IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 5, 2018, accepted September 14, 2018, date of publication September 28, 2018, date of current version October 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2872765

Phenotyping Ex-Combatants From EEG Scalp Connectivity

ANDRÉS QUINTERO-ZEA[1], JOSÉ D. LÓPEZ[1], KEITH SMITH[2,3], NATALIA TRUJILLO[4,5], MARIO A. PARRA[6,7], AND JAVIER ESCUDERO[3], (Member, IEEE)

[1]SISTEMIC, Facultad de Ingeniería, Universidad de Antioquia, Medellín 050010, Colombia
[2]Usher Institute for Population Health Science and Informatics, Medical School, The University of Edinburgh, Edinburgh EH16 4UX, U.K.
[3]School of Engineering, Institute for Digital Communications, The University of Edinburgh, Edinburgh EH9 3FB, U.K.
[4]GISAME, Facultad Nacional de Salud Pública, Universidad de Antioquia, Medellín 050010, Colombia
[5]Neuroscience Group, Universidad de Antioquia, Medellín 050010, Colombia
[6]School of Social Sciences, Psychology Department, Heriot-Watt University, Edinburgh EH14 4AS, U.K.
[7]Universidad Autónoma del Caribe, Barranquilla 080020, Colombia

Corresponding author: Andrés Quintero-Zea (andres.quintero@udea.edu.co)

This work was supported in part by Colciencias under Grant 122266140116 and Grant 111577757638, in part by CODI-UDEA under Grant INV518-16, and in part by the Newton-Caldas Fund under Grant BC027-EDU2016. The work of A. QUINTERO-ZEA was supported by the Colciencias Doctoral Fellowship Call 647 (2014). The work of K. Smith was supported by an HDRUK UKRI Innovation Fellowship under Grant Mr/S004122/1.

ABSTRACT Being involved in war experiences may have severe consequences in mental health. This exposure has been associated in Colombian ex-combatants with risk of proactive aggression modulating emotional processing. However, the extent of the cognitive processes underlying aggressive behavior is still an open issue. In this paper, we propose a support vector machine-based processing pipeline to identify different cognitive phenotypes associated with atypical emotional processing, based on canonical correlation analysis of EEG network features, and cognitive and behavioral evaluations. Results show the existence of cognitive phenotypes associated with differences in the mean value of leaf fraction and diameter of EEG networks across groups. The ability of identifying phenotypes in these otherwise healthy subjects opens up the possibility to aid in the development of specific interventions aimed to reduce expression of proactive aggression in ex-combatants and assessing the efficacy of such interventions.

INDEX TERMS EEG phenotypes, ex-combatants, functional connectivity, graph theory, machine learning, scalp EEG.

I. INTRODUCTION

Colombia is an upper-middle-income country that has been challenged by the adverse effects of internal conflict for nearly six decades [1]. Recently, the Reincorporation and Normalization Agency (ARN)[1] has been leading disarmament, demobilization, and reintegration (DDR) programs for illegal armed groups in Colombia [2]. Ex-combatants involved in war experiences have been exposed to numerous forms of extreme violence [3]. This exposure has been associated with risk of aggressive behavior, especially proactive aggression [4]. Proactive aggression traits have been identified to be highly correlated with recidivism of ex-combatants, and might affect their experience of reintegration into Colombian society [5], [6].

Moreover, this aggressive behavior generally underlies atypical emotional processing (EP) [7]–[10]. EP is crucial for an adequate interpersonal functioning, and it is related to the ability of understanding emotional information conveyed by stimuli or situations, including facial expressions and words with emotional content (faces and words from now on) [11]. Research into EP remains an open issue given the lack of markers applied to ex-combatants studies.

Abnormalities in EP can be assessed using psychological evaluation such as rating scales and experimental tasks in populations with neurological and psychiatric disorders [12], [13]. However, these studies do not provide accurate knowledge about functional reorganization in the brain, *i.e.* changes in neurophysiological functioning elicited by such disorders, and their contribution to atypical EP. EP can be estimated using task events synchronized with neural acquisition techniques such as electroencephalography (EEG) and magnetoencephalography (MEG). EEG and MEG are preferred over other functional imaging techniques (fMRI, PET, etc.), because several studies have evidenced that emotional stimuli

[1]http://www.reintegracion.gov.co/en

are recognized and differentiated within the first 200–250 ms after their presentation, and fMRI is known to have a poor temporal resolution [14].

EEG is a noninvasive, portable, and low-cost brain-imaging technique widely used to characterize brain function in healthy subjects as well as in neurological and psychiatric patients with high temporal resolution. EEG analysis aims to extract spatio-temporal information, such as functional connectivity (FC) patterns. Particularly, a FC pattern reveals the synchronization of distant neural areas or regions via correlations and dependencies in neurophysiological measures of brain activity [15] and can be characterized using a set of neurobiologically meaningful network metrics [16]. If an EEG pattern is consistent within a behavioral homogeneous group, it can be interpreted as a neurophysiological marker or phenotype that allows characterizing the majority of the population under experimental circumstances [17], [18].

Previous studies showed that EEG phenotypes can be used to identify functional impairments across a wide variety of clinical conditions such as Alzheimer's disease [19], attention deficit hyperactivity disorder [20], impulse control disorders [21], and post-traumatic stress disorder [22]. For EP impairments, applications in patients with depression [23] and schizophrenia [24] show the high interest of phenotyping. Although these studies provide evidence to consider phenotyping schemes as powerful tools to assess neurophysiological impairments, EP phenotypes have yet to be considered in non-clinical populations such as ex-combatants.

We hypothesize that a joint phenotyping scheme using both psychological and neurophysiological measures may improve the EP characterization in Colombian ex-combatants. They represent a non-clinical group with well-documented atypical EP, and investigating their cognitive phenotype may provide insights into how combat experience affects EP. Such a phenotyping scheme would be useful to design psycho-therapeutic interventions to address solutions to the proactive aggression drivers in ex-combatants, and to help them towards their reintegration process.

In this paper, we introduce a methodology that identifies potential neurophysiological phenotypes related to EP of faces and words in ex-combatants. First, we use graph theory to characterize the EEG-FC and to identify changes in the ex-combatants' functional connectivity network (FCN), in comparison with a non-combat-experienced control group. Second, we use canonical correlation analysis to assess the relation between the network analysis, and social cognition and behavior (SCB) evaluations. Finally, a support vector machine (SVM) is used to obtain class representative patterns. Such patterns were used to empirically identify potential phenotypes.

This paper is structured as follows. Section II presents the background and details of the proposed methodology. Section III presents the experimental results. Finally, discussion and conclusion are presented in Sections IV and V.

II. MATERIALS AND METHODS

Fig. 1 shows the pipeline diagram of the proposed methodology for phenotyping ex-combatants using EEG-FC. Details of each stage are presented in the following sections.

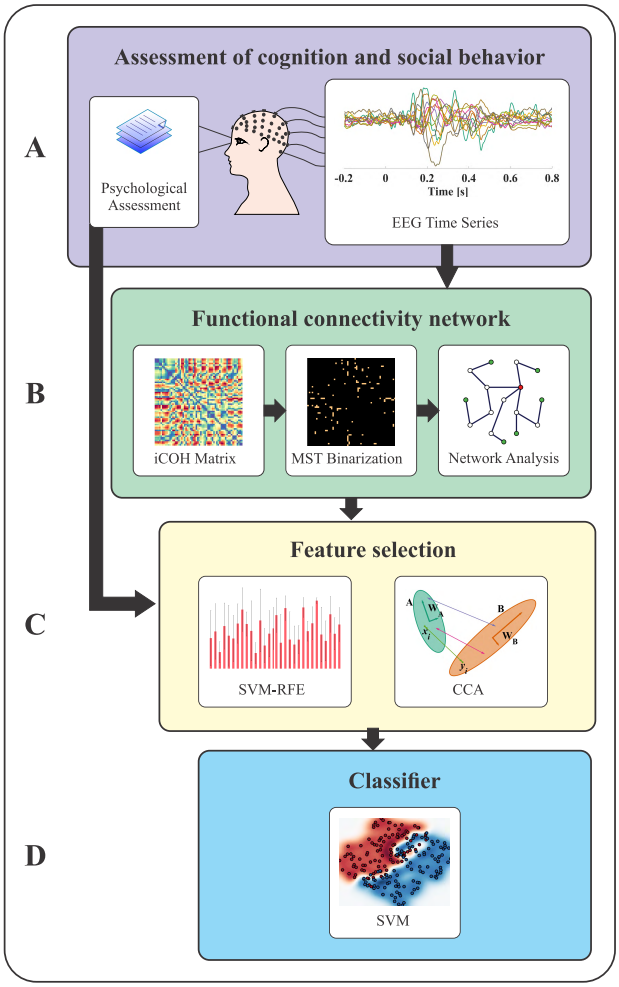

FIGURE 1. Diagram of the analysis pipeline. (A) EEG recordings and psychological assessments were acquired from each subject. (B) EEG time series were used to obtain the imaginary part of coherency (iCOH) matrix for each subject. iCOH matrices were binarized by means of minimum spanning tree (MST). With this procedure only the stronger links constituted the brain graph. The topology of the brain network was quantified by graph metrics. (C) A feature selection scheme was applied to graph metrics. Further, canonical correlation analysis (CCA) was performed over the top FCN features and SCB scores. (D) The resultant discriminant features from CCA were used as the input features in the SVM classifier. The output of the SVM was used to identify the potential phenotypes.

A. PARTICIPANTS

The sample consisted of $N_p = 50$ participants. Of these, thirty were Colombian ex-combatants (two female) involved in the DDR program of the ARN, and 20 Colombian citizens without combat experience (paired by sex, age and school level). Volunteers that informed having confirmed psychiatric and neurological disorders were excluded from the study. All subjects participated voluntarily and signed an informed consent in agreement with the Helsinki declaration. The research

protocol was approved by ethical committee of University of Antioquia (Medellín, Colombia). Demographic information is provided in Table 1. Age and education level are given in years as mean (M) ± standard deviation (SD) and evaluated with Student's t-test. Gender is given as female to male proportion and evaluated with χ^2.

TABLE 1. Demographic information in ex-combatant and civilian participants.

	Ex-combatants $n = 30$	Civilians $n = 20$	
	$M \pm SD$	$M \pm SD$	$t/Chi2$ (p)
Age (Years)	37.50±8.22	36.15±9.17	0.543 (0.589)
Education (Years)	10.23±3.02	11.05±2.14	-1.118 (0.269)
Gender (Female:Male)	2:28	2:18	0.181 (0.670)

B. ASSESSMENT OF COGNITION AND SOCIAL BEHAVIOR

1) EXPERIMENTAL SETUP

We implemented an emotion recognition task (ERT) [8] which requires identifying faces and words with emotional content. The stimuli consisted of 90 pictures of female and male faces taken from the MMI Facial Expression Database [25]. Additionally, 90 words were selected from the linguistic corpus generated by the communications department at University of Antioquia (Medellín, Colombia) [26]. Fig. 2 shows the task sequence of a single trial. Each trial started with a fixation cross presented for 1000 ms followed by the stimulus display presented for 200 ms. Following face or word presentation, participants had to categorize the valence of the stimulus displayed on the computer screen into one of three response categories (positive, negative or neutral). Correct responses were followed by a inter stimulus interval (ISI) with a random duration between 700 and 1000 ms. A red X letter is shown as feedback in error trials, followed by the described ISI. Further details about the experiment protocol have been previously described [7], [8].

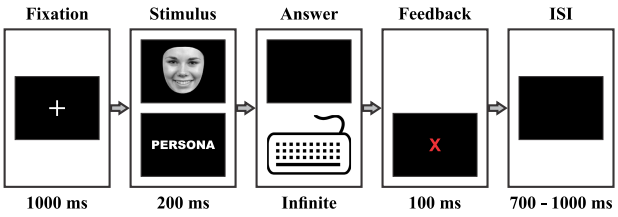

FIGURE 2. Example trial of the Emotion Recognition Task.

2) EEG DATA

EEG signals were sampled at 1000 Hz from a 64-channel Neuroscan SynAmps2 amplifier and band-pass filtered between 0.1 and 30 Hz. We selected 60 electrodes for this study (see Fig. 3 for the schematic of electrode placement. HEO, VEO, CB1 and CB2 were excluded as they do not record neural activity). Continuous EEG data were segmented from −200 ms prior to the stimulus to 800 ms after it.

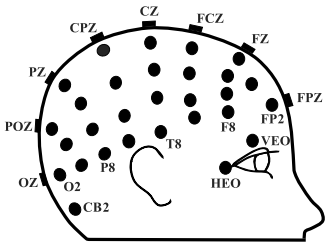

FIGURE 3. Schematic of electrode placement based on the international 10-20 system. Right lateral view of the placement of 60 electrodes for brain activity registration, two cerebellar electrodes (CB1 –left side– and CB2) and two EOG electrodes (HEO and VEO).

Epochs were baseline corrected using the −200 to 0 ms window, downsampled to 500 Hz, and offline re-referenced to average. Bad trials were rejected based on visual inspection, no more than 5% of the trials were marked as bad for each subject. Electrooculography (EOG) artifacts were removed using Independent Component Analysis (ICA) and visual inspection. This preprocessing stage was performed using FieldTrip Toolbox [27] for MATLAB.

3) PSYCHOLOGICAL TESTS

Both ex-combatants and civilians completed an evaluation that included two psychological tests. First, the social ability scale of Gismero (EHS) [28]. The EHS is a 33-item self-report instrument that inquiries individuals about their ability to interact with others in a variety of situations. Larger values of this score suggest reduced social assertion. Second, the reactive-proactive aggression questionnaire (RPQ) [29]. The RPQ is a valid 23-item questionnaire designed to measure reactive and proactive aggression. Higher scores obtained on each dimension indicate higher levels of aggression.

C. FUNCTIONAL CONNECTIVITY NETWORK

For each subject, we estimated the FCN by computing the imaginary part of coherency (iCOH) between each pair of EEG preprocessed time series, $\mathbf{Z} \in \mathbb{R}^{N_c \times N_t}$, where N_c is the number of channels and N_t is the number of time samples. We selected iCOH as the functional connectivity measure to reduce the volume conduction effects [30].

Coherency is a measure of the linear relation between two EEG channels at a specific frequency [31]. For two different time series of channels i and j, the coherency at each frequency (f) is defined as:

$$\mathrm{COH}_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)\, S_{jj}(f)}}, \tag{1}$$

where $S_{ij}(f)$ is the cross spectrum of the signals acquired from channels i and j, while $S_{ii}(f)$ and $S_{jj}(f)$ are the respective autospectra for a given frequency f. From this formula we are only interested in the absolute value of the imaginary part of the resultant coherency value.

From a physiological point of view, EEG coherency reflects functional interaction among the brain areas under

study [32]. Therefore, higher values of coherency denote that two brain regions are working synchronized at a specific frequency.

For each of the four independent EEG frequency bands selected for this analysis: delta (0.1–4 Hz), theta (4–7 Hz), alpha (7–13 Hz), and beta (13–30 Hz); we calculated the spectrum of each data channel. As cross-spectral quantities can be readily computed using multitaper methods [33], we used the multitaper method based on discrete prolate spheroidal sequences as tapers to estimate spectra (see [27] for further details). These spectra were then used to calculate iCOH between all pairs of data channels within each task (faces and words). Both spectral and iCOH calculations were performed using FieldTrip [27].

Furthermore, we used Kruskal's algorithm [34] to compute the minimum spanning tree (MST) from every weighted connectivity matrix. The MST is a simple acyclic (i.e., without loops) sub-network of the original weighted network that connects all nodes in the network, while maximizing connectivity within the tree [35]. MST captures important topological information in the original network, and can be used to directly compare networks with the same number of nodes. MSTs can be characterized with a set of relatively simple measures [35]. We used the Brain Connectivity Toolbox [36] to calculate three common metrics for each MST matrix [37], [38]:

- Leaf fraction (LF): Fraction of leaf nodes in the MST network, where a leaf node is defined as a node with degree one. This measure is particularly useful to quantify the extent to which a tree is more pathlike (LF $= 2/N_c$) or more starlike (LF $= 1 - 1/N_c$).
- Maximum degree (MD): Degree of the node with the highest number of links. A high value of MD denotes the existence of hubs within the tree. In a pathlike tree MD is two, while in a star MD is equal to the number of edges ($N_c - 1$).
- Diameter (D): The largest-shortest path length between any two nodes in the graph. The global efficiency in the communication of the tree can be defined to be inversely proportional to D. A pathlike tree has the worst efficiency with a diameter equal to ($N_c - 1$). In the case of a star the diameter is two, leading to a high efficiency.

D. FEATURE SELECTION

1) SUPPORT VECTOR MACHINE RECURSIVE FEATURE ELIMINATION

Support vector machine recursive feature elimination (SVM-RFE) is an embedded feature selection algorithm proposed in [39]. In this algorithm, the features are discarded recursively based on the SVM classifier weights and a rank score list is generated.

The support vector machine (SVM) classifier is a binary classifier algorithm that aims to find an optimal hyperplane to separate the points of two classes [40]. SVM takes a training dataset $\{\mathbf{x}_i, y_i\} \in \mathbb{R}^{N_p} \times \{-1, 1\}$, where $\mathbf{x}_i$ are the training examples and y_i the class labels; and solves the constrained optimization problem:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} y_i y_j \alpha_i \alpha_j \kappa(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j}^{N_p} \alpha_j$$
$$\text{subject to } \sum_{i=1}^{N_p} \alpha_i y_i = 0,$$
$$0 \leq \alpha_i \leq C,\ i = 1, \ldots, N_p, \tag{2}$$

where α_i are Lagrange multipliers, $\kappa(\mathbf{x}_i, \mathbf{x}_j)$ represents a kernel function, and C is the penalty factor that controls the complexity of the SVM. The decision function is:

$$f(\mathbf{x}) = \text{sgn}\left(\sum_{i}^{N_p} \alpha_i y_i \kappa(\mathbf{x}_i, \mathbf{x}) + b \right) \tag{3}$$

For the C-SVM model specification, the Gaussian (or radial basis) kernel is defined as:

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \left\| \mathbf{x}_i - \mathbf{x}_j \right\|^2 \right) \tag{4}$$

The penalty factor C and the parameter γ of the kernel must be tuned by minimizing the estimated generalization error.

As presented in [41], the ranking criterion for feature k is given by:

$$J(k) = \frac{1}{2} \sum_{i,j=1}^{N_p} y_i y_j \alpha_i \alpha_j \kappa(\mathbf{x}_i, \mathbf{x}_j)$$
$$- \frac{1}{2} \sum_{i,j=1}^{N_p} y_i y_j \alpha_i \alpha_j \kappa\left(\mathbf{x}_i^{(-k)}, \mathbf{x}_j^{(-k)} \right), \tag{5}$$

where $\mathbf{x}_i^{(-k)}$ stands for the training example $\mathbf{x}_i$ after removing the k-th feature.

The SVM-RFE algorithm is stated in Algorithm 1.

Algorithm 1 SVM-RFE [39]

Require: Initial feature set, G
Ensure: Rank list, R
  $R \leftarrow \{\}$
  repeat
    Train SVM using G
    Compute and sort the ranking criteria
    Update the ranking list R
    Remove the feature with smallest rank from G
  until G is empty
  return R

We performed SVM-RFE to select the most discriminative subset of features. SVM-RFE was not used to increase classification accuracy or to lead to over-optimistic results, hence accuracies from this method only estimate the ability of the subset of feature to classify into one of the two classes; they do not reflect the overall cross-validation accuracy.

2) CANONICAL CORRELATION ANALYSIS

The reduced feature set given by the top features selected by the SVM-RFE for the FCN metrics ($\mathbf{A} \in \mathbb{R}^{N_p \times N_a}$, with N_a being the cardinality of the selected feature set) and the SCB scores ($\mathbf{B} \in \mathbb{R}^{N_p \times N_b}$, with N_b being the cardinality of the SCB scores) were combined to obtain a single set of features, which was more discriminant than any of the input datasets. This was achieved by using a feature fusion technique based on Canonical Correlation Analysis (CCA) [42].

We computed the canonical correlations of matrices $\mathbf{A}$ and $\mathbf{B}$ using the linear algebraic formulation presented in [43]. We first performed a compact singular value decomposition (SVD) of matrices $\mathbf{A} = \mathbf{U_A}\Sigma_\mathbf{A}\mathbf{V_A}^\top$ and $\mathbf{B} = \mathbf{U_B}\Sigma_\mathbf{B}\mathbf{V_B}^\top$. Then, the canonical projective matrices were computed as $\mathbf{W_A} = \mathbf{V_A}\Sigma_\mathbf{A}^{-1}\mathbf{U}$ and $\mathbf{W_B} = \mathbf{V_B}\Sigma_\mathbf{B}^{-1}\mathbf{V}$, where $\mathbf{U}$ and $\mathbf{V}$ were found as $\mathbf{U_A}^\top\mathbf{U_B} = \mathbf{U}\Sigma\mathbf{V}^\top$.

As defined in [42], a feature-level fusion was performed either by concatenation or summation of the transformed feature spaces $\mathbf{X} = \mathbf{A}\mathbf{W}_A$ and $\mathbf{Y} = \mathbf{B}\mathbf{W}_B$:

$$\mathbf{Z}_1 = \begin{pmatrix}\mathbf{X} & \mathbf{Y}\end{pmatrix} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\end{pmatrix}\begin{pmatrix}\mathbf{W}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_B\end{pmatrix}, \tag{6a}$$

$$\mathbf{Z}_2 = \mathbf{X} + \mathbf{Y} = \mathbf{A}\mathbf{W}_A + \mathbf{B}\mathbf{W}_B, \tag{6b}$$

where $\mathbf{Z}_1 \in \mathbb{R}^{N_p \times 2r}$ and $\mathbf{Z}_2 \in \mathbb{R}^{N_p \times r}$, with $r = \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$, are called the canonical correlation discriminant features (CCDF).

E. CLASSIFIER

After selecting the discrimination features, the final task was to inquire about the existence of cognitive phenotypes. The classifier adopted here was the SVM [40], as it is one of the most robust and accurate classifiers.

As we have an imbalanced dataset, the soft margin SVM classifier in Eq. (2) must be reformulated [44] to account for this imbalance:

$$\begin{aligned}\min_{\alpha}\ & \frac{1}{2}\sum_{i=1}^{N_p}\sum_{j=1}^{N_p} y_i\, y_j\, \alpha_i\, \alpha_j\, \kappa(\mathbf{x}_i, \mathbf{x}_j) - \sum_{j=1}^{N_p}\alpha_j, \\ \text{subject to}\ & \sum_{i=1}^{n}\alpha_i y_i = 0, \\ & 0 \le \alpha_i \le C^+, \quad i \in \mathcal{P}, \\ & 0 \le \alpha_i \le C^-, \quad i \in \mathcal{N},\end{aligned} \tag{7}$$

where $C^+ = \frac{N_\mathcal{P}}{N_p}C$ and $C^- = \frac{N_\mathcal{N}}{N_p}C$ are the error penalization constants, and $\mathcal{P}$ and $\mathcal{N}$ represent the positive and negative training data points, respectively.

To evaluate the performance of the classifier we performed a stratified 10-fold cross-validation strategy. The SVM was trained using a Gaussian kernel. The regularization parameter C in Eq. (2) and the kernel width γ in Eq. (4) were tuned for each fold separately using a nested cross-validated 3-stage grid-search. We repeated each 10-fold cross-validation 50 times to account for its stability.

We computed accuracy, sensitivity, and specificity to quantify the performance of the classifier based on the results of the cross-validation stage. Accuracy is the proportion of correctly classified subjects in all subjects. Sensitivity is the proportion of correctly classified ex-combatants in all ex-combatants, while specificity is the proportion of correctly classified civilians in all civilians.

F. STATISTICAL ANALYSIS

Social cognition and behavior (SCB) data were analyzed using the Statistical Package for Social Sciences (IBM SPSS version 23.0 for Windows). Independent samples t-tests or χ^2 (for gender) were performed to assess differences between-groups for demographic variables and SCB scores. Wilcoxon rank-sum tests were used to measure the differences in whole brain-averaged level of iCOH, and in the network metrics between the ex-combatants and civilians. p-values were corrected for multiple comparisons using the two-stage FDR method [45]. For all statistical analyses, a p (or q) value of less than 0.05 was considered to be statistically significant.

Furthermore, we conducted ROC curve analyses to establish the effectiveness of the SCB tests in classifying individuals as belonging to a given group [46]. The area under the ROC curve (AUC) was used as a measure of discriminatory accuracy for such analyses.

III. RESULTS

A. SOCIAL COGNITION AND BEHAVIOR

We found significant between-group differences in both the global social skills score (GSSS) from EHS, and proactive and reactive aggression scores from RPQ. Table 2 shows the mean, standard deviation from both groups, t-tests, and ROC analyses results of SCB scores.

As SCB scores were statistically significant for the two groups, we used a SVM classifier to estimate the prediction accuracy using only these three features. We obtained an accuracy of M = 70.00%, SD = 5.06%.

B. FUNCTIONAL CONNECTIVITY NETWORK

To visualize the connections between all pairs of electrodes, we used the procedure originally introduced in [31]. In this procedure the scalp is depicted as a large circle. Small circles are placed at the location of each electrode also representing the scalp and containing the absolute value of iCOH for the respective electrode (marked as a black dot) with all other electrodes. From the qualitative comparison of maps shown in Fig. 4, the iCOH for ex-combatants shows a spatial pattern similar to that for civilians.

Between-group differences were not observed for whole brain-averaged level of iCOH. Both groups were therefore indistinguishable, meaning that the overall level of connectivity did not define a phenotype. Table 3 summarizes the mean value, the standard deviation, and the test results per task per band.

For the word task, significant differences between groups were obtained for leaf fraction (LF) and maximum degree (MD) for the delta band; and LF, MD and diameter (D) for the

TABLE 2. Group results of the SCB scores and *t* (and *p*) values for between group comparisons based on Student's *t*-test.

	Ex-combatants	Civilians		
	M±*SD*	*M*±*SD*	*t* (*p*)	*AUC* and 0.95 *CI*
GSSS	68.10±15.86	80.40±25.56	-2.10 (**0.041**)	0.709 (0.561–0.858)
Proactive	3.63±3.40	1.50±1.91	2.83 (**0.007**)	0.737 (0.596–0.877)
Reactive	7.33±3.17	5.50±2.91	2.07 (**0.044**)	0.687 (0.531–0.843)

*Bold values indicate significant *p*-values.

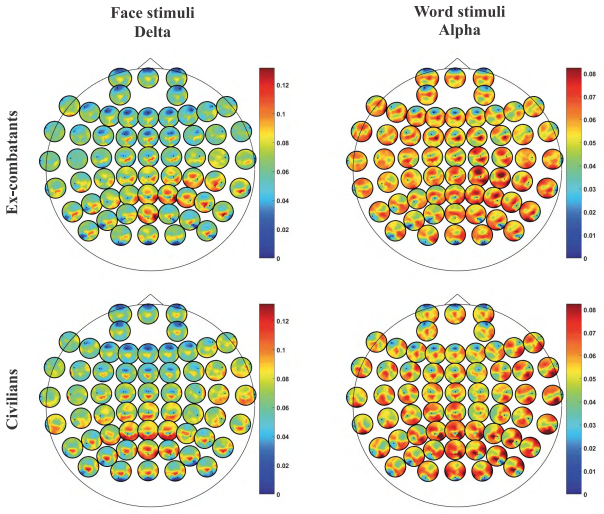

FIGURE 4. Absolute value of the imaginary part of coherency. These maps correspond to the face stimuli in the delta band, and to the word stimuli in the alpha band, as they were further validated to be the top-ranked discriminative features.

TABLE 3. Group comparisons based on Wilcoxon rank-sum test. Results for iCOH, *Z*-scores, and *p*-values.

		Ex-combatants	Civilians	
Band	Task	*M*±*SD*	*M*±*SD*	*Z* (*p*)
Delta	Words	0.061±0.020	0.061±0.019	-0.139 (0.890)
	Faces	0.067±0.016	0.071±0.022	-0.297 (0.766)
Theta	Words	0.077±0.022	0.077±0.037	-0.891 (0.373)
	Faces	0.094±0.028	0.097±0.025	-0.535 (0.593)
Alpha	Words	0.056±0.014	0.055±0.016	-0.574 (0.566)
	Faces	0.065±0.020	0.066±0.017	-0.673 (0.501)
Beta	Words	0.035±0.011	0.035±0.006	-0.693 (0.488)
	Faces	0.038±0.012	0.037±0.008	-0.257 (0.797)

alpha band. For the face task, we found significant differences in LF for the delta band; LF, MD and D for the theta band; MD for the alpha band; and LF for the beta band. Table 4 shows the mean, standard deviation, and test results for the network metrics per task per band.

C. FEATURE SELECTION

We defined a 24-dimensional neurophysiological set comprised of the FCN network metrics. We performed SVM-RFE to determine which of them were most predictive of ex-combatant or civilian classification. Classification accuracies for this step are reported in Fig. 5. We found that the subset comprised of leaf fraction and diameter for the face stimuli in

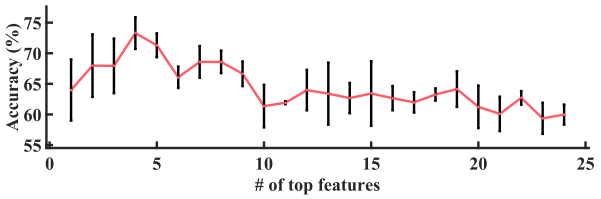

FIGURE 5. SVM-RFE results on FCN features. Average classification accuracy of the top-ranked features. The error bars represent the standard deviations.

delta band, and leaf fraction and diameter for the word stimuli in alpha band showed the highest discriminability.

Furthermore, the CCA was performed using the top four FCN ($N_a = 4$) and the 3-dimensional set comprised of the SCB scores ($N_b = 3$). With this final step, we conformed six different sets of features, presented in column one of Table 5.

D. SVM CLASSIFICATION

We implemented an SVM classifier based on the libsvm library [47]. To tune the parameters of SVM, we performed a 3-level grid search using growing sequences in the range $[2^{-5}, 2^{17}]$ for C, and $[2^{-17}, 2^{3}]$ for γ. Additionally, to account for imbalances in the dataset, we set the option `wi` to be the ratio $\frac{N_{\mathcal{P}}}{N_{\mathcal{N}}} = 1.5$ for the positive class (ex-combatants) and 1.0 for the negative one (civilians). Generalization of classification performances was ensured by nested 10-fold cross-validation. Classification performances of the SVM classifier with each set of features are summarized in Table 5.

IV. DISCUSSION

In this paper, we presented a methodological pipeline that aimed to assess EP in ex-combatants using a SVM to classify individuals. We identified the most discriminative features for accurate classification using a feature fusion technique based on CCA over selected features from EEG-based FCN metrics and SCB scores. We chose iCOH to estimate the FCN because it has proven to be robust to volume conduction effects, as it is only non-zero for non-zero phase delays [31]. We used the MST to binarize the FCN as its usage avoids the introduction of arbitrary thresholds.

The SCB yielded significant differences between groups. ROC analyses on SCB scores resulted in acceptable discriminatory accuracy, with AUC values between 0.687 and 0.709, and 0.95 CI excluding 0.5. Used as classification features,

TABLE 4. Group results of the network metrics, *Z*-scores, and corrected *p*-values for their group comparisons based on Wilcoxon rank-sum test.

		Leaf fraction			Max. degree			Diameter		
		Ex-combatants	Civilians		Ex-combatants	Civilians		Ex-combatants	Civilians	
Band	Task	$M\pm SD$	$M\pm SD$	Z (q)	$M\pm SD$	$M\pm SD$	Z (q)	$M\pm SD$	$M\pm SD$	Z (q)
Delta	Words	0.749±0.060	0.711±0.063	-3.389 (**0.004**)	18.578±6.253	16.400±6.416	-2.363 (**0.036**)	9.622±2.009	10.217±2.084	1.500 (0.164)
	Faces	0.761±0.053	0.733±0.055	-2.864 (**0.017**)	19.733±8.067	17.317±5.803	-1.650 (0.136)	9.167±1.874	10.100±2.097	2.675 (**0.021**)
Theta	Words	0.753±0.060	0.736±0.060	-1.409 (0.184)	18.956±6.413	17.667±6.027	-1.306 (0.211)	9.500±1.973	9.767±2.174	0.629 (0.467)
	Faces	0.810±0.049	0.779±0.052	-3.728 (**0.002**)	25.789±8.288	21.783±7.461	-2.832 (**0.017**)	7.978±1.635	8.683±1.818	2.366 (**0.036**)
Alpha	Words	0.740±0.057	0.709±0.048	-3.789 (**0.002**)	17.111±6.306	14.300±4.666	-2.529 (**0.028**)	9.789±1.963	11.117±2.415	3.521 (**0.003**)
	Faces	0.760±0.051	0.740±0.056	-1.826 (0.115)	19.189±5.904	17.817±7.808	-2.183 (0.053)	9.244±2.063	9.617±2.132	0.634 (0.467)
Beta	Words	0.717±0.069	0.699±0.057	-1.759 (0.116)	15.700±6.942	14.850±5.963	-0.525 (0.508)	10.644±2.745	10.917±2.472	0.884 (0.378)
	Faces	0.722±0.064	0.695±0.061	-2.719 (**0.021**)	15.011±5.049	14.150±5.665	-1.758 (0.116)	10.667±2.308	10.750±2.297	0.285 (0.570)

Bold values indicate significant q-values.

TABLE 5. Accuracy, sensitivity and specificity reached with the proposed methodology.

Feature Selection Algorithm	# Features	Accuracy (%)	Sensitivity (%)	Specificity (%)
FCN features	30	60.00 (2.53)	67.22 (1.85)	50.07 (3.06)
FCN + SCB features	33	61.33 (3.50)	67.69 (3.44)	51.66 (4.39)
SVM-RFE	4	75.29 (1.01)	80.60 (1.73)	68.58 (2.94)
SVM-RFE + SCB features	7	78.67 (2.42)	82.34 (3.27)	73.30 (2.18)
SVM-RFE + CCA concatenation	6	81.00 (4.15)	83.21 (2.64)	77.66 (6.64)
SVM-RFE + CCA summation	**3**	**84.33 (1.97)**	**85.61 (2.24)**	**82.32 (2.34)**

*Bold values indicate best results. Values are given as M (SD).

SCB scores outperformed the classification rate obtained with the raw FCN features.

The use of SVM-RFE to identify the top-ranked FCN features allowed improving the classification rates from 60.00% to 75.29%. With the inclusion of CCA to fuse FCN and SCB features, the classification accuracy reached values of up to 84.33% for the summation fusion. This value outperformed the results obtained in previous works (in [7] we achieved 80.00%, and in [10] a mean accuracy of 58.4%). Furthermore, SVM-RFE allowed identifying a potential phenotype given by the pair Leaf fraction–Diameter in the delta band elicited by face stimuli, and in the alpha band elicited by word stimuli. Specifically, controls have higher diameter values that reflect the presence of networks with reduced global efficiency [38]. Conversely, higher values of leaf fraction in ex-combatants show the presence of clustered nodes that dominate the network topology, and it is thought to confer high integration of information within the network for specialized processing [16].

This work provides new empirical knowledge on the reorganization of EEG-FCN in EP of ex-combatants. On one hand, most of the previous EP studies focused on comparing well-defined groups, such as healthy controls vs. clinical population [12], [13]. Here, the data came from two non-clinical groups which are only distinguishable based on whether they have war experiences or not, and not on any demographic, psychiatric or neurological background measures. On the other hand, previous studies detected atypical functioning of EP in ex-combatants linked to EEG spatio-temporal patterns [7], and regional values of FC [10]. Here, we proposed a methodology to identify potential cognitive phenotypes linked to atypical EP using SCB features and graph metrics from EEG FCN.

While the current study found the leaf fraction and diameter to be significantly discriminative between groups, it remains to be seen if these traits may therefore prove to be effective biomarkers for impairments in EP. In addition to the promise as a diagnostic marker, these features merit further investigation as functional markers of response to psychological interventions conducted to reduce the prevalence of aggression attitudes in ex-combatants.

V. CONCLUSION

This paper presented an SVM-based system to estimate cognitive phenotypes related with emotional processing in ex-combatants. To our knowledge, this is the first phenotyping scheme based on behavioral and brain FCN measures of emotional processing in this population. In our experiments, we reached the highest accuracy using a method of feature fusion based on canonical correlation analysis.

Results show the existence of a cognitive phenotype related to increased values of the leaf fraction and reduced values of the diameter of the FCN in ex-combatants, in comparison with controls. This demonstrates that combat experience modulates the EP in ex-combatants by forcing a reorganization of their FCN.

If further validated, we would expect this approach to help in the monitoring of DDR programs in countries with internal conflicts, such as Colombia, and to provide a new research approach to characterize affective problems in populations that do not fall in the classical definitions of disease.

ACKNOWLEDGMENT

The authors appreciate the assistance of Reincorporation and Normalization Agency (ARN), and thank the volunteers for their effort and patience.

REFERENCES

[1] M. Denissen, "Reintegrating Ex-combatants into civilian life: The case of the paramilitaries in Colombia," *Peace Change*, vol. 35, no. 2, pp. 328–352, Apr. 2010.

[2] O. Kaplan and E. Nussio, "Community counts: The social reintegration of Ex-combatants in Colombia," *Conflict Manage. Peace Sci.*, vol. 35, no. 2, pp. 132–153, Nov. 2015.

[3] A. Köbach, C. Nandi, A. Crombach, M. Bambonyé, B. Westner, and T. Elbert, "Violent offending promotes appetitive aggression rather than posttraumatic stress—A replication study with burundian ex-combatants," *Frontiers Psychol.*, vol. 6, p. 1755, Dec. 2015.

[4] R. Weierstall, C. Patricia, B. Castellanos, F. Neuner, and T. Elbert, "Relations among appetitive aggression, post-traumatic stress and motives for demobilization: A study in former Colombian combatants," *Conflict Health*, vol. 7, p. 9, Apr. 2013.

[5] M. R. López, E. Andreouli, and C. Howarth, "From Ex-combatants to citizens: Connecting everyday citizenship and social reintegration in Colombia," *J. Social Political Psychol.*, vol. 3, no. 2, pp. 171–191, 2015.

[6] O. Kaplan and E. Nussio, "Explaining recidivism of Ex-combatants in Colombia," *J. Conflict Resolution*, vol. 62, no. 1, pp. 64–93, May 2016.

[7] A. Quintero-Zea *et al.*, "Characterization framework for Ex-combatants based on EEG and behavioral features," in *Proc. 7th Latin Amer. Congr. Biomed. Eng. (CLAIB)*, in IFMBE Proceedings, vol. 60, I. Torres, J. Bustamante, and D. A. Sierra, Eds. Singapore: Springer, 2017, pp. 205–208.

[8] S. P. Trujillo *et al.*, "Atypical modulations of N170 component during emotional processing and their links to social behaviors in Ex-combatants," *Frontiers Hum. Neurosci.*, vol. 11, pp. 1–12, May 2017.

[9] S. Trujillo *et al.*, "Social cognitive training improves emotional processing and reduces aggressive attitudes in Ex-combatants," *Frontiers Psychol.*, vol. 8, pp. 1–13, Apr. 2017.

[10] M. V. Rodríguez-Calvache, A. Quintero-Zea, S. P. Trujillo-Orrego, N. Trujillo-Orrego, and J. D. López-Hincapié, "Detecting atypical functioning of emotional processing in Colombian Ex-combatants," *Tecno Lógicas*, vol. 20, no. 40, pp. 83–96, 2017.

[11] L. Carretié, J. Iglesias, T. Garcia, and M. Ballesteros, "N300, P300 and the emotional processing of visual stimuli," *Clin. Neurophysiol.*, vol. 103, no. 2, pp. 298–303, Aug. 1997.

[12] P. L. Rock, G. M. Goodwin, and C. J. Harmer, "The common adolescent bipolar phenotype shows positive biases in emotional processing," *Bipolar Disorders*, vol. 12, no. 6, pp. 606–615, Sep. 2010.

[13] A. C. Miu, S. E. Pană, and J. Avram, "Emotional face processing in neurotypicals with autistic traits: Implications for the broad autism phenotype," *Psychiatry Res.*, vol. 198, no. 3, pp. 489–494, Aug. 2012.

[14] G. G. Knyazev, J. Y. Slobodskoj-Plusnin, and A. V. Bocharov, "Gender differences in implicit and explicit processing of emotional facial expressions as revealed by event-related theta synchronization," *Emotion*, vol. 10, no. 5, pp. 678–687, 2010.

[15] K. J. Friston, "Functional and effective connectivity in neuroimaging: A synthesis," *Hum. Brain Mapping*, vol. 2, nos. 1–2, pp. 56–78, 1994.

[16] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, Mar. 2009.

[17] E. R. John, L. S. Prichep, and M. Almas, "Subtyping of psychiatric patients by cluster analysis of QEEG," *Brain Topography*, vol. 4, no. 4, pp. 321–326, 1992.

[18] J. Johnstone, J. Gunkelman, and J. Lunt, "Clinical database development: Characterization of EEG phenotypes," *Clin. EEG Neurosci.*, vol. 36, no. 2, pp. 99–107, Apr. 2005.

[19] S. Afshari and M. Jalili, "Directed functional networks in Alzheimer's disease: Disruption of global and local connectivity measures," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 4, pp. 949–955, Jul. 2017.

[20] M. Arns, J. Gunkelman, M. Breteler, and D. E. Spronk, "EEG phenotypes predict treatment outcome to stimulants in children with ADHD," *J. Integrative Neurosci.*, vol. 7, no. 3, pp. 421–438, 2008.

[21] L. Schmüser, A. Sebastian, A. Mobascher, K. Lieb, B. Feige, and O. Tüscher, "Data-driven analysis of simultaneous EEG/fMRI reveals neurophysiological phenotypes of impulse control," *Hum. Brain Mapping*, vol. 37, no. 9, pp. 3114–3136, Sep. 2016.

[22] M. Tursich, T. Ros, P. A. Frewen, R. C. Kluetsch, V. D. Calhoun, and R. A. Lanius, "Distinct intrinsic network connectivity patterns of post-traumatic stress disorder symptom clusters," *Acta Psychiatrica Scandinavica*, vol. 132, no. 1, pp. 29–38, Jul. 2015.

[23] Y. Li, D. Cao, L. Wei, Y. Tang, and J. Wang, "Abnormal functional connectivity of EEG gamma band in patients with depression during emotional face processing," *Clin. Neurophysiol.*, vol. 126, no. 11, pp. 2078–2089, Nov. 2015.

[24] H. Cao *et al.*, "Altered functional subnetwork during emotional face processing," *JAMA Psychiatry*, vol. 73, no. 6, pp. 598–605, Jun. 2016.

[25] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2005, pp. 317–321.

[26] M. González, Estudio sociolingüístico de Medellín. Fase 1. Corpus sociolingüístico de Medellín. [Online]. Available: http://comunicaciones.udea.edu.co/corpuslinguistico/

[27] R. Oostenveld, P. Fries, E. Maris, and J. M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, p. 156869, Dec. 2011.

[28] E. Gismero, *Escala de Habilidades Sociales (EHS)*. Madrid, Spain: TEA Publicaciones de Psicologia Aplicada., 2000.

[29] A. Raine *et al.*, "The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys," *Aggressive Behav.*, vol. 32, no. 2, pp. 159–171, Apr. 2006.

[30] A. Khadem and G.-A. Hossein-Zadeh, "Comparing the robustness of brain connectivity measures to volume conduction artifact," in *Proc. 20th Iranian Conf. Biomed. Eng. (ICBME)*, Dec. 2013, pp. 209–214.

[31] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, "Identifying true brain interaction from EEG data using the imaginary part of coherency," *Clin. Neurophysiol.*, vol. 115, no. 10, pp. 2292–2307, Oct. 2004.

[32] C. Babiloni *et al.*, "Intra-hemispheric functional coupling of alpha rhythms is related to golfer's performance: A coherence EEG study," *Int. J. Psychophysiol.*, vol. 82, no. 3, pp. 260–268, 2011.

[33] B. Babadi and E. N. Brown, "A review of multitaper spectral analysis," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1555–1564, May 2014.

[34] J. B. Kruskal, Jr., "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc. Amer. Math. Soc.*, vol. 7, no. 1, pp. 48–50, Feb. 1956.

[35] C. J. Stam, P. Tewarie, E. Van Dellen, E. C. W. Van Straaten, A. Hillebrand, and P. Van Mieghem, "The trees and the forest: Characterization of complex brain networks with minimum spanning trees," *Int. J. Psychophysiol.*, vol. 92, pp. 129–138, Jun. 2014.

[36] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[37] V. Latora and M. Marchiori, "Efficient behavior of small-world networks," *Phys. Rev. Lett.*, vol. 87, p. 198701, Oct. 2001.

[38] P. Tewarie, E. van Dellen, A. Hillebrand, and C. J. Stam, "The minimum spanning tree: An unbiased method for brain network analysis," *NeuroImage*, vol. 104, pp. 177–188, Jan. 2015.

[39] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, nos. 1–3, pp. 389–422, 2002.

[40] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[41] A. Rakotomamonjy, "Variable selection using SVM-based criteria," *J. Mach. Learn. Res.*, vol. 3, pp. 1357–1370, Mar. 2003.

[42] Q.-S. Sun, S.-G. Zeng, Y. Liu, P.-A. Heng, and D.-S. Xia, "A new method of feature fusion and its application in image recognition," *Pattern Recognit.*, vol. 38, no. 12, pp. 2437–2448, Dec. 2005.

[43] G. H. Golub and H. Zha, "Perturbation analysis of the canonical correlations of matrix pairs," *Linear Algebra Appl.*, vol. 210, pp. 3–28, Oct. 1994.

[44] K. Veropoulos, C. Campbell, and N. Cristianini, "Controlling the sensitivity of support vector machines," in *Proc. Int. Joint Conf. Artif. Intell.*, 1999, pp. 55–60.

[45] Y. Benjamini, A. M. Krieger, and D. Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, vol. 93, no. 3, pp. 491–507, Sep. 2006.

[46] G. Tripepi, K. J. Jager, F. W. Dekker, and C. Zoccali, "Diagnostic methods 2: Receiver operating characteristic (ROC) curves," *Kidney Int.*, vol. 76, no. 3, pp. 252–256, Aug. 2009.
[47] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011. [Online]. Available: http://www.csie.ntu. edu.tw/~cjlin/libsvm

ANDRÉS QUINTERO-ZEA received the bachelor's degree in electronic engineering from the Universidad de Antioquia, Medellín, Colombia, in 2006, where he is currently pursuing the Ph.D. degree in electronic engineering. He was a full-time faculty with the Universidad de Medellín from 2010 to 2014. His research interests include machine learning, signal processing, and cognitive neuroscience.

JOSÉ D. LÓPEZ received the bachelor's degree in electronic engineering and the master's degree in electronics from the Universidad de Antioquia, Medellín, Colombia, and the Ph.D. degree from the Universidad Nacional de Colombia in 2013. He was a full-time Faculty with the Universidad de Medellín for three years, finishing in 2009. He was a Coordinator of the Electronic Engineering Program at the Universidad de Antioquia from 2012 to 2014, where he has been a full-time Professor (Bioengineering Program) since 2014. His research interests are related to large scale systems, digital signal processing, neuroimaging, and computational psychology. Since 2018, he has also been an Honorary Research Associate with the Wellcome Centre for Human Neuroimaging, UCL, U.K.

KEITH SMITH received the bachelor's degree (Hons.) in mathematics from University of Glasgow, the master's degree in actuarial science from Heriot-Watt University, Edinburgh, and the Ph.D. degree in biomedical engineering from The University of Edinburgh. He is currently an HDRUK UKRI Innovation Fellow with the Usher Institute for Population Health Science and Informatics, The University of Edinburgh. He is involved in research in applied mathematics, complex networks, neuroimaging, and health data science. His most recent publication is *Accounting for the Complex Hierarchical Topology of EEG Phase-Based Functional Connectivity in Network Binarisation*.

NATALIA TRUJILLO received the bachelor's degree from the Psychology Department, San Buenaventura University, in 2005, the master's degree (specialist) in forensic psychology from San Buenaventura University in 2007, and the Ph.D. degree in biomedical basic science with a focus on neuroscientific research from the Neuroscience Group, University of Antioquia, Colombia, in 2011. She was a Young Researcher at Neuropsychology in 2006. She conducted a Research Group at the University of San Buenaventura, Medellín, Colombia. In 2010, she became a full-time Professor with the Psychology Department, Fundacion Universitaria Maria Cano. From 2011 to 2014, she was a full-time Researcher with the Neuroscience Group, University of Antioquia, where she has been a full-time Professor in Public Health Program since 2014.

MARIO A. PARRA graduated as a Medical Doctor in 1993 and a Clinical Neurophysiologist in 1997. He received the Ph.D. degree in psychology from The University of Edinburgh in 2009. He spent 12 years practicing and teaching in clinical settings. During his clinical work, his focused on neurophysiological aspects of dementia syndromes and other neurological disorders. His Ph.D. study was supported by the European Program ALBAN. He held two consecutive post-doctoral fellowships—the first post-doctoral fellowship was supported by the Neuroscience Program of Campagnia di San Paolo Foundation and the second post-doctoral fellowship was supported by Alzheimer's Society. He has been supporting Clinical Research as a Research Officer of the Scottish Dementia Clinical Research Network, U.K., since 2009. In 2015, he was appointed as an Assistant Professor in psychology at Heriot Watt University.

JAVIER ESCUDERO (S'07–M'10) received the M.Eng. and Ph.D. degrees in telecommunications engineering from the University of Valladolid, Spain, in 2005 and 2010, respectively. He held a post-doctoral position at University of Plymouth, U.K., until 2013. He is currently a tenured Faculty Member (Chancellor's Fellow) with the School of Engineering, The University of Edinburgh, U.K., where he leads a research group in biomedical signal processing with particular interests in non-linear analysis, network theory, and multiway decompositions. He has authored over 45 scientific articles. In 2016, he was elected as a member of the Young Academy of Scotland. He is the President of the Society of Spanish Researchers in the U.K. for the 2018–2019 term. He received the Third Prize of the EMBS Student Paper Competition in 2007 and the Best Ph.D. Thesis Award in healthcare technologies from the Spanish Organization of Telecommunications Engineers in 2010.