Pitfalls in using ML to predict executive function performance by linguistic variables

Gianna Kuhles[1,2], Julia A. Camilleri[1,2], Sami Hamdan[1,2], Stefan Heim[3,4], Simon B. Eickhoff[1,2], Kaustubh R. Patil[1,2] & Susanne Weis[1,2]

[1]Institute of Systems Neuroscience, Heinrich Heine University Düsseldorf, Düsseldorf, Germany;
[2]Institute of Neuroscience and Medicine (INM-7: Brain and Behaviour), Research Centre Jülich, Jülich, Germany
[3]Department of Psychiatry, Psychotherapy und Psychosomatics, Medical Faculty, RWTH Aachen University, Aachen, Germany.
[4]Institute of Neuroscience and Medicine (INM-1 Structural and functional organization of the brain), Research Center Jülich, Jülich, Germany.
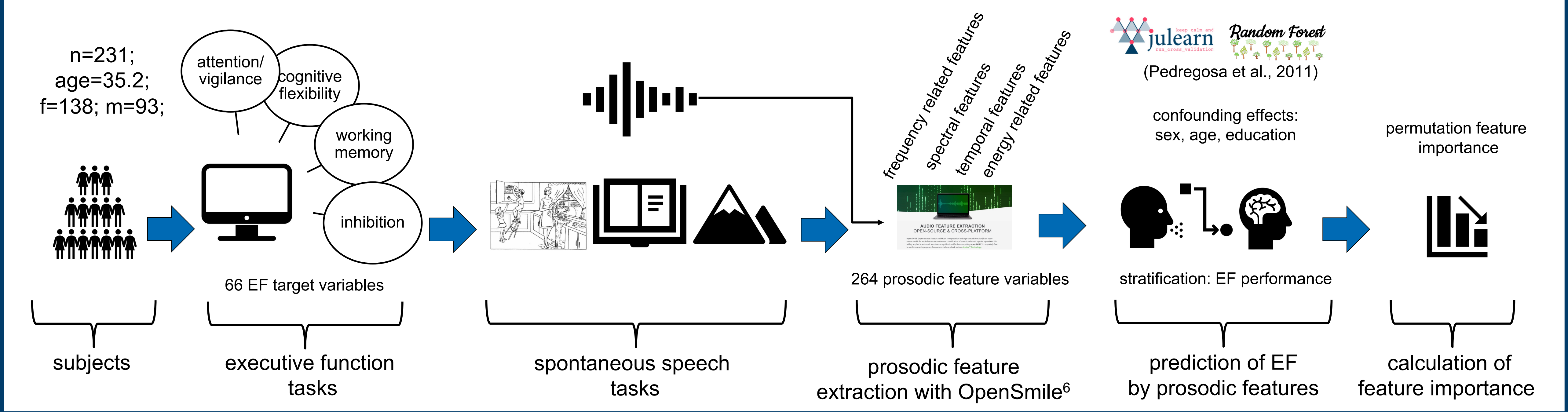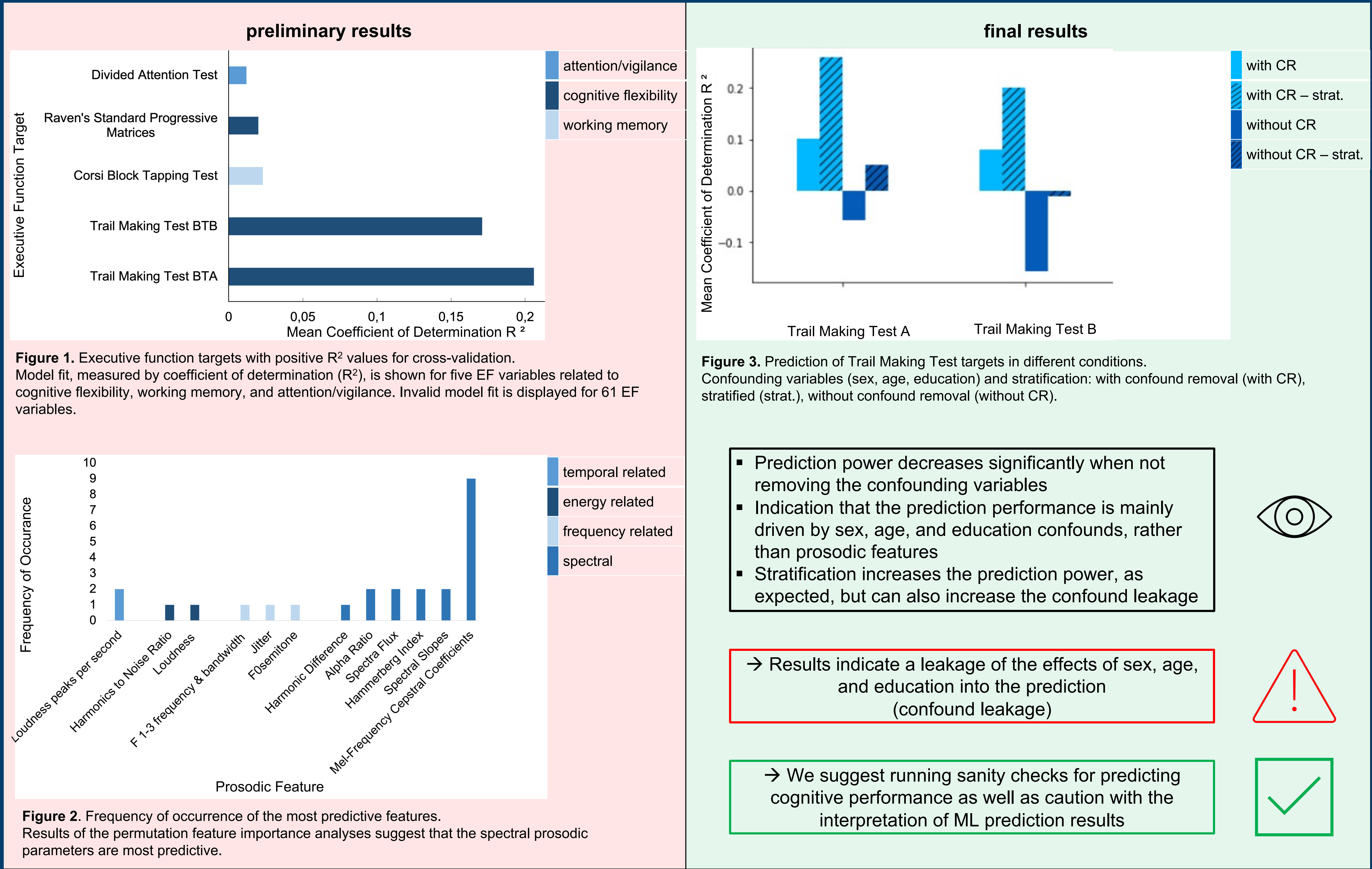
Introduction

- Link between executive function (EF) performance and prosody in numerous mental disorders (Filipe et al., 2018; Le et al., 2011; Nevler et al., 2017)

- Most findings based on patient studies, not on healthy subjects - unclear how different subdomains of EF and prosody are related to each other

- Aim of present study: Explore the relationships of EF and prosody using a machine learning (ML) regression approach by predicting EF performance from various prosodic features

Methods

n=231; age=35.2; f=138; m=93;

subjects — executive function tasks (66 EF target variables) — spontaneous speech tasks — prosodic feature extraction with OpenSmile[6] (264 prosodic feature variables) — prediction of EF by prosodic features (confounding effects: sex, age, education; stratification: EF performance) — calculation of feature importance (permutation feature importance)

(Pedregosa et al., 2011)

Results

preliminary results

Figure 1. Executive function targets with positive R^2 values for cross-validation. Model fit, measured by coefficient of determination (R^2), is shown for five EF variables related to cognitive flexibility, working memory, and attention/vigilance. Invalid model fit is displayed for 61 EF variables.

Figure 2. Frequency of occurrence of the most predictive features. Results of the permutation feature importance analyses suggest that the spectral prosodic parameters are most predictive.

final results

Figure 3. Prediction of Trail Making Test targets in different conditions. Confounding variables (sex, age, education) and stratification: with confound removal (with CR), stratified (strat.), without confound removal (without CR).

- Prediction power decreases significantly when not removing the confounding variables
- Indication that the prediction performance is mainly driven by sex, age, and education confounds, rather than prosodic features
- Stratification increases the prediction power, as expected, but can also increase the confound leakage

→ Results indicate a leakage of the effects of sex, age, and education into the prediction (confound leakage)

→ We suggest running sanity checks for predicting cognitive performance as well as caution with the interpretation of ML prediction results

Discussion

- Study provides insights into the specific relationships between prosody (mainly spectral features) and executive function performance

- Findings about connection between prosody and EF are in line with previous studies (Filipe et al., 2018; Le et al., 2011; Nevler et al., 2017)

- Further research should investigate whether the predictive power of prosody can serve as a biomarker of executive dysfunction

- **Main outcome:** Our findings examplify possible pitfalls with the use of ML & strongly suggest caution with the interpretation of ML prediction results

References:
- Filipe M G et al. Front. Psych. (2018) 9: 359.
- Le P et al. Speech Comm. (2011) 54: 540-551.
- Nevler N et al. Neurol. (2017) 89: 650-656.
- Pedregosa et al. JMLR (2011) 12: 2825-2830.