

# Using acoustic and linguistic markers from spontaneous speech to predict scores on the Montreal Cognitive Assessment (MoCA)

Aparna Balagopalan, MSc(c)<sup>2,3</sup>, Maria Yancheva, MSc<sup>2</sup>, Jekaterina Novikova, PhD<sup>2</sup>, William Simpson, PhD<sup>1,2</sup>

(1) Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada, (2) Winterlight Labs, Toronto, ON, Canada, (3) Department of Computer Science, University of Toronto, Toronto, ON, Canada

## Background

Recent clinical trials in Alzheimer's disease have been overwhelmingly negative. This has spurred development of composite metrics and novel biomarkers with the intent of more accurately capturing changes in cognitive function. Computational analysis of speech and language represents one such group of biomarkers. The objective of this study was to examine the construct validity of speech and language markers by using them to predict scores on the MoCA.

## Methods

- 219 individuals, aged 55-95, were recruited from the community and senior care facilities as part of a normative data collection study for a digital cognitive assessment.
- A picture description task and a MoCA were administered by a trained psychometrist.
- Using natural language processing techniques, audio samples and transcripts were analyzed to generate markers pertaining to sample acoustics, syntax, grammar, cohesion, discourse and content described.
- Feature selection and linear regression with 10-fold cross validation were used to generate an estimated MoCA Total Score
- Using the same methodology, estimates were generated for MoCA index scores – namely: Memory, Executive Functioning, Visuospatial, Attention, Language, and Orientation.

## Results

Metric	Mean	SD	Value Range	Scale Range
MoCA Total	24.4	4.0	10-30	0-30
Memory Index	11.2	3.7	0-15	0-15
Executive Function Index	10.9	2.2	1-13	0-13
Visuospatial Index	5.9	1.0	3-7	0-7
Attention Index	15.8	2.7	4-18	0-18
Language Index	4.9	1.1	2-6	0-6
Orientation Index	5.8	0.6	3-6	0-6

Figure 1: Distribution of errors in MoCA Total Score Predictions

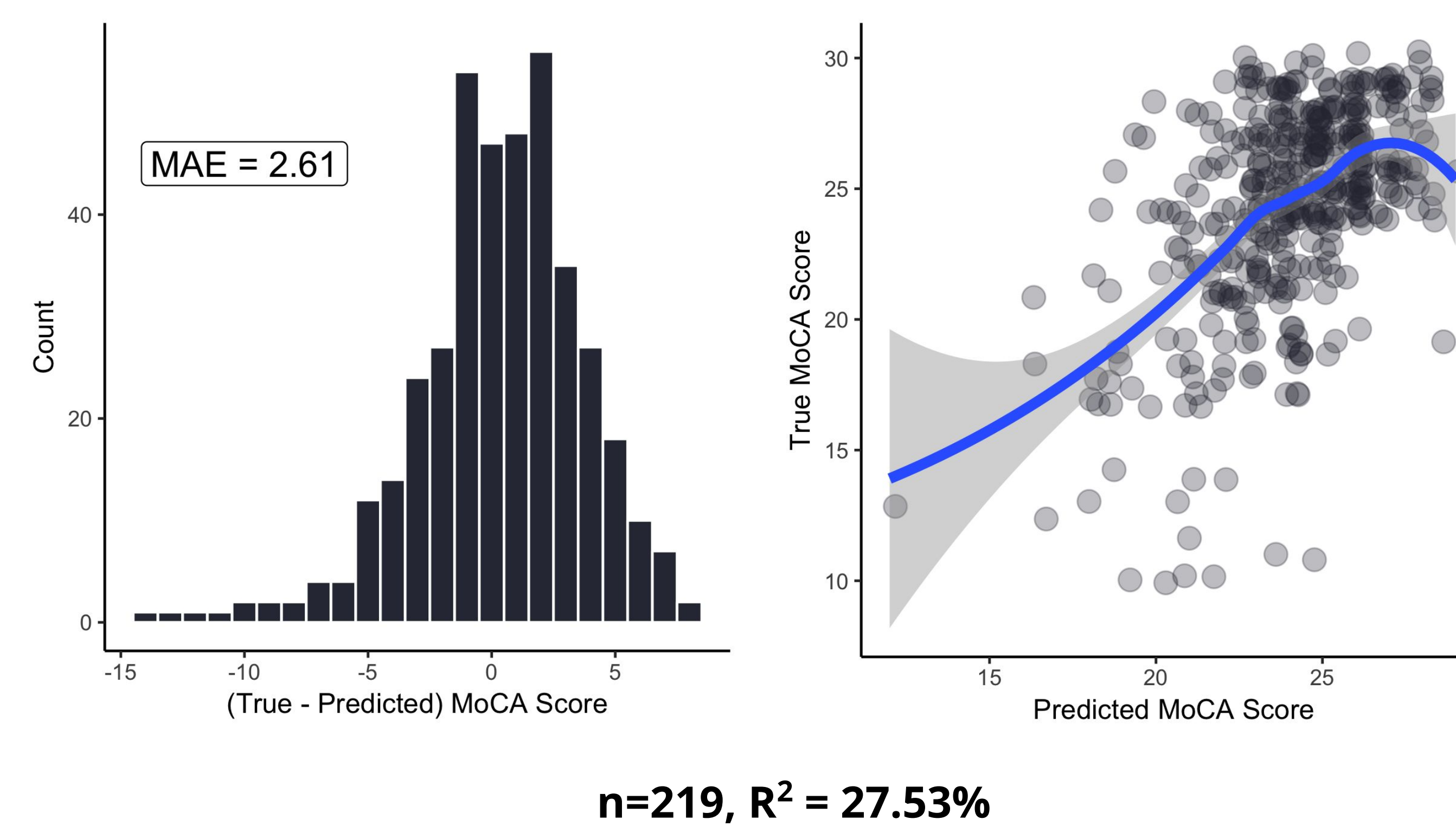
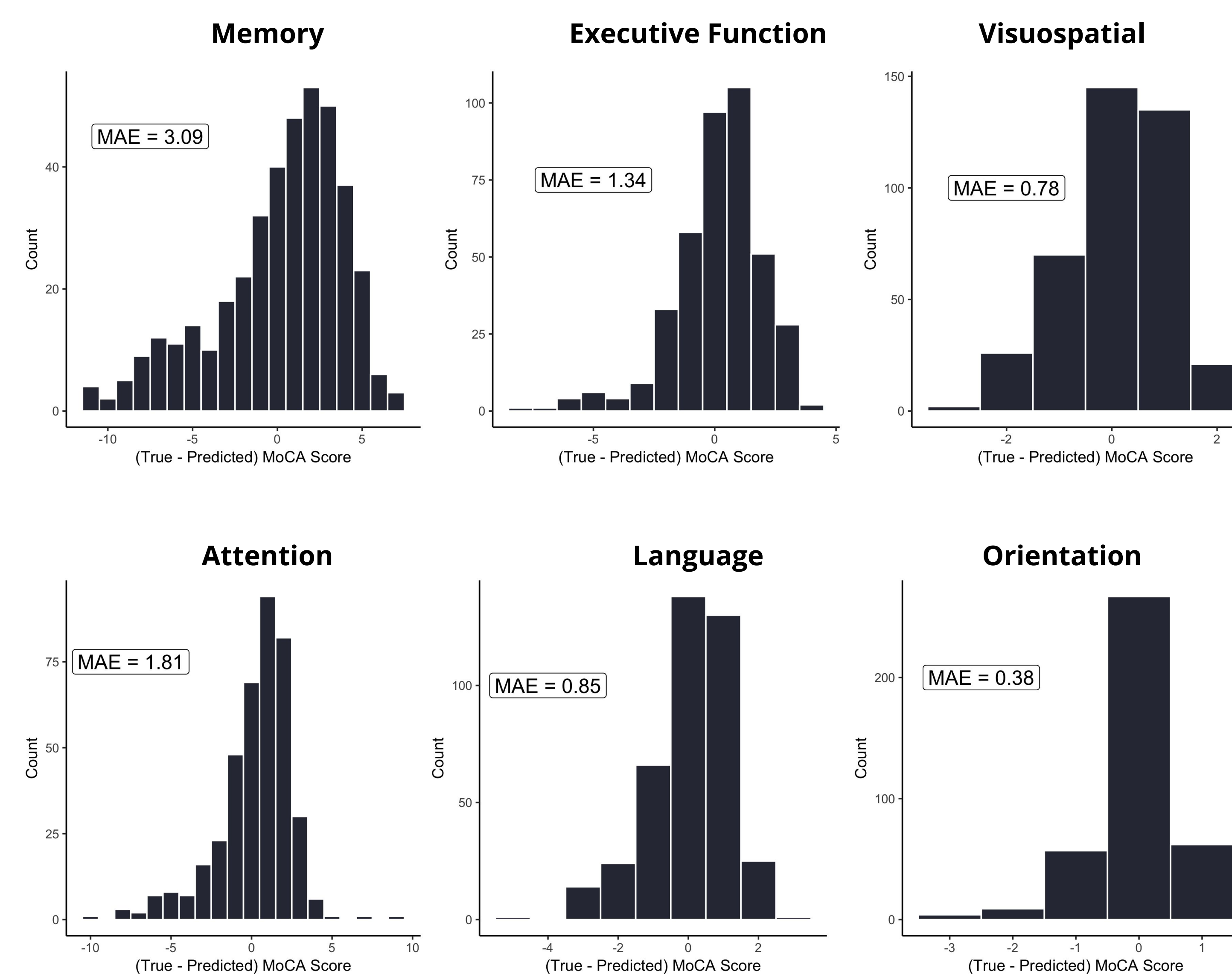


Figure 2: Distribution of errors in MoCA Index Score predictions



## Results Continued

Metric	MAE (%)	N Markers	Important Markers
MoCA Total	8.7%	15	Speech duration, rate, pauses, discourse quality, spectral composition
Memory	20.6%	13	Speech duration, rate, type-token ratio, spectral composition
Executive	10.4%	12	Word length, content units, speech rate, coherence, spectral composition
Visuospatial	11.2%	9	Coherence, pauses, spectral composition
Attention	10.1%	5	Discourse, spectral composition
Language	14.3%	9	Speech rate, syntax, content units, spectral composition
Orientation	6.4%	4	Spectral Composition

## Conclusions

The low error in total score prediction and moderate  $R^2$  suggests that speech and language variables obtained from the picture description task map to the constructs measured by the MoCA. This supports the further development of novel, speech-based digital biomarkers, and composite metrics for assessing cognitive status.

## References

- Slegers, A., Filiou, R.-P., Montembeault, M. & Brambati, S. M. Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *J. Alzheimers. Dis.* 65, 519-542 (2018).
- Yancheva, M., Fraser, K. & Rudzicz, F. Using linguistic features longitudinally to predict clinical scores for Alzheimer's disease and related dementias. in *Proceedings of SLPAT 2015: 6th Workshop on Speech and Language Processing for Assistive Technologies (2015)*. doi:10.18653/v1/w15-5123
- Fraser, K. C., Meltzer, J. A. & Rudzicz, F. Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *J. Alzheimers. Dis.* 49, 407-422 (2016).