

1 **Machine learning Classification of Dyslexic Children based**
2 **on EEG Local Network Features**

3 Rezvani, Z.¹, Zare, M.², Žarić, G.³, Bonte, M.³, Tijms, J.^{4,5}, Van der Molen, M.W.⁶, Fraga
4 González, G.⁷

5
6 ¹Institute for Cognitive and Brain Sciences, Shahid Beheshti University G.C., Tehran, Iran

7 ²School of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

8 ³Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience,
9 Maastricht University, Netherlands

10 ⁴Rudolfs Berlin Center, Amsterdam, Netherlands

11 ⁵IWAL Institute, Amsterdam, Netherlands

12 ⁶Department of Psychology, University of Amsterdam, Netherlands

13 ⁷Amsterdam Brain and Cognition, University of Amsterdam, Netherlands

14

15

16 **Abstract**

17 Machine learning can be used to find meaningful patterns characterizing individual
18 differences. Deploying a machine learning classifier fed by local features derived from graph
19 analysis of electroencephalographic (EEG) data, we aimed at designing a neurobiologically-
20 based classifier to differentiate two groups of children, one group with and the other group
21 without dyslexia, in a robust way. We used EEG resting-state data of 29 dyslexics and 15 typical
22 readers in grade 3, and calculated weighted connectivity matrices for multiple frequency bands
23 using the phase lag index (PLI). From the connectivity matrices, we derived weighted
24 connectivity graphs. A number of local network measures were computed from those graphs, and
25 37 False Discovery Rate (FDR) corrected features were selected as input to a Support Vector
26 Machine (SVM) and a common K Nearest Neighbors (KNN) classifier. Cross validation was
27 employed to assess the machine-learning performance and random shuffling to assure the
28 performance appropriateness of the classifier and avoid features overfitting. The best
29 performance was for the SVM using a polynomial kernel. Children were classified with 95%
30 accuracy based on local network features from different frequency bands. The automatic
31 classification techniques applied to EEG graph measures showed to be both robust and reliable in
32 distinguishing between typical and dyslexic readers.

33 **Keywords:** EEG; Machine Learning; Support Vector Machine (SVM); KNN classifier; Local
34 Network measures; Dyslexia; Children.

35

36

37

38 **1. Introduction**

39 Developmental dyslexia is a specific reading and spelling disability with a genetic and
40 neurobiological component and relatively high prevalence rates around 5% (Blomert 2005;
41 Snowling 2013). Neuroimaging studies have investigated biomarkers of dyslexia using structural
42 and functional network analyses and there is a growing interest in connectivity abnormalities
43 between different brain systems that may result in impaired reading (e.g., Finn et al. 2014;
44 Schurz et al. 2014; Liu et al. 2015). Diffusion tensor imaging studies show reduced connectivity
45 in the main white matter pathways (see review in Vandermosten et al. 2012). Functional
46 magnetic resonance imaging (fMRI) studies reported reduced long-range connectivity across
47 brain systems specialized for reading (e.g., Pugh et al. 2000; Shaywitz et al. 2003; van der Mark
48 et al. 2011; Schurz et al. 2014), and associated functional networks (Wolf et al. 2010; Finn et al.
49 2014). Whole brain connectivity studies observed that dyslexics differ from typical readers in
50 whole brain networks organization, showing increased local processing and less long-range
51 communication (Liu et al. 2015) and lower global efficiency in dyslexics (Dimitriadis et al.
52 2018). Collectively, this evidence supports the view that a widespread network of brain regions
53 may be compromised in developmental dyslexia (Martin et al. 2016).

54 An important objective of the neuroscientific search for biomarkers of dyslexia is to
55 contribute to the diagnosis and early detection or prediction of reading disabilities. The
56 classification of individuals suffering from several neuropsychological disorders may take
57 advantage of using machine learning methods (e.g., Duda et al. 2016; Kessler et al. 2016). These
58 methods are particularly useful when applied to neuroimaging data as they allow for using
59 widely distributed information to improve the classification of clinical groups or individuals at
60 risk. A recent MRI study used a machine learning classifier and reported above chance levels in

61 discriminating dyslexic adults from controls based on gray matter differences (Tamboer et al.
62 2016). Two related studies employed multivariate pattern analysis of brain activity during a
63 phonological task to identify poor readers (Tanaka et al. 2011) and predict long-term outcomes
64 in dyslexic children based on whole-brain activation (Hoeft et al. 2011). The latter study showed
65 that methods using brain measures outperformed procedures relying on behavioral measures in
66 predicting reading improvements across the 2.5 years following the experiment. These studies
67 illustrate the potential of using machine learning techniques in combination with neuroimaging
68 data to improve the classification and early detection of dyslexia.

69 The present study uses a ML classifier to discriminate between dyslexic and typically reading
70 children based on functional connectivity using measures derived from the electroencephalogram
71 (EEG). It has been demonstrated previously that task-independent EEG activity contains
72 information about how different brain systems communicate and how functional networks may
73 be intrinsically organized (van den Heuvel and Hulshoff Pol 2010). Importantly, earlier studies
74 related neural activity at rest to reading ability in children and adults, showing that a resting-state
75 paradigm can be profitably used to study language networks (Hampson et al. 2006; Koyama et
76 al. 2010).

77 Given the highly interactive and complex nature of reading, the study of its neurobiology
78 might benefit from an integrative and holistic view of brain function conceptualized as a
79 complex network (Bullmore and Sporns 2009). Within that framework, graph theoretical
80 analysis allows for modeling whole-brain functional connectivity networks as a set of nodes
81 (vertices) and the connections between them (edges). The multiple measures that can be derived
82 from a graph are used to describe the network in terms of information transfer and balance
83 between ‘segregation’ and ‘integration’ (see reviews in Bullmore and Sporns 2009, 2012). Two

84 magnetoencephalographic (MEG) studies of dyslexia examined graph measures and found
85 dysfunctional long- and short-range functional connectivity in dyslexics during a reading task
86 (Vourkas et al. 2011) and less organized connectivity at rest (Dimitriadis et al. 2013). In a
87 previous study, we applied graph analysis to resting-state EEG to compare dyslexics and
88 typically reading children in grade 3 (Fraga González et al. 2016). The results suggested group
89 differences in several *global* graph metrics in the theta band suggesting a reduced network
90 integration and communication between the nodes in dyslexics compared to typical readers.

91 In our previous graph study, we examined global properties of the network that were
92 described by graph measures averaged across electrodes (Fraga González et al. 2016). For the
93 current analysis, we will employ data from that study, however, we use a different analysis
94 approach by extracting local features, i.e., computed per node, and take a step further to clinical
95 application and deploy artificial intelligence. This local information could reflect aspects of
96 regional connectivity organization relevant to network development in dyslexia (Liu et al. 2015)
97 and may provide neural features for the benefit of the SVM classifier performance. Basically, an
98 SVM is a discriminative classifier formally defined by a hyperplane. Given labeled training data
99 (supervised learning), the algorithm outputs an optimal hyperplane that categorizes new
100 examples. Due to its ability to manage large datasets, the algorithm is widely used for binary
101 classification problems in machine learning. For more details on SVM see (Hsu et al. 2003). We
102 will use methodological approach that is similar to the one applied in a previous study that, using
103 resting-state EEG and an SVM classifier, identified 6-month-old infants at familial risk for a
104 language learning disorder (Zare et al. 2016). The current study uses SVM and KNN to classify
105 children in 3rd grade as dyslexics or typical readers, based upon a large number of local features
106 derived from functional connectivity matrices in the different frequency bands of the EEG.

107 Cross-validation is employed to assess the resulting classification, and random shuffling is
108 deployed to assure that the classifying performance is not due to bias in feature selection criteria.
109 We aimed at assessing the utility of machine learning techniques to find best distinguishable
110 characteristics in reading difficulties based on functional EEG networks. To the best of our
111 knowledge, this is the first study to use local EEG features to designing a classifier for dyslexia.

112 **2. Methodology**

113 The current analysis is performed on data from a previous study (Fraga González et al. 2016).
114 We refer to that article for a more extensive description of the EEG recordings and summarize
115 here only the information needed for the current study.

116 **2.1. Participants**

117 The participants of the current study were 44 subjects including 29 third-grade dyslexic children
118 (mean age = 8.96; SD = 0.40); with a percentile score of 10 or lower on a standard reading test,
119 and 15 third-grade children (8.75 ± 0.31 years old) in the control group with the same socio-
120 demographical background as dyslexic; with no history of reading difficulties and had a
121 percentile score of 25 or higher on standard reading tests. The participants of the current study
122 were part of a larger sample of 62 children participating in a larger study. Due to young age of
123 our participants some participants were excluded due to excessive movement or other artifacts in
124 the data, or did not complete the resting-state recordings. All participants were native Dutch
125 speakers, received two and a half years of formal reading instruction in primary education. The
126 study was approved by the Ethical Review Board of the University and all parents or caretakers
127 signed informed consent before the children participated. Demographic characteristics and
128 reading scores of the complete sample are included in S1 Table.

129 **2.2. EEG recording and preprocessing**

130 The total duration of the eyes-closed resting state data collected was 2 minutes. EEG data were
131 collected using 64 channels with sampling frequency of 250 Hz; Biosemi ActiveTwo system.
132 Data was imported in Brain Vision Analyzer (Version 2.01.5528 © Brain Products) where spline
133 interpolation was applied to channels with excessive artifacts and segmented in 30 epochs of 4
134 seconds. Epochs containing excessive noise and artifacts were visually inspected and removed.
135 For each subject, 10 artifact-free epochs were selected and exported to ASCII files. The data
136 were imported to Brainwave v0.9.117 (developed by Prof. Cornelis Stam' s research group;
137 freely available at <http://home.kpn.nl/stam7883/brainwave.html>) where it was re-referenced to
138 the common average, submitted to power spectral analysis using Fast Fourier Transform (FFT)
139 and filtered into four frequency bands [delta (0.5-4Hz); theta (4-8Hz); alpha (8-13Hz); beta (13-
140 30Hz)]. Then, the functional connectivity matrices were obtained in Brainwave and based upon
141 those matrices weighted network measures were computed using custom MATLAB code
142 (R2016; The Mathworks, Natick, MA), as well as with functions available as part of the
143 MATLAB Machine learning and Statistics Toolboxes and the Brain Connectivity Toolbox
144 (BCT) (Rubinov and Sporns 2010).

145 **3. EEG analysis**

146 **3.1. Functional connectivity**

147 Functional connectivity was assessed for each segment (10 segments of 4096 data points per
148 subject) and frequency band with the phase lag index (PLI). The PLI measures the asymmetry of
149 the distribution of phase differences between two signals. PLI provides a measure of statistical
150 interdependencies between time series, which reflects the strength of coupling. The major aim of

151 using the phase lag index is to obtain reliable estimates of phase synchronization that are
152 invariant of the presence of common sources, a feature that may be absent in other connectivity
153 measures (Stam et al. 2007). Asymmetry of the phase difference distribution indicates that the
154 likelihood of phase difference $\Delta\varphi$ being in the interval $-\pi < \Delta\varphi < 0$ is different from the
155 likelihood of being in the interval $0 < \Delta\varphi < \pi$. This asymmetry indicates a phase difference (or
156 'lag') between the two time series (Stam et al. 2007). The adjacency matrix is constructed using
157 formula

$$158 \quad PLI(t) = | \langle \text{sign}(\Delta\varphi(t_k)) \rangle |$$

159 where $\Delta\varphi$ is the phase difference and t_k is the k^{th} sequence. PLI ranges between 0 and 1 where
160 zero indicates no coupling. Angle between each pair of time series is calculated using Hilbert
161 transform.

162 **3.2. Weighted network and local feature selection**

163 From the previous step, weighted connectivity matrices were derived for each subject. For each
164 of the individual 64 by 64 matrices, we calculated several local features from weighted network
165 measures per node. In a weighted graph, each electrode represents a node and all the nodes are
166 connected by links with a specific weight representing the strength of connectivity (Barrat et al.
167 2004). The network measures used in this study are summarized in Table 1. We focus on the
168 main network features that are commonly used in the literature and are measured at local level in
169 the network (Bullmore and Sporns 2009; Rubinov and Sporns 2010). The chosen measures aim to
170 quantify how a node contributes in the information flow in the network, in the current study we
171 try to exploit these local characteristics for classification. The features derived from the weighted
172 network were averaged across segments for every subject. Note that for each subject, we

173 repeated the calculations for each connectivity matrices per frequency band and segment. For all
174 network features, the significance of FDR (false discovery rate; Benjamini and Hochberg 1995)
175 corrected group differences ($p < 0.05$) were examined with t-tests defined as ($t(F_{TYP}, F_{DYS}), p$ -
176 value), where F stands for feature name, and then FDR corrected. Then, to demonstrate that this
177 method worked well in feature selection, we employed a random shuffling technique. These
178 steps are described in Fig 1.

Table 1. Network metrics and their definitions.

Network measure	Definition
Characteristic path length	The number of edges mediating between two nodes
Degree	The number of edges incident to a node or sum of weights of incident edges (flow)
Efficiency	Quantifies a network's resistance to failure on a small scale
Clustering coefficient	Quantifies how close a given node and its neighbors are to being a clique
Modularity	Fraction of the edges that fall within the given groups minus the expected fraction if edges were distributed at random
Eccentricity	The maximum distance between a reference's node and node i of a graph
Betweenness centrality (BC)	The number of shortest paths running through a node i

179

180 **SVM classification and cross-validation**

181 The primary goal of this study was to use local network features for classifying children into two
182 separate groups. In the current study, 37 features passed the FDR corrected group comparisons.
183 Those features were used to train the classifier (see Results). Usually, it is difficult to determine
184 in advance which classifier, and in particular which kernel function, fits the SVM classifier best
185 from simpler to more complexity degree for a particular set of data. Therefore, starting from non-
186 parametric classifiers to parametric classifiers, we tested several other classifiers. Here we report
187 the classification using two kernel functions that resulted in good classification performance:

188 linear and polynomial of degree 3. In addition, we compared the SVM to a simpler common
189 classifier, i.e., k-nearest neighbors algorithm (KNN) with $k = 3$ and $k = 7$. In KNN
190 classification, an object is classified by a majority vote of its neighbors, with the object being
191 assigned to the class most common among its k nearest neighbors.

192 Subsequently, the leave-one-out-cross-validation (LOOCV) was deployed to assess the
193 performance of the classifiers averaging performance across N sets (see Fig 1). LOOCV is the
194 most common procedure for cross-validation where the number of sets equals the number of
195 instances in the data set. This approach is used to avoid overfitting and promotes the reliability
196 and generalizability of the results to a new set of data. The selected features were used as input to
197 an SVM that performed supervised classification, mapping children into two groups: dyslexic

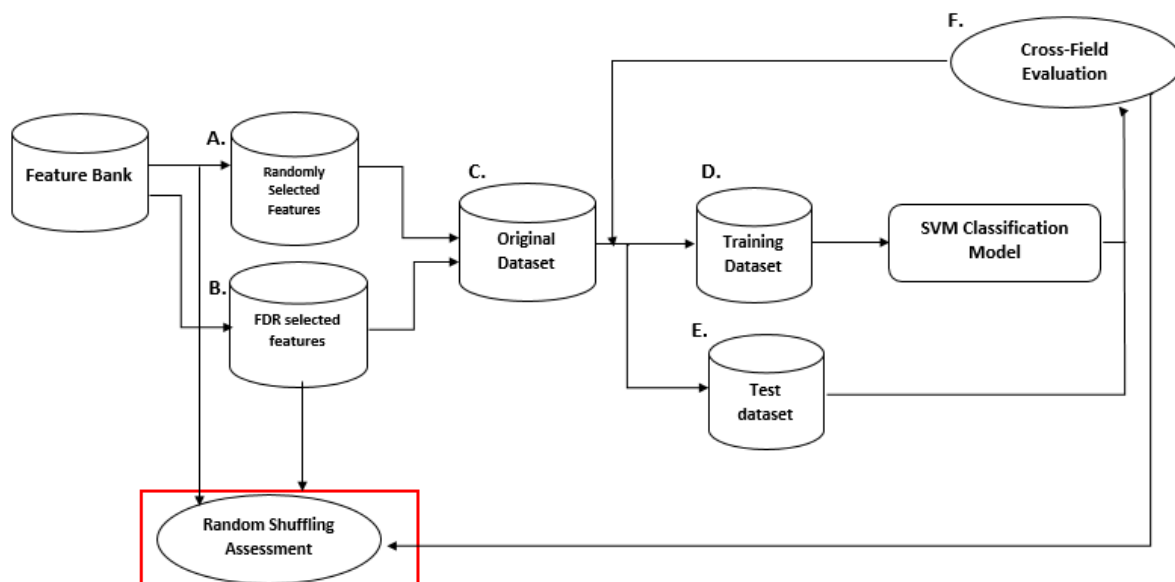


Fig 1. SVM classification and performance assessment by Random Shuffling. We followed two approaches to select features from those available: random selection (A) and selection via t -tests (B). In both cases the dataset was then divided into a Training set (D) and a Test set (E) using cross-validation. We assessed each selected feature with the SVM classifier. Finally, a random shuffling cross-fold evaluation (F) was performed to ensure the t -test selected features were the most relevant for classification.

198 children (DYS) and typical readers (TYP).

199 We assessed the classifier's performance using the conventional measures of precision,
200 specificity, sensitivity, and accuracy. The dyslexic reader group is designated positive and the
201 typical readers are categorized as negative. The correct detection or "true classification" of
202 dyslexia is then a true positive (TP). Likewise, correct classification of the typical readers is true
203 negative (TN). Precision or Positive predictive value (PPV) indicates the proportion of positives,
204 which are correctly identified as such (e.g., the percentage of children who are correctly
205 identified as Dyslexic readers). Specificity refers to the percentage of participants correctly
206 classified as typical readers, which is also known as the true negative rate (TNR). These
207 performance indices are further described in (Zare et al. 2016).

208 The computations related to the SVM classifier were carried out in MATLAB 10.8.0 (R2016;
209 The Mathworks, Natick, MA), as well as with functions available as part of the MATLAB
210 Machine Learning Toolboxes, and in-house MATLAB code.

211 **4. Results**

212 **4.1. PLI and network features extraction**

213 The PLI weighted matrices that were used to extract the local features presented in Table 1. We
 214 extracted a total of 1792 features (4 frequency bands \times 7 network measures \times 64 channels) for
 215 each child, which were used to perform the group comparisons. Applying FDR correction,

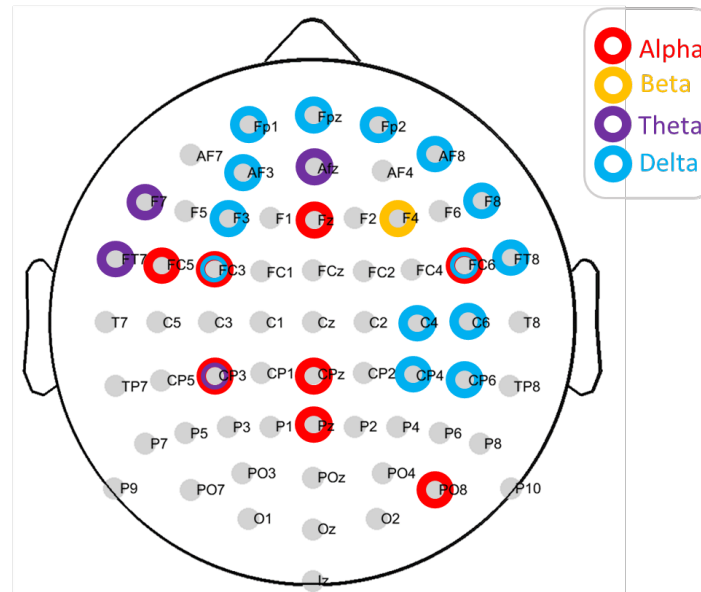


Fig 2. Location on scalp of EEG channels. Highlights indicate the channels where FDR corrected features were selected. The color map indicates frequency band.

216 significant differences ($p < 0.05$) were obtained for a total of 37 features. The results, per
 217 frequency band, separately, are shown in Table 2. In order to visualize our results, we show an
 218 EEG channel scalp map with a color map indicating the electrode sites for which group
 219 differences were found in any of the features per frequency band (see Fig 2).

Table 2. FDR corrected features from weighted matrices derived from PLI.

Feature no.	Frequency Band	Network measure	Channel	p -value	t -value
1	Delta	Betweenness centrality	'F8'	0.0081	2.7824
2			'C4'	0.0031	-3.1404
3		Clustering coefficient	'FP2'	0.0027	-3.1988
4			'AF8'	0.0095	-2.7208
5			'F8'	0.0135	-2.5826
6		Degree	'F3'	0.0151	-2.5358
7			'FP2'	0.0012	-3.4773
8			'AF8'	0.0038	-3.0734

9			'C6'	0.0173	-2.4807
10		Modularity	'AF3'	0.0084	-2.7708
11		Efficiency	'FPZ'	0.0157	-2.5194
12			'FP2'	0.0015	-3.3961
13			'AF8'	0.0063	-2.8815
14			'F8'	0.0100	-2.7027
15			'C6'	0.0182	-2.4603
16			Eccentricity	'FP1'	0.0092
17		'FC3'		0.0152	-2.5350
18		'CP1'		0.0088	-2.7497
19		'AF8'		0.0000	-4.8302
20		'FT8'		0.0081	-2.7834
21		'FC6'		0.0123	-2.6207
22		'C6'		0.0072	-2.8285
23		'CP6'		0.0137	-2.5756
24	Theta	Betweenness centrality	'AFZ'	0.0014	-3.4276
25		Modularity	'F7'	0.0059	2.9055
26			'FT7'	0.0049	2.9740
27			CP3'	0.0142	-2.5610
28			'CP4'	0.0089	2.7466
29	Alpha	Betweenness centrality	'FZ'	0.0037	-3.0773
30			'PO8'	0.0047	-2.9881
31		Modularity	'FC5'	0.0197	-2.4273
32			'FC3'	0.0051	-2.9589
33			'CP3'	0.0112	-2.6577
34			'PZ'	0.0161	-2.5099
35			'CPZ'	0.0014	-3.4211
36			'FC6'	0.0060	2.8975
37	Beta	Modularity	'F4'	0.0152	-2.5345

220

221 Machine Performance

222 The results of machine performance after applying LOOCV are indicated in Table 3. The table
 223 shows that an SVM with a linear kernel provides the statistically best performance among the
 224 classifiers used to identify children as typical readers and dyslexic readers with high specificity,
 225 sensitivity, precision, and accuracy.

Table 3. Machine performance after LOOCV. Complexity of the model is reduced from left to right.

Classifier	SVM Polynomial kernel	SVM Linear kernel	KNN, K=7	KNN, K=3
Accuracy	90.69	95.34	86.04	81.39
Sensitivity	90.00	96.42	86.66	85.71
Specificity	92.30	93.33	84.61	73.33
Precision	96.42	96.42	92.85	85.71

226

227 As input to the machine, we used a total of 37 features vectors, i.e., approximately 2% of the
228 features, based in the FDR corrected group comparison. In order to ensure that the subset of
229 selected features is suitable and unbiased, we used a random-shuffling method (Zare et al. 2016).
230 In this method: a) all network features enter the classifier's pool irrespective of the feature
231 selection criteria. Among all network features, 37 features are randomly chosen to feed the
232 classifier. This way, we assure that feature selection is not biased by the selection method (i.e.,
233 significant difference across two groups over a network feature); b), machine performance is
234 evaluated using the conventional measures of precision, specificity, sensitivity, and accuracy; c)
235 a histogram analysis has to be performed for each of the machine performance parameters (See
236 Fig. 3). The distribution of parameters shows whether the performance parameters are rare and
237 incidental ($p < 0.05$), and finally iv) features that contribute to optimal performance are extracted
238 and compared to those chosen by the initial selection criteria. If those features are fully matched,
239 we conclude that our feature selection is robust and reliable. In particular, for 1000 rounds, we a)
240 randomly chose 37 vectors out of the 1792 features; b) calculated precision, sensitivity,
241 specificity, and accuracy of the machine; d) saved the results in appropriate vectors and drew 4
242 histograms from 1000 element vectors of precision, sensitivity, specificity, and accuracy. Finally,
243 as shown in Fig 3, the distribution of accuracy, precision, sensitivity and specificity show that
244 our selected features are statistically, not random.

245

246

247

248

249

250

251

252

253

254

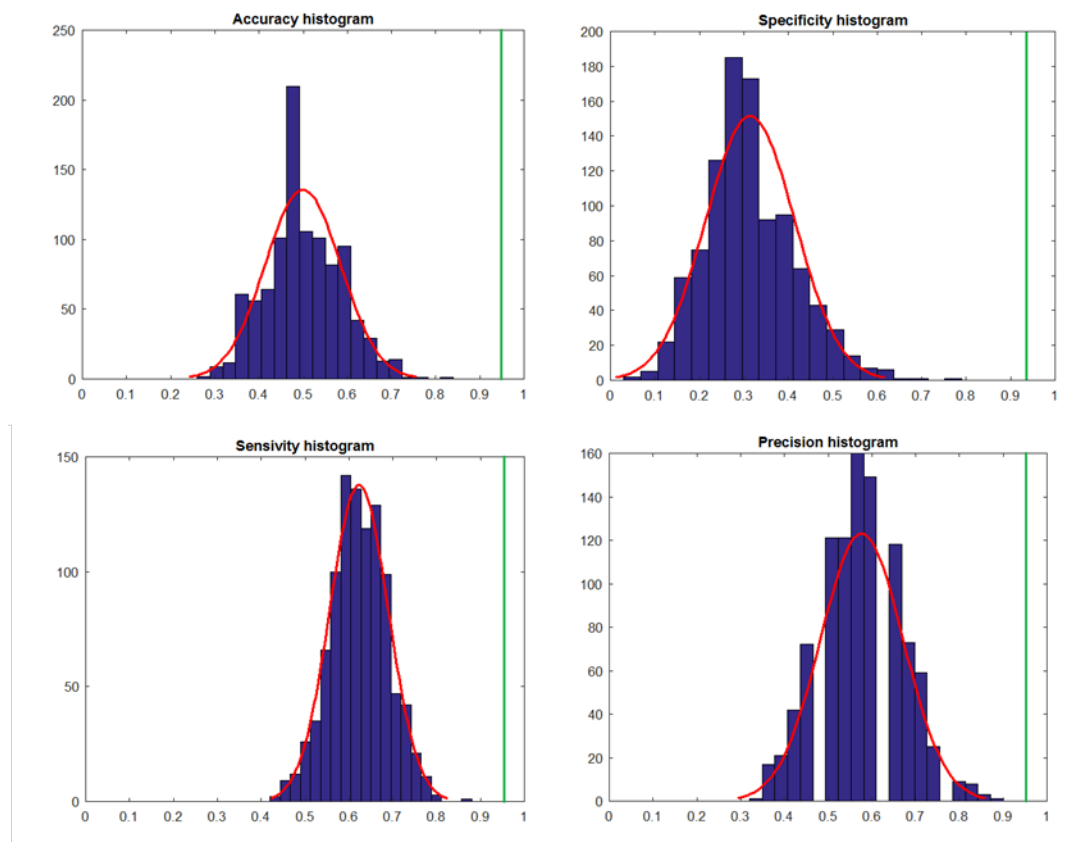


Fig 3. Random shuffling histogram. Red curve shows the distribution of accuracy, specificity, sensitivity and precision for 1000 rounds with the random shuffling of 37 features. The green line shows the results using the 37 FDR corrected features.

255

256 Our random shuffling results indicate that the selected features based on FDR corrected

257 significant differences were robust and reliable.

258 **5. Discussion**

259 We applied two classifier to discriminate between dyslexics and typical readers based on local

260 connectivity features derived from EEG resting-state. In our previous study, we showed group

261 differences in how connectivity is organized between dyslexics and typical readers using graph

262 measures that related to global properties of the network (Fraga González et al. 2016, 2018). In

263 the current analysis, we focus on classification of subjects by using a larger number of features
264 computed per node (i.e., scalp electrode). First, we found group differences associated with
265 several local network measures that were selected to feed an automatic classifier. Then, our
266 cross-validation analysis showed that the classifier could identify dyslexic readers with high
267 accuracy (> 90%). Our random shuffling analysis showed that the selected features, most of
268 which were associated with the delta frequency band, were useful to obtain a high accuracy
269 classification that was not possible to achieve by a random selection of the network features. The
270 current results are consistent with the data of a previous MEG study using sensor-level
271 information (Dimitriadis et al. 2018). The results of this study showed that a higher classification
272 of adults with reading difficulties was possible using sensor-specific network measures
273 compared to using global (averaged across nodes) measures. The current results extend those
274 and previous findings relating resting-state activity to reading abilities (Koyama et al. 2011) and
275 suggesting differences in connectivity organization in dyslexia (Finn et al. 2014; Schurz et al.
276 2014).

277 Our classifier used network features derived from connectivity matrices in different
278 frequency bands. The local network features describe different aspects of connectivity between
279 the nodes, for example, how relevant a node is in the whole network or how well connected it is
280 to its neighbors (e.g., Stam 2014). Importantly, oscillations at different frequency bands have
281 different temporo-spatial features. Higher frequencies seem to relate to more local activity, i.e.,
282 smaller networks, while slower oscillations modulate larger networks with more widespread
283 activity (Buzsáki and Draguhn 2004). It is likely that different aspects of dyslexia are manifested
284 at multiple levels of neural organization. Indeed, previous research shows that reading
285 impairments in dyslexia may result from a heterogeneous cluster of cognitive characteristics

286 (Leinonen et al. 2001; Menghini et al. 2010; Pacheco et al. 2014). This is supported by evidence
287 suggesting similarly complex and heterogeneous neurobiological profiles (e.g. Richlan et al.
288 2009; Finn et al. 2014). The current results suggest that EEG features provide important
289 information about the underlying neurocognitive profiles in dyslexia, which can be used to
290 classify individuals based on task-unspecific electrophysiological data. Recently, a study used
291 structural and functional MRI data to classify a sample of 22 dyslexics and 27 typically reading
292 adults (Tamboer et al. 2016). The study reported accuracy levels around 80% and detected
293 specific occipital and parietal brain regions that contributed most reliably to the classification.
294 We obtained high classification accuracy using local EEG features supporting the potential
295 utility of using automatic classifiers in identifying individuals with dyslexia. However, the use of
296 scalp EEG does not allow to draw conclusions about the underlying sources defining those
297 features. Future studies using source modeling in MEG recordings could bring new insights into
298 what sources are more relevant for classification.

299 For the present findings to have clinical implications, e.g., in diagnostics and early detection
300 of reading disabilities, the generalizability of classification should be further investigated. In the
301 study by Tamboer and colleagues (2016), classification performance dropped to around 60%
302 when the trained classifier was applied to an independent sample. We only performed
303 classification within the available dataset and it would be important to assess the current
304 classifier's performance within additional datasets (Pulini et al. 2018). In addition, to further
305 advance on the relation between EEG network features and cognitive deficits in dyslexia,
306 longitudinal studies should examine whether the current approach could be used to predict
307 reading improvements or treatment outcomes. In relation to this, a previous study used structural
308 and functional MRI data to predict future performance in children with dyslexia and found

309 higher accuracy in prediction using brain measures compared to behavioral tests (Hoeft et al.
310 2007).

311 This study has several limitations. First, the current study included almost twice as many
312 typical readers as readers with dyslexia. This bias is relevant to machine-learning based
313 diagnostics as an imbalance sample results in a disproportionate representation of one of the
314 classes in the training (Mazurowski et al. 2008). Secondly, this and other studies include a gap
315 between typical readers (i.e., readers with a reading performance $> 25\%$) and readers with
316 dyslexia (i.e., readers with a reading performance $< 10\%$). Obviously, such a gap is not present
317 in the real world and might compromise our classifier when applied in a natural setting. Thirdly,
318 the ratio of readers with dyslexia to typical readers is far off from the prevalence rates in the
319 general population (1:20). Future studies should determine the sensitivity of the machine
320 classifiers to changing odd ratios.

321 **General conclusions**

322 This study builds upon the notion that dysfunctional connectivity between local specialized
323 networks may be involved in dyslexia rather than global network measures studied in (Fragza et.
324 al., 2016). We therefore focus on local network properties derived from EEG functional
325 connectivity at rest and show that they can be used to classify individuals as dyslexics and
326 typical readers. The current study presents an interesting and novel biologically-based method to
327 analyze multidimensional data derived from EEG functional connectivity networks. Further
328 research should help elucidating the clinical applicability of EEG-based classification and
329 functional significance of these measures in relation to reading deficits in dyslexia. The current
330 study adds to previous studies that encourage the application of machine learning techniques to

331 neuroimaging data in order to improve subject classification in the context of reading disabilities
332 (Tamboer et al. 2016; Dimitriadis et al. 2018). This approach may benefit from increasingly
333 large data sets available for data-driven analysis.

334 **Acknowledgements**

335 We would like to express our gratitude to all the children and parents for participating in the
336 study.

337 **References**

- 338 Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. The architecture of complex
339 weighted networks. *Proc Natl Acad Sci U S A*. 2004;101(11):3747–52.
- 340 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful
341 Approach to Multiple Testing. *J R Stat Soc Ser B(Methodological)*. 1995;57(1):289–300.
- 342 Blomert L. Dyslexie in Nederland. Amsterdam Uitg Nieuwezijds. 2005;241.
- 343 Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and
344 functional systems. *Nat Rev Neurosci*. 2009;10(3):186–98.
- 345 Bullmore E, Sporns O. The economy of brain network organization. *Nat Rev Neurosci*.
346 2012;13(5):336–49.
- 347 Buzsáki G, Draguhn A. Neuronal Oscillations in Cortical Networks. *Science* (80-).
348 2004;304(5679):1926–9.
- 349 Dimitriadis SI, Laskaris NA, Simos PG, Micheloyannis S, Fletcher JM, Rezaie R, et al. Altered
350 temporal correlations in resting-state connectivity fluctuations in children with reading
351 difficulties detected via MEG. *Neuroimage*. 2013;83:307–17.
- 352 Dimitriadis SI, Simos PG, Fletcher JM, Papanicolaou AC. Aberrant resting-state functional brain
353 networks in dyslexia: Symbolic mutual information analysis of neuromagnetic signals. *Int J*
354 *Psychophysiol*. 2018;(February).
- 355 Duda M, Ma R, Haber N, Wall DP. Use of machine learning for behavioral distinction of autism
356 and ADHD. *Transl Psychiatry*. 2016;6(2):e732.
- 357 Finn ES, Shen X, Holahan JM, Scheinost D, Lacadie C, Papademetris X, et al. Disruption of
358 functional networks in dyslexia: a whole-brain, data-driven analysis of connectivity. *Biol*
359 *Psychiatry*. 2014;76(5):397–404.
- 360 Fraga González G, Van der Molen MJW, Žarić G, Bonte M, Tijms J, Blomert L, et al.
361 Corrigendum to “Graph analysis of EEG resting state functional networks in dyslexic
362 readers” [*Clin Neurophysiol*. 127(9) (2016) 3165–3175]. *Clin Neurophysiol*.
363 2018;129(1):339–40.
- 364 Fraga González G, van der Molen MW, Žarić G, Bonte M, Tijms J, Blomert L, et al. Graph
365 analysis of EEG resting state functional networks in dyslexic readers. *Clin Neurophysiol*.
366 2016;127(9):3165–75.
- 367 Hampson M, Tokoglu F, Sun Z, Schafer RJ, Skudlarski P, Gore JC, et al. Connectivity-behavior
368 analysis reveals that functional connectivity between left BA39 and Broca’s area varies with
369 reading ability. *Neuroimage*. 2006;31(2):513–9.
- 370 van den Heuvel MP, Hulshoff Pol HE. Exploring the brain network: a review on resting-state
371 fMRI functional connectivity. *Eur Neuropsychopharmacol*. 2010;20(8):519–34.
- 372 Hoeft F, McCandliss BD, Black JM, Gantman A, Zakerani N, Hulme C, et al. Neural systems
373 predicting long-term outcome in dyslexia. *Proc Natl Acad Sci U S A*. 2011;108(1):361–6.

- 374 Hoeft F, Ueno T, Reiss AL, Meyler A, Whitfield-Gabrieli S, Glover GH, et al. Prediction of
375 children's reading skills using behavioral, functional, and structural neuroimaging
376 measures. *Behav Neurosci.* 2007;121(3):602–13.
- 377 Hsu C-W, Chang C-C, Lin C-J. A Practical Guide to Support Vector Classification [Internet].
378 2003.
- 379 Kessler RC, van Loo HM, Wardenaar KJ, Bossarte RM, Brenner LA, Cai T, et al. Testing a
380 machine-learning algorithm to predict the persistence and severity of major depressive
381 disorder from baseline self-reports. *Mol Psychiatry.* 2016;21(10):1366–71.
- 382 Koyama MS, Kelly C, Shehzad Z, Penesetti D, Castellanos FX, Milham MP. Reading networks
383 at rest. *Cereb cortex.* 2010;20(11):2549–59.
- 384 Koyama MS, Di Martino A, Zuo X-N, Kelly C, Mennes M, Jutagir DR, et al. Resting-state
385 functional connectivity indexes reading competence in children and adults. *J Neurosci.*
386 2011;31(23):8617–24.
- 387 Leinonen S, Müller K, Leppänen PHT, Aro M, Ahonen T, Lyytinen H. Heterogeneity in adult
388 dyslexic readers : Relating processing skills to the speed and accuracy of oral text reading.
389 *Read Writ.* 2001;14(3–4):265–96.
- 390 Liu K, Shi L, Chen F, Waye MMY, Mok VCT, Chu WCW, et al. Altered topological
391 organization of brain structural network in Chinese children with developmental dyslexia.
392 *Neurosci Lett.* 2015;589:169–75.
- 393 van der Mark S, Klaver P, Bucher K, Maurer U, Schulz E, Brem S, et al. The left
394 occipitotemporal system in reading: disruption of focal fMRI connectivity to left inferior
395 frontal and inferior parietal language areas in children with dyslexia. *Neuroimage.*
396 2011;54(3):2426–36.
- 397 Martin A, Kronbichler M, Richlan F. Dyslexic brain activation abnormalities in deep and
398 shallow orthographies: A meta-analysis of 28 functional neuroimaging studies. *Hum Brain*
399 *Mapp.* 2016;37(7):2676–99.
- 400 Mazurowski MA, Habas PA, Zurada JM, Lo JY, Baker JA, Tourassi GD. Training neural
401 network classifiers for medical decision making: The effects of imbalanced datasets on
402 classification performance. *Neural Networks.* 2008;21(2–3):427–36.
- 403 Menghini D, Finzi A, Benassi M, Bolzani R, Facoetti A, Giovagnoli S, et al. Different
404 underlying neurocognitive deficits in developmental dyslexia: A comparative study.
405 *Neuropsychologia.* 2010;48(4):863–72.
- 406 Pacheco A, Reis A, Araújo S, Inácio F, Petersson KM, Faisca L. Dyslexia heterogeneity:
407 cognitive profiling of Portuguese children with dyslexia. *Read Writ.* 2014;27(9):1529–45.
- 408 Pugh KR, Mencl W, Shaywitz B, Shaywitz S, Fulbright RK, Constable RT, et al. The angular
409 gyrus in developmental dyslexia: task-specific differences in functional connectivity within
410 posterior cortex. *Psychol Sci.* 2000;11(1):51–6.
- 411 Pulini A, Kerr WT, Loo SK, Lenartowicz A. Classification accuracy of neuroimaging
412 biomarkers in Attention Deficit Hyperactivity Disorder: Effects of sample size and circular

- 413 analysis. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2018;
- 414 Richlan F, Kronbichler M, Wimmer H. Functional abnormalities in the dyslexic brain: a
415 quantitative meta-analysis of neuroimaging studies. *Hum Brain Mapp*. 2009;30(10):3299–
416 308.
- 417 Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and
418 interpretations. *Neuroimage*. 2010;52(3):1059–69.
- 419 Schurz M, Wimmer H, Richlan F, Ludersdorfer P, Klackl J, Kronbichler M. Resting-State and
420 Task-Based Functional Brain Connectivity in Developmental Dyslexia. *Cereb cortex*.
421 2014;184.
- 422 Shaywitz S, Shaywitz B, Fulbright RK, Skudlarski P, Mencl W, Constable RT, et al. Neural
423 systems for compensation and persistence: young adult outcome of childhood reading
424 disability. *Biol Psychiatry*. 2003;54(1):25–33.
- 425 Snowling MJ. Early identification and interventions for dyslexia: a contemporary view. *J Res*
426 *Spec Educ Needs*. 2013;13(1):7–14.
- 427 Stam CJ. Modern network science of neurological disorders. *Nat Rev Neurosci*.
428 2014;15(10):683–95.
- 429 Stam CJ, Nolte G, Daffertshofer A. Phase lag index: assessment of functional connectivity from
430 multi channel EEG and MEG with diminished bias from common sources. *Hum Brain*
431 *Mapp*. 2007;28(11):1178–93.
- 432 Tamboer P, Vorst HCM, Ghebreab S, Scholte HS. Machine learning and dyslexia: Classification
433 of individual structural neuro-imaging scans of students with and without dyslexia.
434 *NeuroImage Clin*. 2016;11:508–14.
- 435 Tanaka H, Black JM, Hulme C, Stanley LM, Kesler SR, Whitfield-Gabrieli S, et al. The Brain
436 Basis of the Phonological Deficit in Dyslexia Is Independent of IQ. *Psychol Sci*.
437 2011;22(11):1442–51.
- 438 Vandermosten M, Boets B, Wouters J, Ghesquière P. A qualitative and quantitative review of
439 diffusion tensor imaging studies in reading and dyslexia. *Neurosci Biobehav Rev*.
440 2012;36(6):1532–52.
- 441 Vourkas M, Micheloyannis S, Simos PG, Rezaie R, Fletcher JM, Cirino PT, et al. Dynamic task-
442 specific brain network connectivity in children with severe reading difficulties. *Neurosci*
443 *Lett*. 2011;488(2):123–8.
- 444 Wolf RC, Sambataro F, Lohr C, Steinbrink C, Martin C, Vasic N. Functional brain network
445 abnormalities during verbal working memory performance in adolescents and young adults
446 with dyslexia. *Neuropsychologia*. 2010;48(1):309–18.
- 447 Zare M, Rezvani Z, Benasich AA. Automatic Classification of 6-month-old Infants at Familial
448 Risk for Language-based Learning Disorder Using a Support Vector Machine. *Clin*
449 *Neurophysiol*. 2016;
- 450

451 Supporting information

S1 Table. Sample characteristics and descriptive statistics of reading accuracy and fluency scores.

	Typical Readers <i>M (SD)</i>	Dyslexics <i>M (SD)</i>	<i>p</i> -value	η^2
N	15	29		
Sex ratio (m:f)	6:9	16:13		
Handedness (L:R)*	2:10	2:27		
Age	8.75 (0.31)	8.96 (0.40)	.088	0.07
RAVEN – IQ test^a	6.70 (1.51)	7.11 (1.51)	.395	0.02
3DM Word reading - accuracy^b				
High Frequency	99.28 (1.05)	93.10 (5.93)	.000	0.27
Low Frequency	98.32 (2.54)	86.31 (14.48)	.003	0.19
Pseudo	88.70 (8.48)	73.33 (17.43)	.003	0.20
Total [T]^c	51.40 (8.00)	33.72 (12.58)	.000	0.37
3DM Word reading - fluency [T]				
High Frequency	54.27 (7.58)	31.38 (6.14)	.000	0.74
Low Frequency	56.80 (8.98)	32.07 (6.46)	.000	0.72
Pseudo	54.93 (9.71)	30.93 (6.37)	.000	0.70
Total	55.93 (9.51)	31.00 (5.40)	.000	0.75
One-Minute Test -fluency [SS]^d	12.07 (2.94)	3.97 (1.97)	.000	0.74
Text Reading - fluency[T]**	55.27 (8.41)	33.21 (6.30)	.000	0.70
3DM Spelling - accuracy[T]	51.73 (8.62)	36.21 (6.70)	.000	0.51
3DM Spelling - fluency[T]	54.33 (9.90)	36.55 (6.01)	.000	0.57
3DM Phoneme deletion - accuracy [T]**	53.73 (8.39)	39.61 (8.32)	.000	0.41
Letter-speech sound associations [T]				
L-SS identificacion - accuracy	46.87 (8.65)	43.34 (12.99)	.350	0.02
L-SS discrimination – accuracy**	50.80 (10.28)	44.43 (9.63)	.050	0.09
L-SS identificacion - fluency	51.53 (7.67)	41.79 (6.97)	.000	0.30
L-SS discrimination - fluency**	51.73 (7.36)	45.46 (8.95)	.025	0.12
3DM Naming speed scores[T]**				
Letters	50.93 (6.95)	36.57 (8.05)	.000	0.45
numbers	52.73 (10.67)	36.21 (8.50)	.000	0.43
Total	50.80 (7.73)	35.54 (9.15)	.000	0.42

^a C scores (M = 5, SD = 2). ^b Raw scores. ^c T scores (M = 50, SD = 10). ^d SS scores (M = 10, SD= 3).

*Data missing for 3 participants; Typical N = 12. ** Data missing for one participant; Dyslexics N = 28