273641 *(2007)*.
60. Vidyasagar, M. Identifying predictive features in drug response using machine learning: Opportunities and challenges. *Annu. Rev. Pharmacol. Toxicol.* **55**, 15–34. https://doi.org/10.1146/annurev-pharmtox-010814-124502 (2015).
61. Sarishvili, A., Winter, J., Luhmann, H. & Mildenberger, E. Probabilistic graphical model identifies clusters of eeg patterns in recordings from neonates. *Clin. Neurophysiol.* **130**, 1342–1350. https://doi.org/10.1016/j.clinph.2019.04.708 (2019).
62. Kazda, L. et al. Overdiagnosis of attention-deficit/hyperactivity disorder in children and adolescents. *JAMA Netw. Open* **4**, e215335. https://doi.org/10.1001/jamanetworkopen.2021.5335 (2021).
63. Ford-Jones, P. C. Misdiagnosis of attention deficit hyperactivity disorder: "Normal behaviour" and relative maturity. *Paediatr. Child Health* **20**, 200–202. https://doi.org/10.1093/pch/20.4.200 (2015).

Acknowledgements

Author contributions

Experimental Study concept and design: M.S., K.K., V.M. Conduction of experiment and Data collection: M.S., V.M. Methodology concept: A.S.F. Analysis: P.D., A.S.F., E.K., C.B. Interpretation of results: P.D., A.S.F., E.K., C.B. Analysis supervision: A.S.F., E.K., C.B. Drafting of the manuscript: P.D., A.S.F., E.K., C.B. All authors reviewed the manuscript.

Competing interests

Neuroelectrics is the company that manufactures the Starstim devices used for stimulation and recording of EEG analysed in this work. We, all authors, declare there are no conflict of interests with this publication and we have all approved this final version.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-34724-5.

Correspondence and requests for materials should be addressed to A.S.-F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.