

# Review of: "Dyslexia biomarker detection with Quantitative electroencephalography (QEEG) data in children: Feasibility, Acceptability, Economic impact, Pilot Study and Survey Results"

Emanuele Casani<sup>1</sup>

<sup>1</sup> Ca' Foscari University of Venice

**Potential competing interests:** No potential competing interests to declare.

This is an ambitious study that investigates the use of a mobile app to self-detect in a few minutes a biomarker of developmental dyslexia among children aged 7 to 10 years.

Firstly, I congratulate the authors on their revisions, which have improved the article dramatically since the first version. The objectives are more realistic now. The title is clearer and more consistent with the abstract and content. More relevant and up-to-date references have been added. Terminology is clearer, including acronyms, which have been expanded and briefly explained. Nevertheless, I still find some substantial and formal weaknesses. The former concern coherence between objectives, results, and conclusions, which should be tuned. The latter concern phrasing and some stylistic issues, which might be improved by submitting the draft to a mother-language speaker. Below are some comments on specific sections.

## Title

If “feasibility, acceptability” and “economic impact” were tested through the survey, there is no need to add the phrase “survey results”. So the title would be more concise and clear.

Please check the use of capital letters.

## Introduction

The authors should clarify that they are dealing with developmental dyslexia instead of referring to “dyslexia” generally.

We have no solid evidence that neuroimaging techniques could replace behavioral diagnoses, so the study impact appears to be overestimated. The authors could consider different theories underpinning developmental-dyslexia research to set the study in a more specific frame. This would probably increase its credibility.

The claims “fMRI scans and eye tracking data yield higher accuracy results than other methods” and “algorithms that use EEG scans do not yield a high accuracy result to go to the market” should be supported by references.

Please use a consistent style to report data (e.g., “89.6% and 89.5 percent” [percent → %]).

There is no need to specify H0 and H1 after defining research questions, especially if H0 and H1 do not add any information.

Research questions could be clearer. In RQ4, for instance, the authors should clarify that they are referring to economic affordability.

## Method

### ***Participants***

It is not clear how controls were classified as typically developing, and how the absence of comorbidities was assessed in children with developmental dyslexia. Parental surveys could be insufficiently reliable for this purpose. Are the authors aware of the possible presence of additional learning difficulties besides reading problems? Please provide as much information as possible about the participants.

The statement “Individuals were instructed to perform a 2-minute resting state EEG measurement for the purpose of data collection” would fit better in the “Procedures” section.

There is no need to reiterate that the “study uses a small sample size” after specifying the participants’ number (which does not seem too small, however). All the more so because this aspect is included in the study limitations (see comments below, too).

The authors correctly specified that they used repeated measures (which was not so clear in the previous versions). But this information would fit better in the “Analyses” section, maybe with some additional information allowing for replication of their models.

### ***Statistical analyses***

Please expand the acronym “ML” at the first occurrence.

Some phrasing might be improved (e.g., “the children with dyslexia and typically developing children’s data...” → “Data of children with dyslexia and typical development...”).

Please make verb tenses consistent (e.g., “The missing values were replaced [...] The data is labeled [...] children’s data were balanced [...] The binary classification [...] is applied”).

It is not necessary to explain the meaning of *true/false positives/negatives*, *sensitivity*, and *specificity*. As the authors do so, I would suggest more transparency. For instance, they could specify that “positive/negative” means “with/without developmental dyslexia”, as well as what “entities” stands for.

No need to specify that Pearson/Spearman correlations are parametric/non-parametric tests.

## Results

This section might be divided into subheadings to increase its readability.

“Repeated measure assessment” would sound clearer than “repeated assessment”.

As the authors say “which is the state-of-the-art in literature for dyslexia biomarker detection with EEG scans”, references would be welcome. The nominal phrase “state of the art” should not be hyphenated.

Please expand the acronym AUC at the first occurrence (Table 1).

The statement “the model states that 27 children with dyslexia sessions as TDC sessions out of 8301 samples” should be rephrased. In general, there is some confusion between “sessions” and “children”, which makes this section unclear. I would suggest revising it. More accurate captions would also help the reader interpret Figure 3.

I do not agree with the statement “misclassifying TDC sessions as children with dyslexia -false positives- sessions would not create problems”, as this misclassification affects the tool specificity.

The authors should increase consistency between research questions and results. For instance, RQ2, concerning feasibility, does not match the fact that “the results of diagnosis correspond to that of the psychiatrists”, which has to do with the tool validity. Feasibility would rather match the following outcome, i.e., “82% of the respondents find the app easily usable at home”.

## Discussion

Please check the use of “let to” in the 1<sup>st</sup> line.

The claims that “the left lateralization of the cortex is delayed for children with dyslexia and they use the right hemisphere” and “the left lateralization of the cortex is established for TDC and they use the left hemisphere” are important results, which would deserve further discussion and possible references to the literature.

The fact that “The biomarker detection app catches [please check the verb tense] non-diagnosed dyslexic cases” should be discussed further.

I suggest caution in stating that “technology is ready to diagnose dyslexia with a high accuracy rate and improve the symptoms of dyslexia with mobile health solutions”. At the moment, we have no evidence that any technological tool might replace thorough behavioral diagnoses. I would mitigate (and support with references) this sentence, which does not take into account the multifactorial nature of developmental dyslexia as well as its possible profiles of variability. As I suggested in the “Introduction” section, the study contextualization in a narrower frame might increase its impact credibility.

Reflections about future developments might be moved to the “Conclusions” section. As the authors expect to use “ML methods to predict [...] the level/scale of dyslexia” in the future, they should provide a frame of reference and some additional details in this respect.

## ***Limitations of the Study and Conclusions***

In the statement “it only includes 207”, the authors should specify they mean participants. As the sample size is their recurring concern, they could specify how many participants such a study would need by referring to a power analysis or previous research.

As the authors say that “the control group was not given an alternative intervention”, so “placebo effects might be a substantial source of improvement”, it is not clear which kind of intervention they are referring to. Does the study aim to test “the feasibility, acceptability, and economic impact of the dyslexia biomarker detection app” or to analyze a longitudinal remedial intervention? It is not clear. Tuning objectives, results, and conclusions would be crucial.