TOMASZ M. RUTKOWSKI
RESEARCH SCIENTIST IN COGNITIVE BEHAVIORAL ASSISTIVE TECHNOLOGY TEAM AT RIKEN AIP, TOMASZ.RUTKOWSKI@RIKEN.JP
RESEARCH FELLOW AT THE UNIVERSITY OF TOKYO, TOKYO, JAPAN, TOMEK@BCI-LAB.INFO
MIHOKO OTAKE-MATSUURA AND QIBIN ZHAO
RIKEN AIP, JAPAN

MACHINE LEARNING APPROACHES IN BRAIN CORRELATES OF DEMENTIA ELUCIDATION – TENSOR MACHINE LEARNING AND BEYOND

AI for Social Good

AIP RIKEN

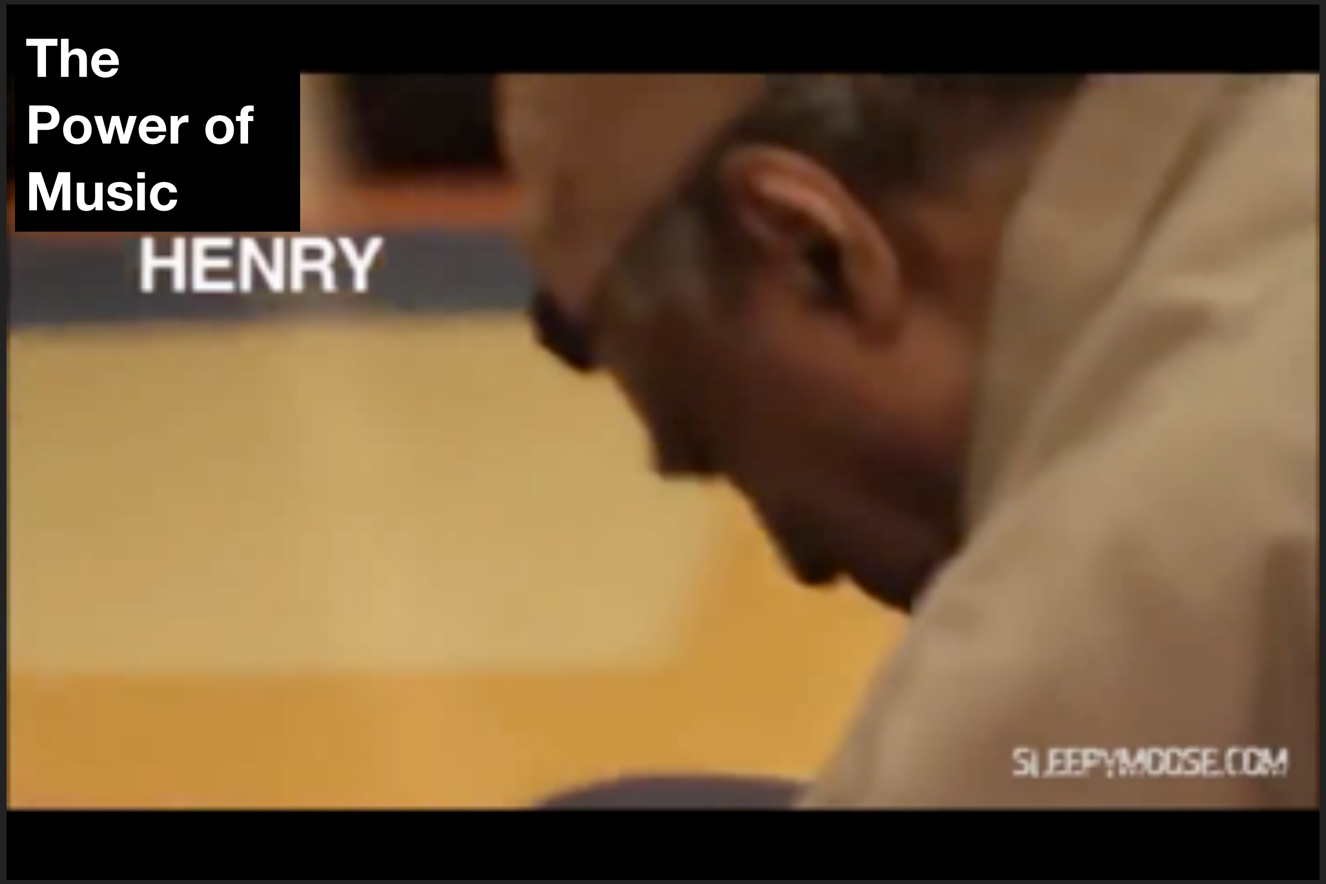
COGNITIVE BEHAVIORAL ASSISTIVE TECHNOLOGY TEAM

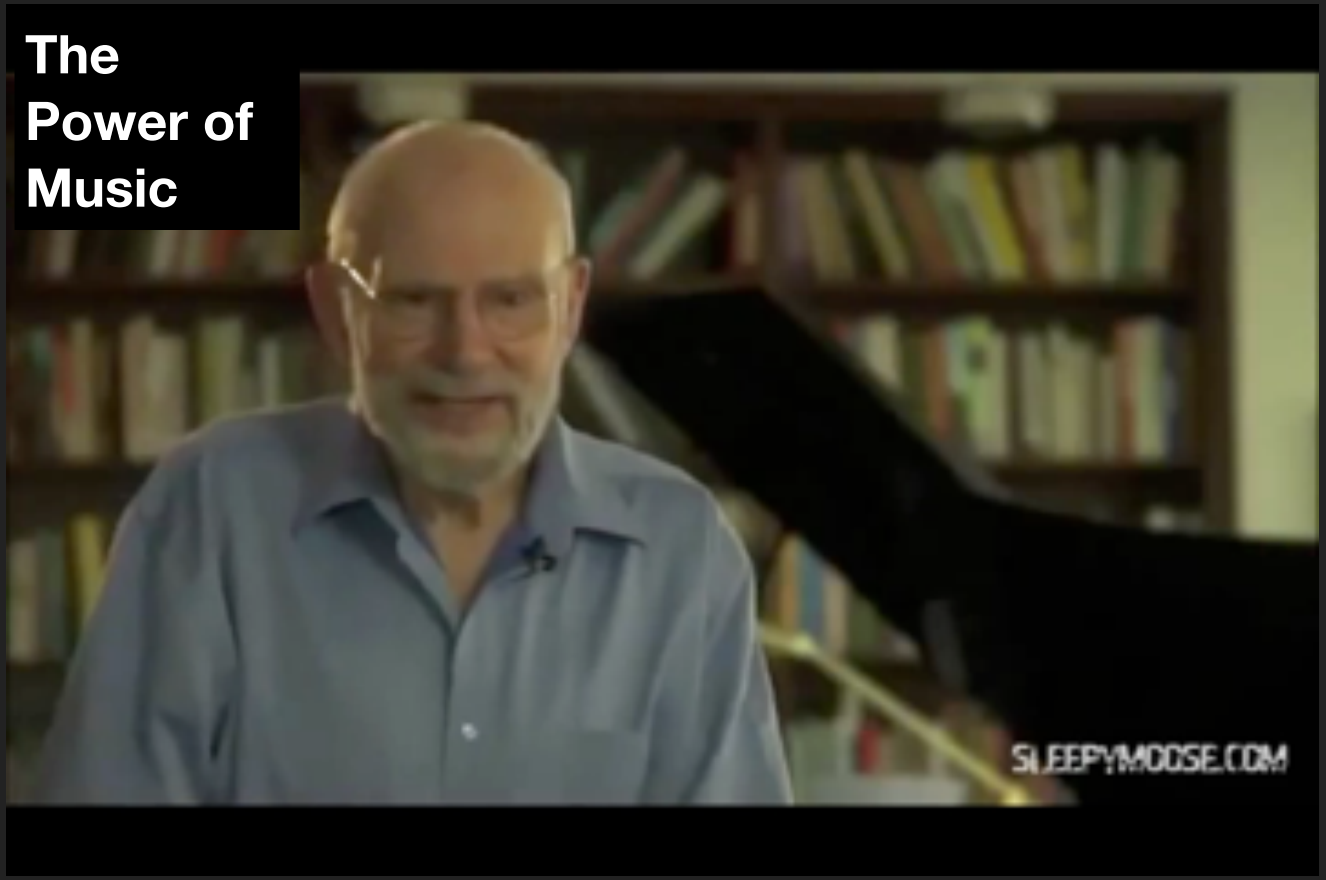
ARTIFICIAL INTELLIGENCE (AI)
VERSUS
INTELLIGENCE AUGMENTATION (IA)

- **Artificial intelligence (AI)** = reproducing of human cognition
- **Intelligence augmentation (IA)** = supplementing and supporting of human thinking, analysis, and planning, leaving the intentionality of a human actor at the heart of the human-computer interaction

An artists with Alzheimer's...

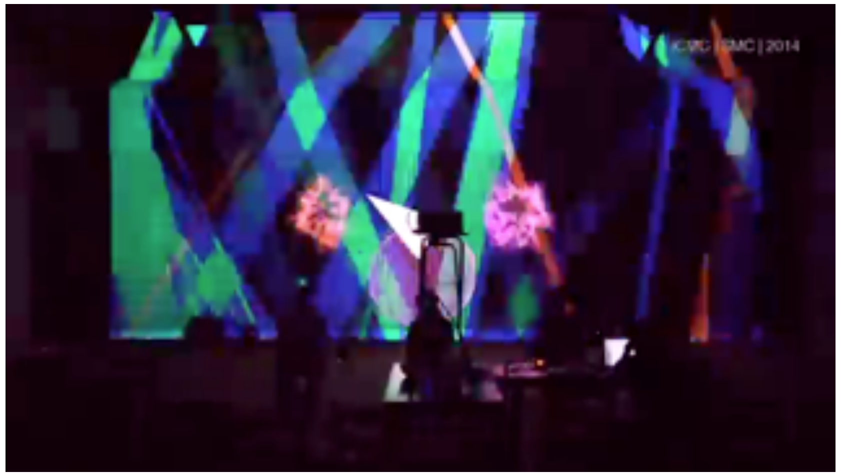
THE POWER OF SENSORY STIMULATION ON DEMENTIA

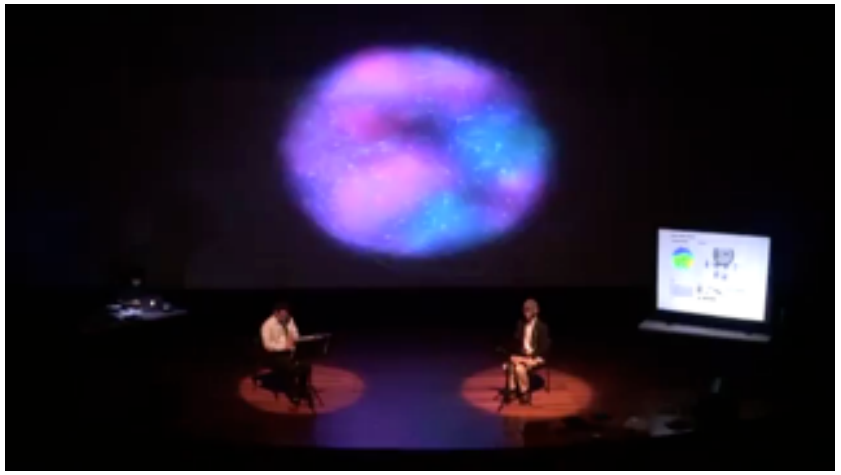




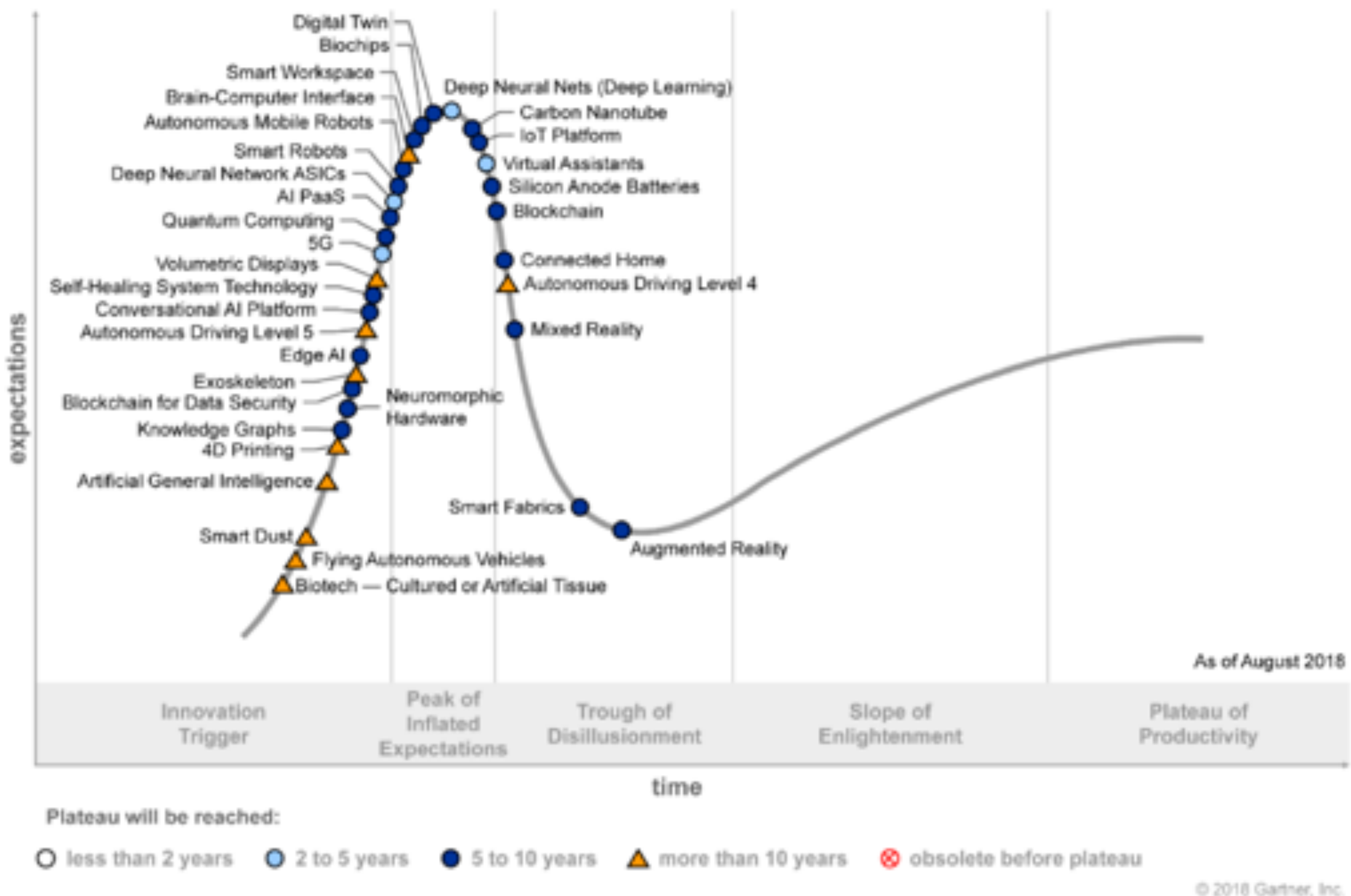
“BRAIN DREAMS MUSIC” PROJECT AT RIKEN BSI

HUMAN BRAIN MEETS AI TO COLLABORATE ON MUSIC COMPOSITION AND PERFORMANCE



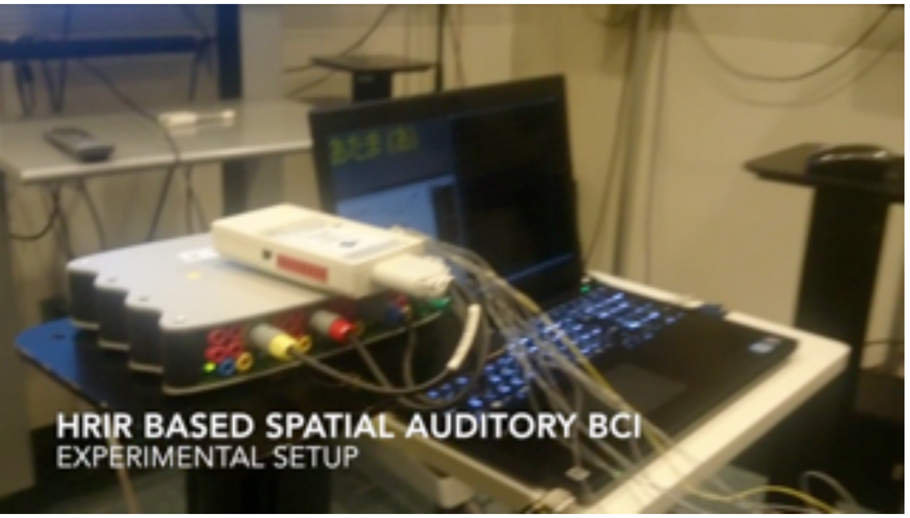


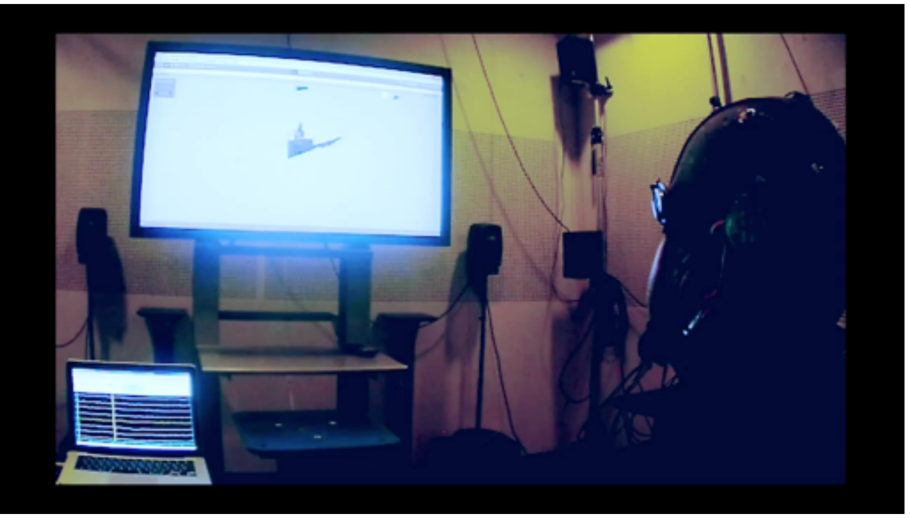
GARTNER HYPE CYCLE FOR EMERGING TECHNOLOGIES, 2018

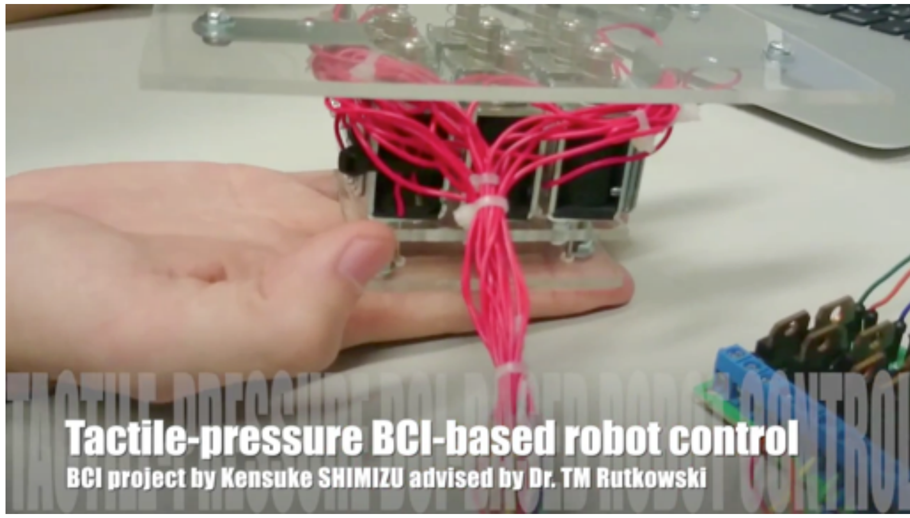


OLD SCHOOL BCI/NEUROTECH

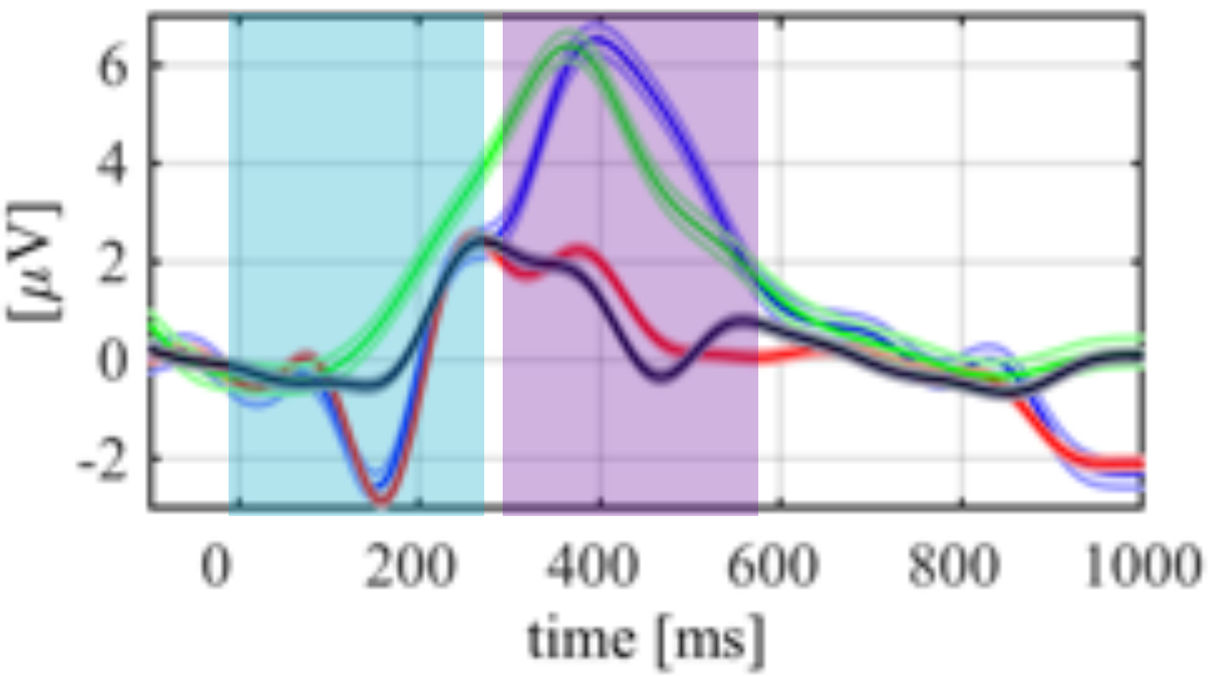
MY PASSIVE BCI/NEUROTECH





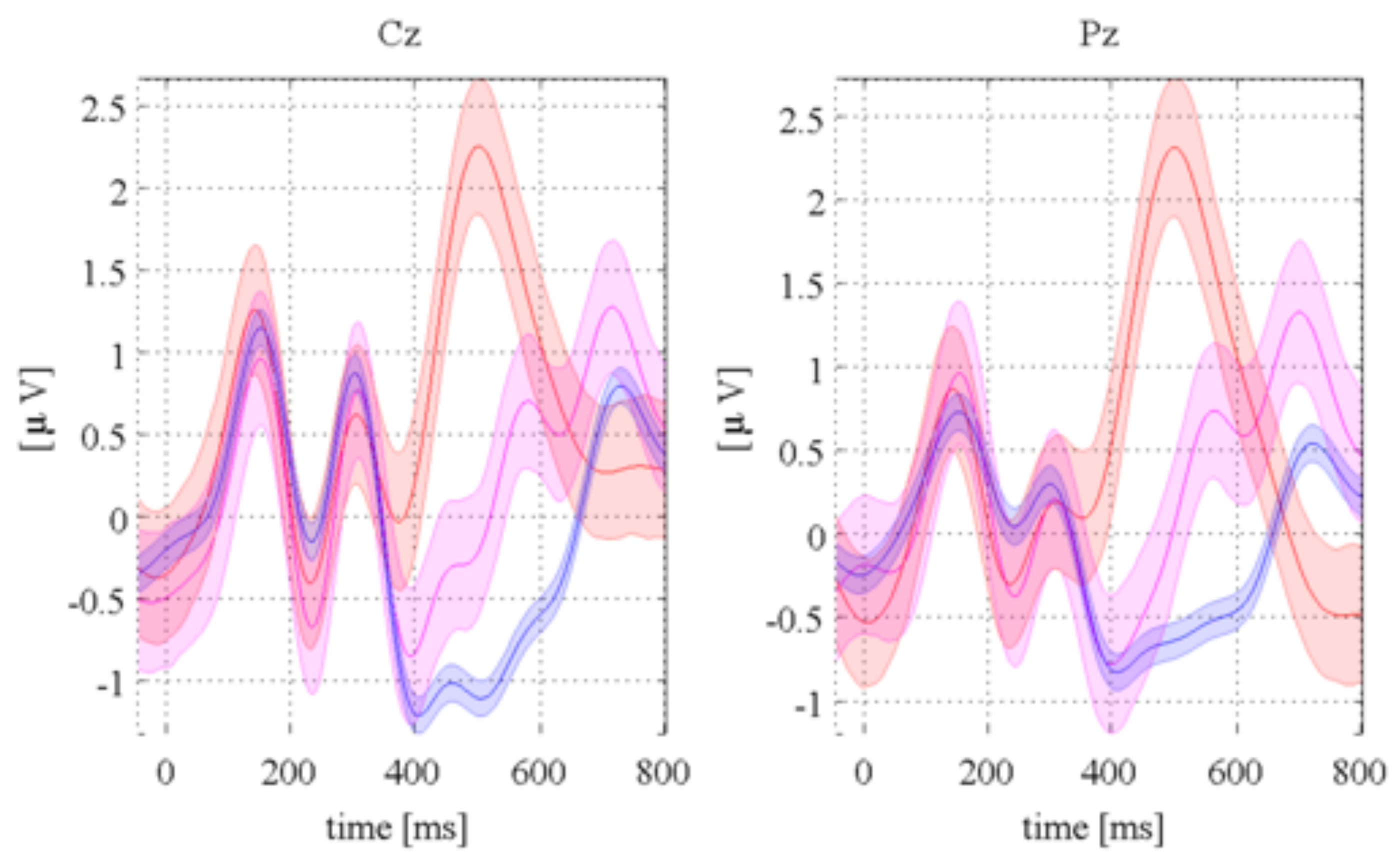


SIGNAL PROCESSING FOR MODALITY AND TASK DEPENDENT NON-LINEAR OR NON-STATIONARY BRAINWAVE ANALYSIS AND NON-BRAIN INTERFERENCE REMOVAL IN REAL TIME



- green = **attended** **vibrotactile**
- black = ignored **vibrotactile**
- blue = **attended** ultrasound tactile
- red = ignored ultrasound tactile

- red = **attended** real sound
- purple = **attended** virtual sound
- blue = ignored sounds

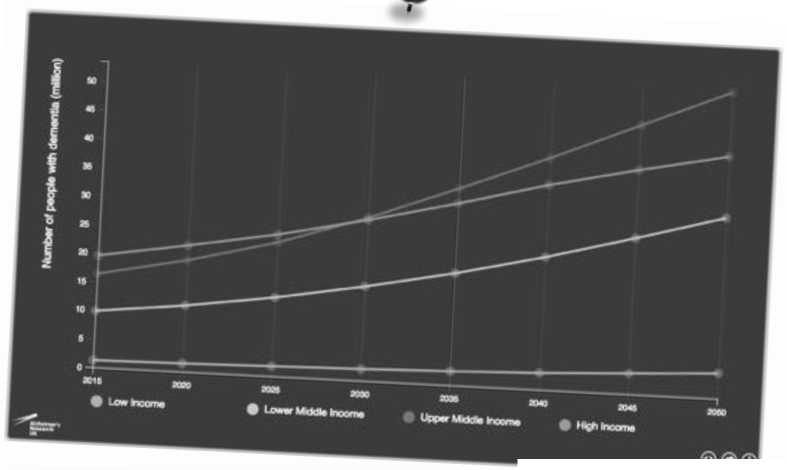


BRAIN–COMPUTER INTERFACE (BCI)

- **reactive BCI –** user intentions are decoded from the brainwaves in realtime using a non-invasive electroencephalography (EEG) and they are translated to a symbiotic robot or VR/AR agent thought-based only control
- **active BCI** – utilizes user intentional modulations of brainwaves in response to self–paced mental processes
- **passive BCI –** which analyzes an arbitrary brain activity and allows for biomarkers monitoring of elderly users for potential dementia progress elucidation

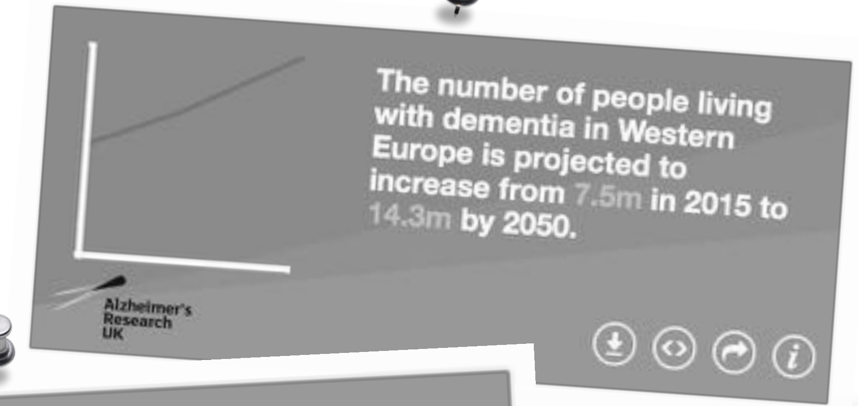
DEMENTIA FACTS

- Dementia and especially Alzheimer's disease (AD) are the most common cause of **cognitive decline in the elderly creating significant medical and economic burden**
- Approximately **47 millions of people** live with neurocognitive disorders worldwide, and **this number is expected to triple by 2050**
- Recent research findings in neuroscience have illustrated that **abnormal tau and amyloid** could modify the pre- and postsynaptic neuronal mechanisms, elevate the neuronal calcium influx leading to **increased excitability, neuronal loss and altered rhythmic patterns**
- Evidence from recent neurophysiological and pathological studies suggests that **impaired synaptic plasticity is a dominant feature of AD**, which is essential in all complex cognitive functions such as learning, abstract thinking and memory
- Dementia and AD has recently been described as a network disconnection disease or as **an oscillopathy**



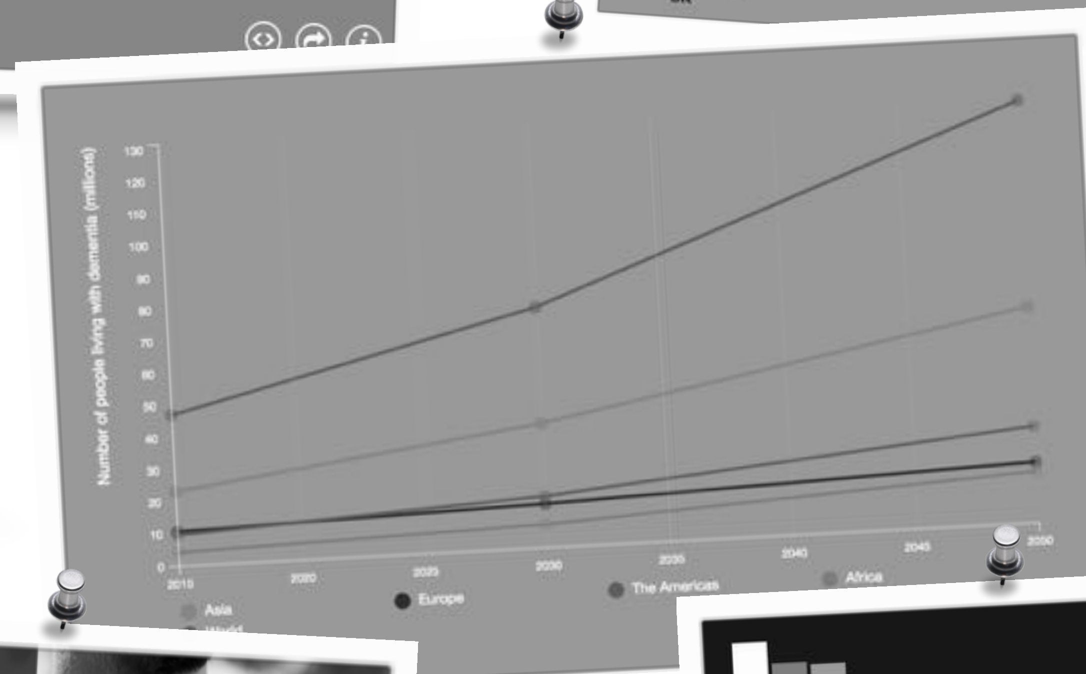
Number of people with dementia (million)
Low Income
Lower Middle Income
Upper Middle Income
High Income

58% of all people with dementia live in lower middle income countries – this will rise to 68% by 2050.
Alzheimer's Research UK



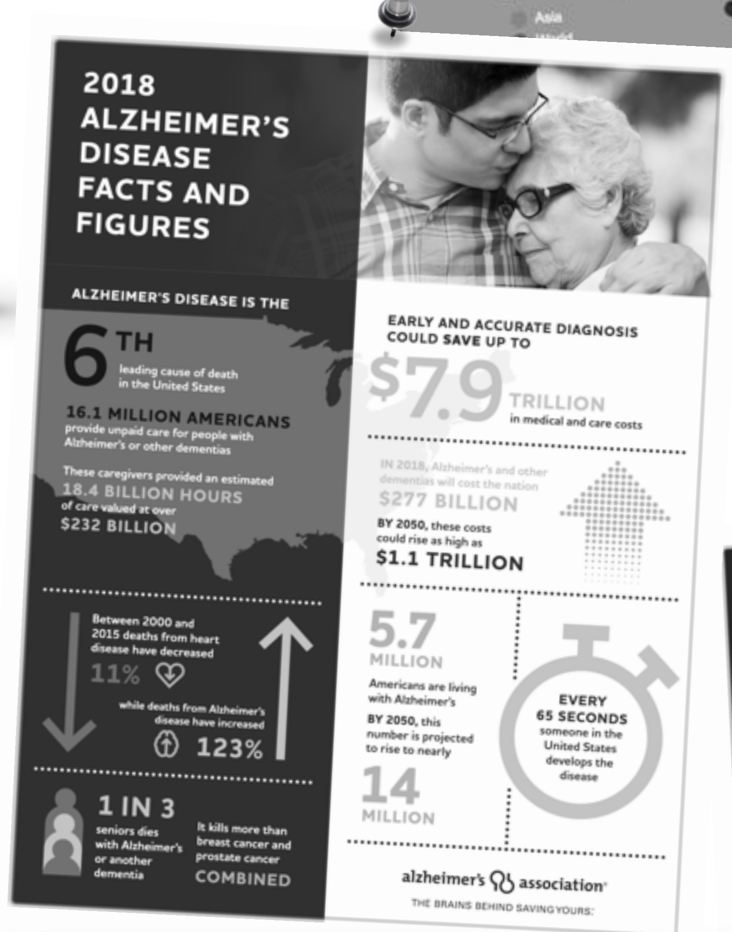
The number of people living with dementia in Western Europe is projected to increase from 7.5m in 2015 to 14.3m by 2050.
Alzheimer's Research UK

5.2%
of people over the age of 60 are living with dementia globally.
Alzheimer's Research

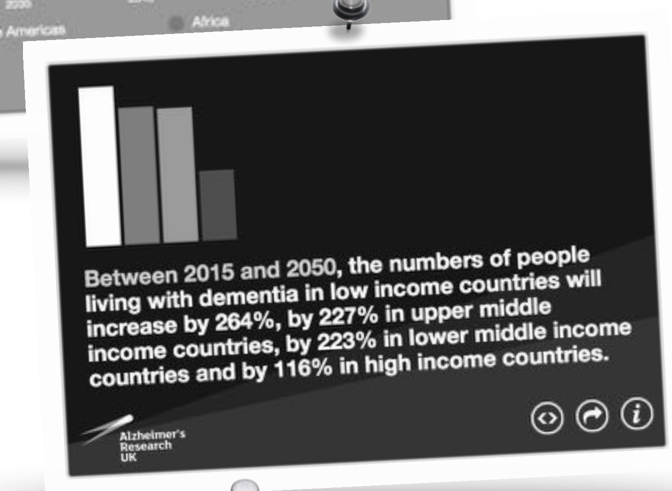


Number of people living with dementia (millions)
Asia
Europe
The Americas
Africa

East Asia is the region with the most people living with dementia (9.8m), followed by Western Europe (7.5m), South Asia (5.1m) and North America (4.8m).
Alzheimer's Research UK



2018 ALZHEIMER'S DISEASE FACTS AND FIGURES
ALZHEIMER'S DISEASE IS THE
6TH leading cause of death in the United States
16.1 MILLION AMERICANS provide unpaid care for people with Alzheimer's or other dementias
These caregivers provided an estimated 18.4 BILLION HOURS of care valued at over $232 BILLION
Between 2000 and 2015 deaths from heart disease have decreased 11%
while deaths from Alzheimer's disease have increased 123%
1 IN 3 seniors dies with Alzheimer's or another dementia
It kills more than breast cancer and prostate cancer COMBINED
EARLY AND ACCURATE DIAGNOSIS COULD SAVE UP TO
$7.9 TRILLION in medical and care costs
IN 2018, Alzheimer's and other dementias will cost the nation $277 BILLION
BY 2050, these costs could rise as high as $1.1 TRILLION
5.7 MILLION Americans are living with Alzheimer's
BY 2050, this number is projected to rise to nearly 14 MILLION
EVERY 65 SECONDS someone in the United States develops the disease
alzheimer's association
THE BRAINS BEHIND SAVING YOURS.



Between 2015 and 2050, the numbers of people living with dementia in low income countries will increase by 264%, by 227% in upper middle income countries, by 223% in lower middle income countries and by 116% in high income countries.
Alzheimer's Research UK

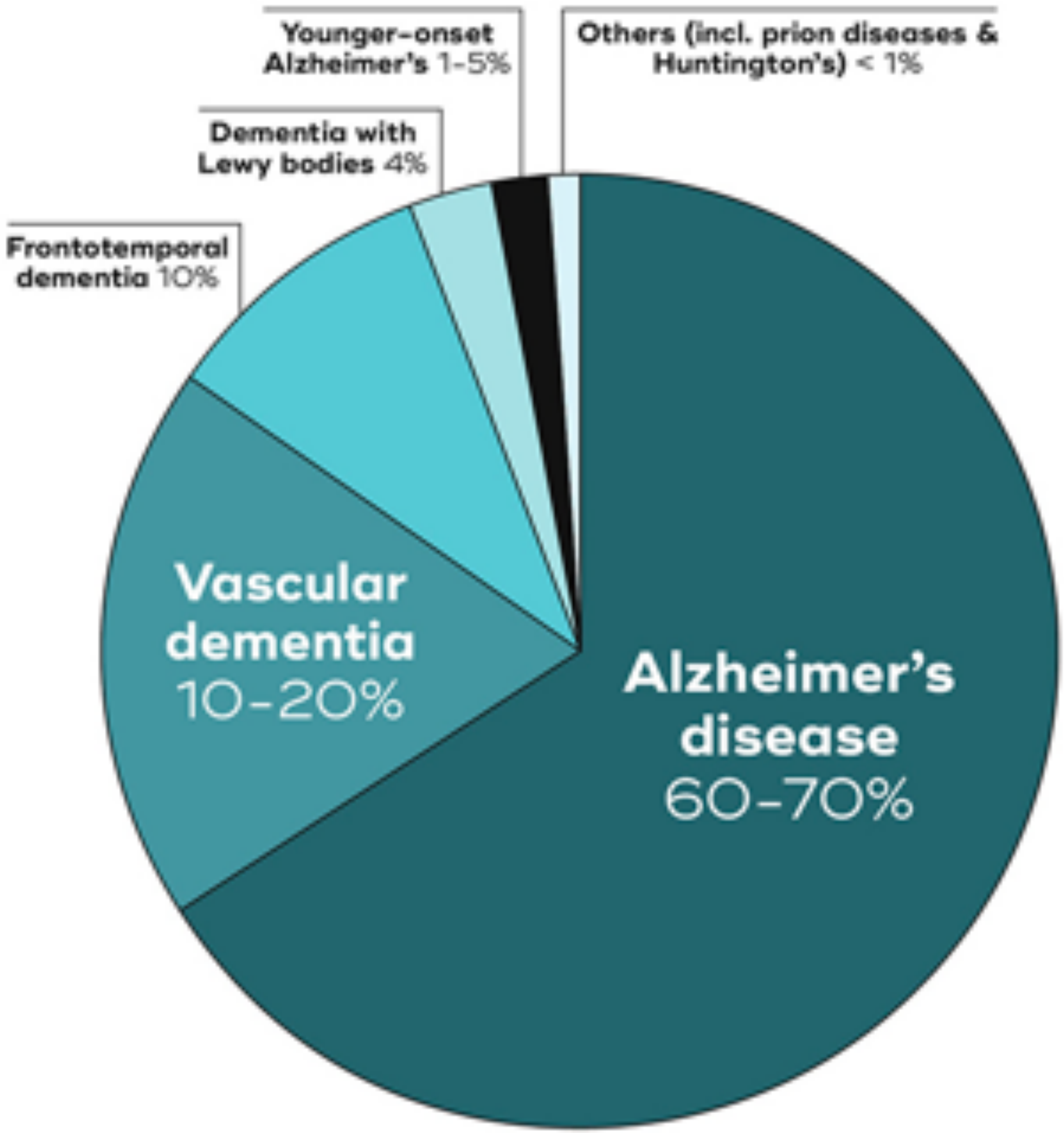
50m
people are living with dementia globally.
Alzheimer's Research UK

Globally, the numbers of people living with dementia will increase from 50m in 2018 to 152m in 2050, a 204% increase.
Alzheimer's Research UK

WHERE ALZHEIMER'S STARTS AND HOW IT SPREADS

- The damage initially appears to take place in the hippocampus, the part of the brain essential in forming memories
- As more neurons die, additional parts of the brain are affected, and they begin to shrink
- By the final stage of Alzheimer's, damage is widespread, and brain tissue has shrunk significantly

NIH National Institute on Aging


Younger-onset Alzheimer's 1-5%
Others (incl. prion diseases & Huntington's) < 1%
Dementia with Lewy bodies 4%
Frontotemporal dementia 10%
Vascular dementia 10-20%
Alzheimer's disease 60-70%

SIGNS AND SYMPTOMS

- Memory problems are typically one of the first signs of cognitive impairment related to Alzheimer's disease (mild cognitive impairment = MCI)
- Movement difficulties and problems with the sense of smell have also been linked to MCI
- The first symptoms of Alzheimer's vary from person to person and often relate to **decline in non-memory aspects of cognition**:
  - word-finding
  - vision/spatial issues
  - impaired reasoning or judgment

NIH National Institute on Aging

Dementia prevention, intervention, and care

Gill Livingston, Andrew Sommerlad, Vasiliki Orgeta, Sergi G Costafreda, Jonathan Huntley, David Ames, Clive Ballard, Sube Banerjee, Alistair Burns, Jiska Cohen-Mansfield, Claudia Cooper, Nick Fox, Laura N Gitlin, Robert Howard, Helen C Kales, Eric B Larson, Karen Ritchie, Kenneth Rockwood, Elizabeth L Sampson, Quincy Samus, Lon S Schneider, Geir Selbæk, Linda Teri, Naaheed Mukadam

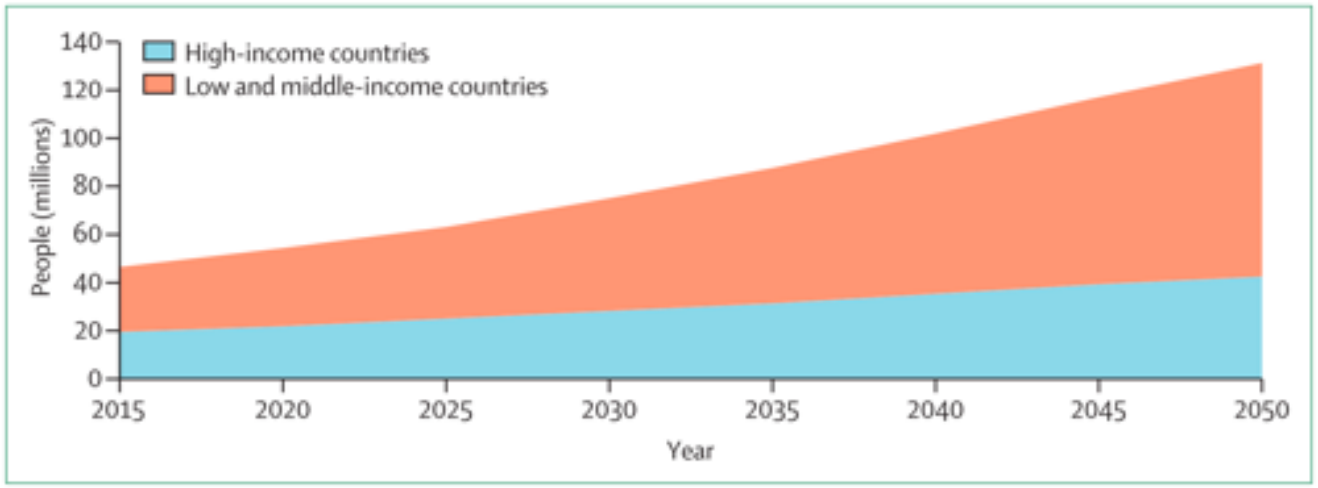


Figure 1: **Growth in numbers of people with dementia in high-income and low and middle-income countries**
Reproduced from Prince and colleagues,[2] by permission of Alzheimer's Disease International.

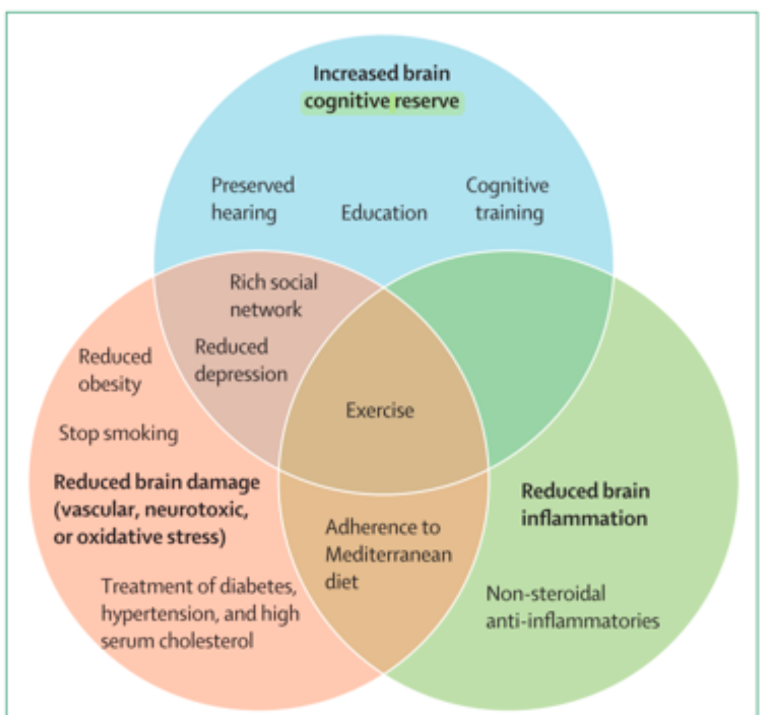


Figure 5: **Potential brain mechanisms for preventive strategies in dementia**

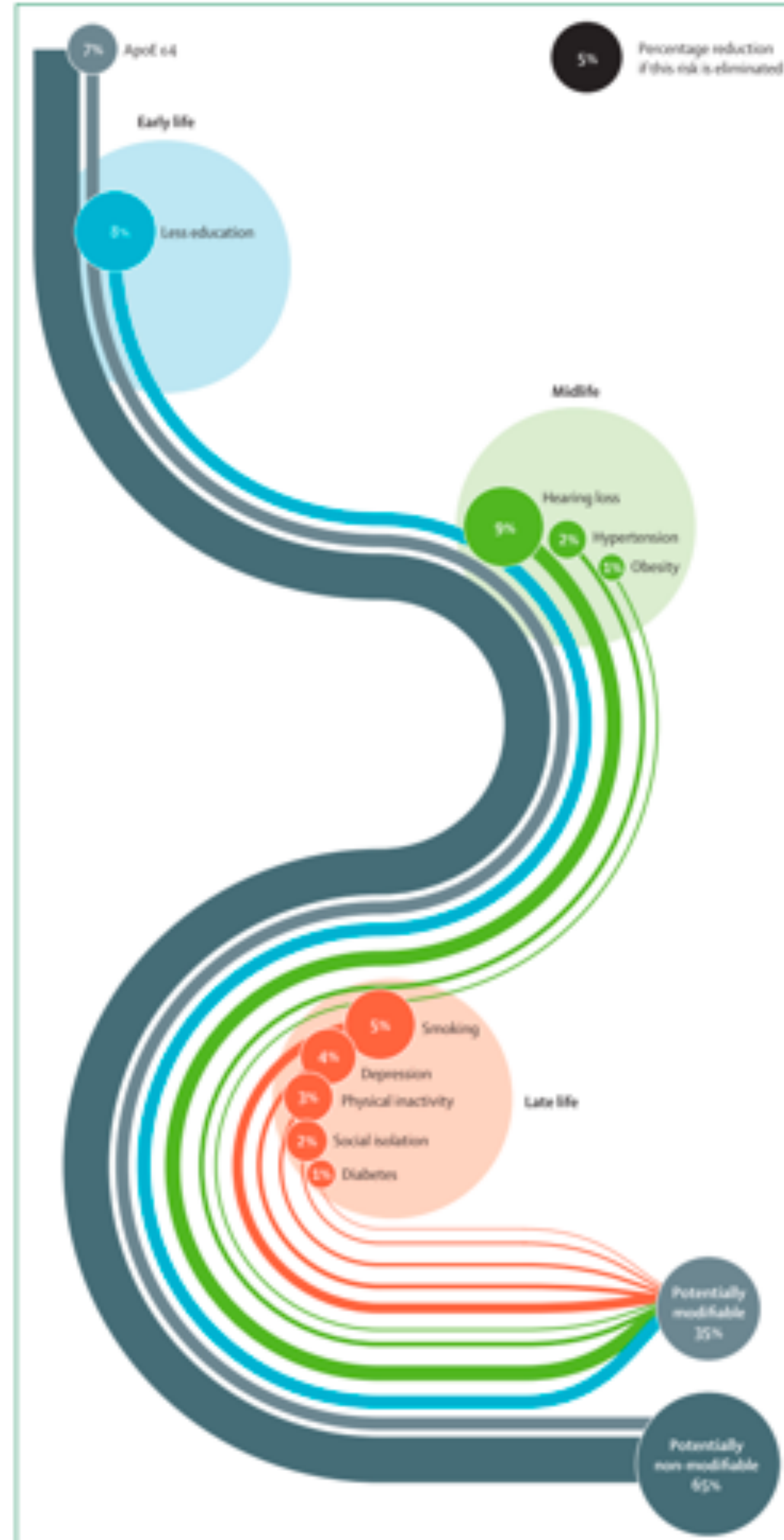
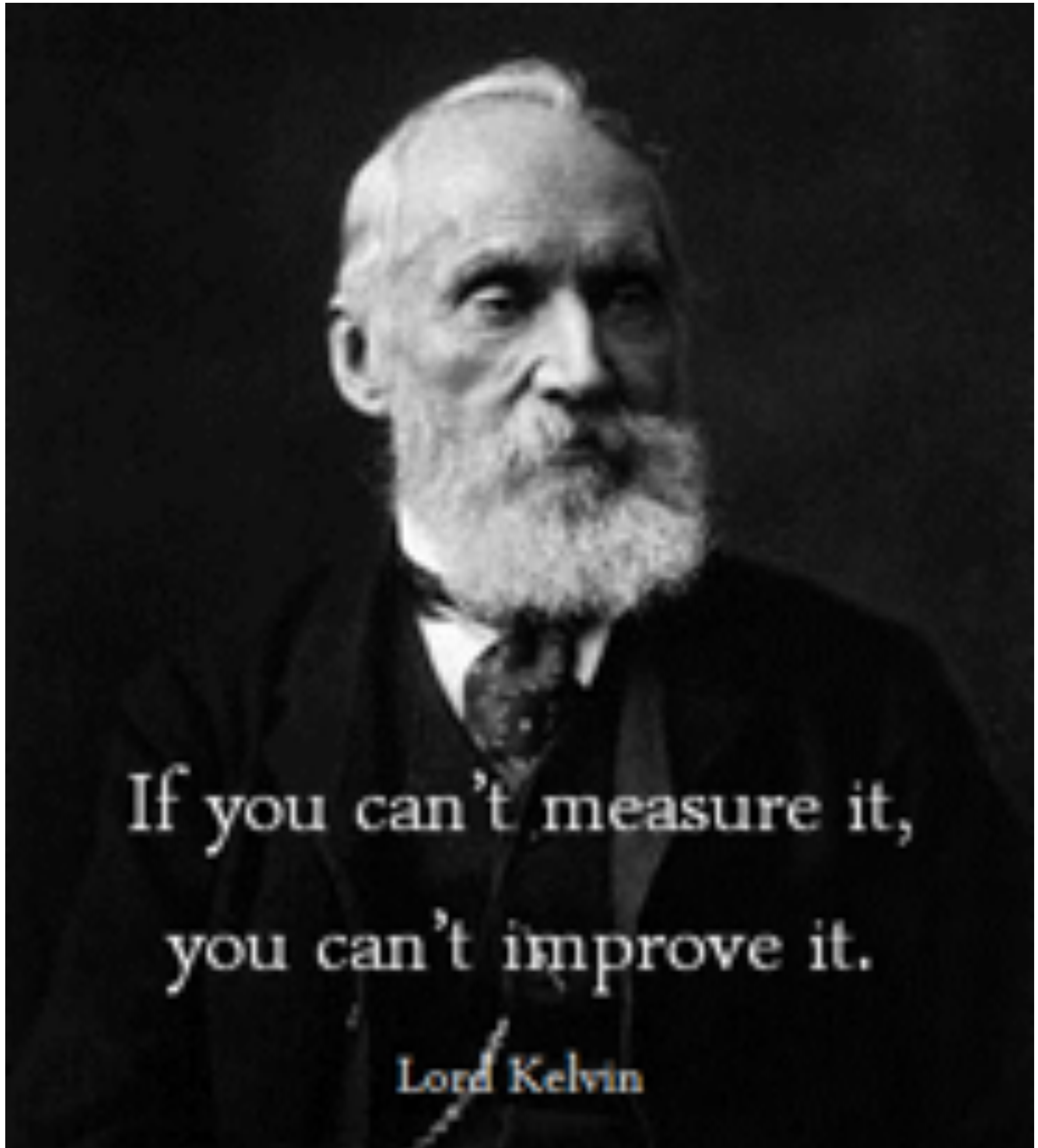


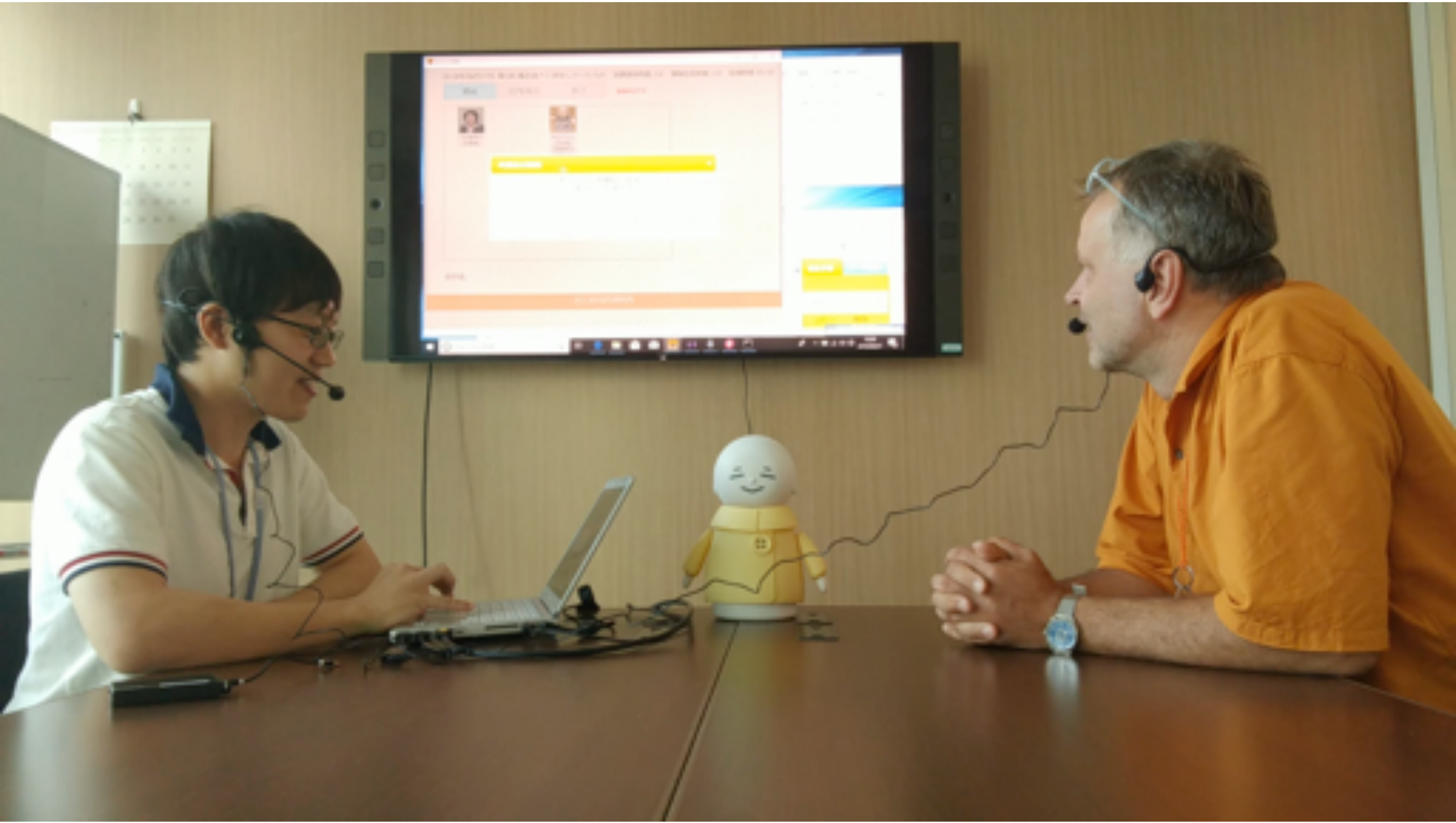
Figure 4: Life-course model of contribution of modifiable risk factors to dementia
Numbers are rounded to nearest integer. Figure shows potentially modifiable or non-modifiable risk factors.

DEMENTIA PREVENTION

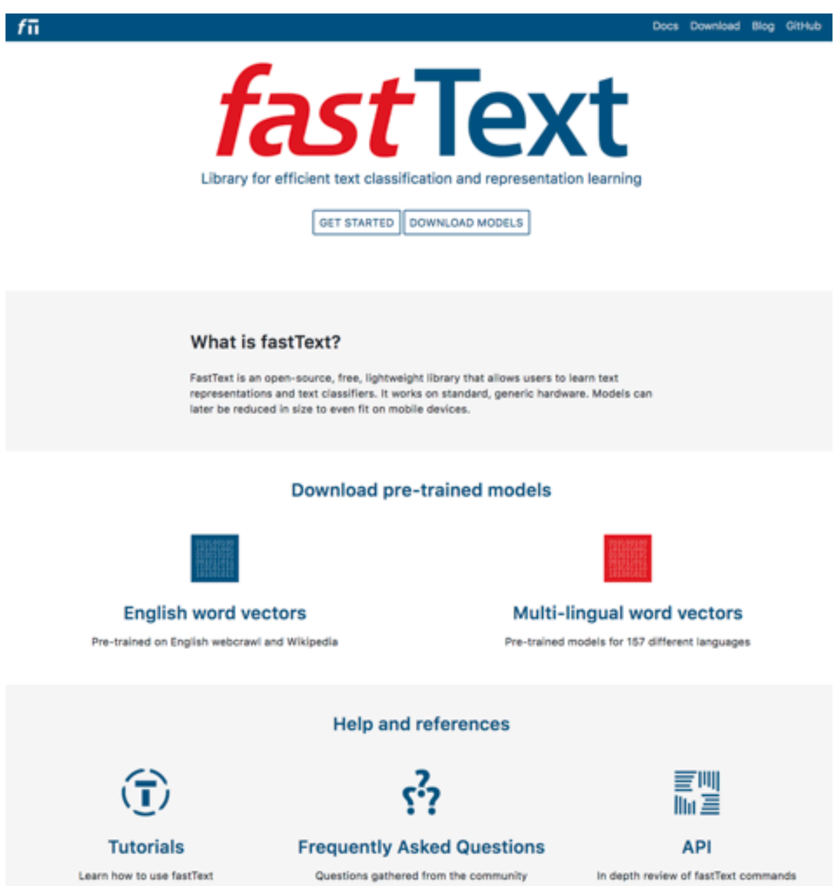
- Pharma companies such as TAKEDA Pharma, Novartis, etc., have “digital pharma” departments for “beyond-a-pill” healthcare
- FONOBONO NPO founded by Dr. Otake–Matsuura
- Pharma spinoff company AIKOMI develops already late dementia cognitive-therapy support
- Improved biomarkers comparing to our previous AD EEG research

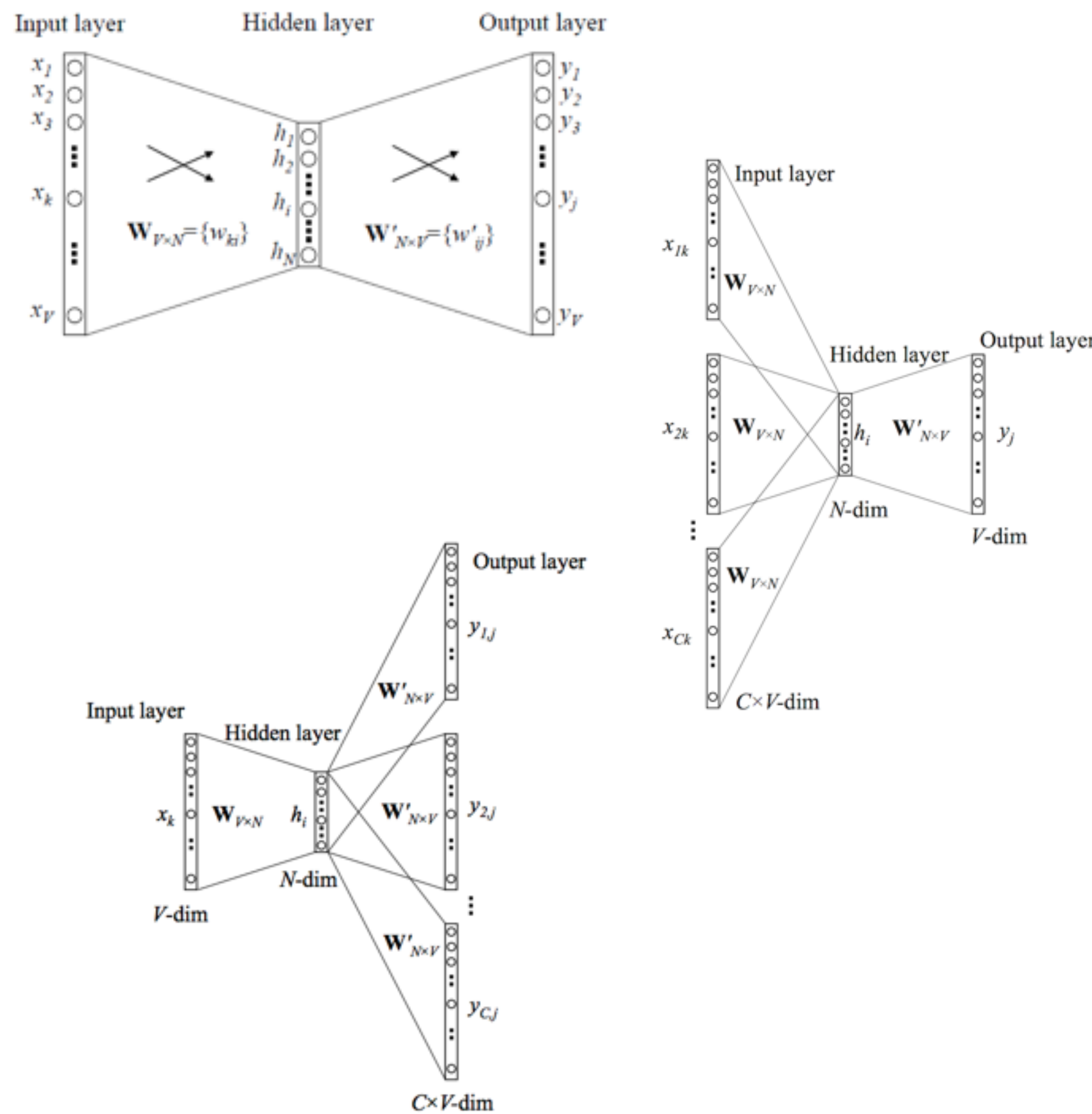

If you can't measure it,
you can't improve it.
Lord Kelvin

CONVERSATIONAL ASSISTIVE AI BASED ON CO-IMAGINATION METHOD



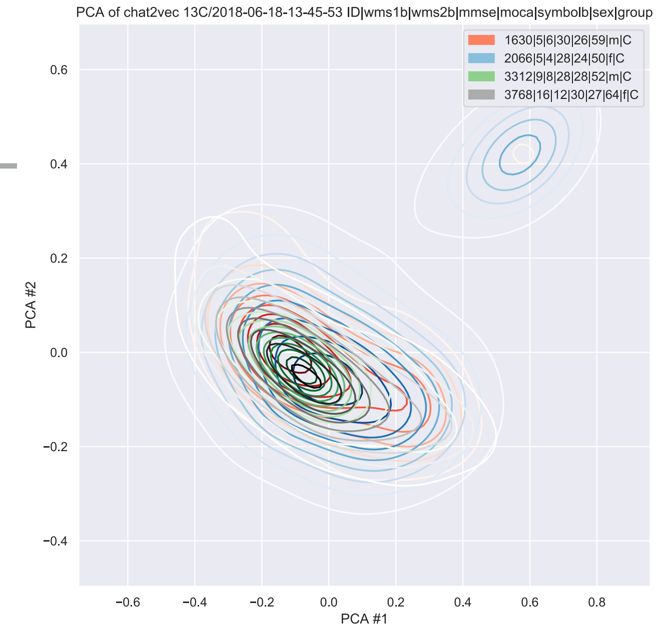
FASTTEXT MODEL USING WORD2VEC (SKIP-GRAM MODEL USING SUB-WORD INFO) APPROACH

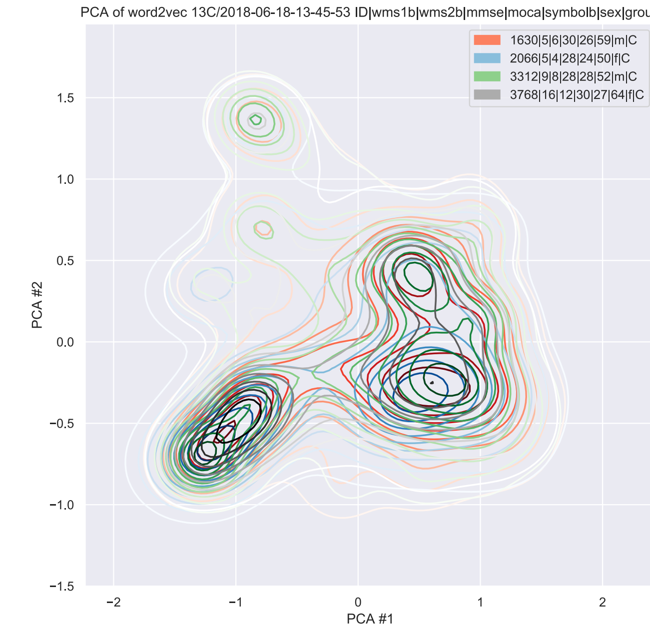


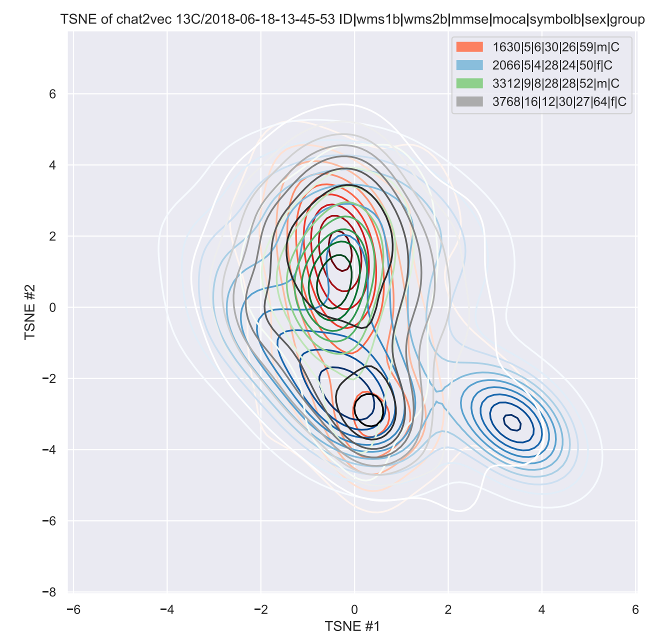


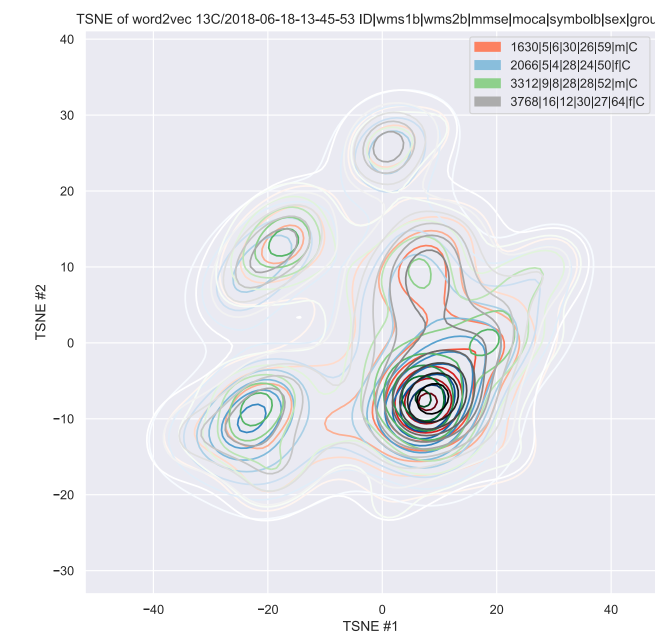
PRELIMINARY RESULTS IN “BAG-OF-WORDS/CHATS” ANALYSIS APPROACH:

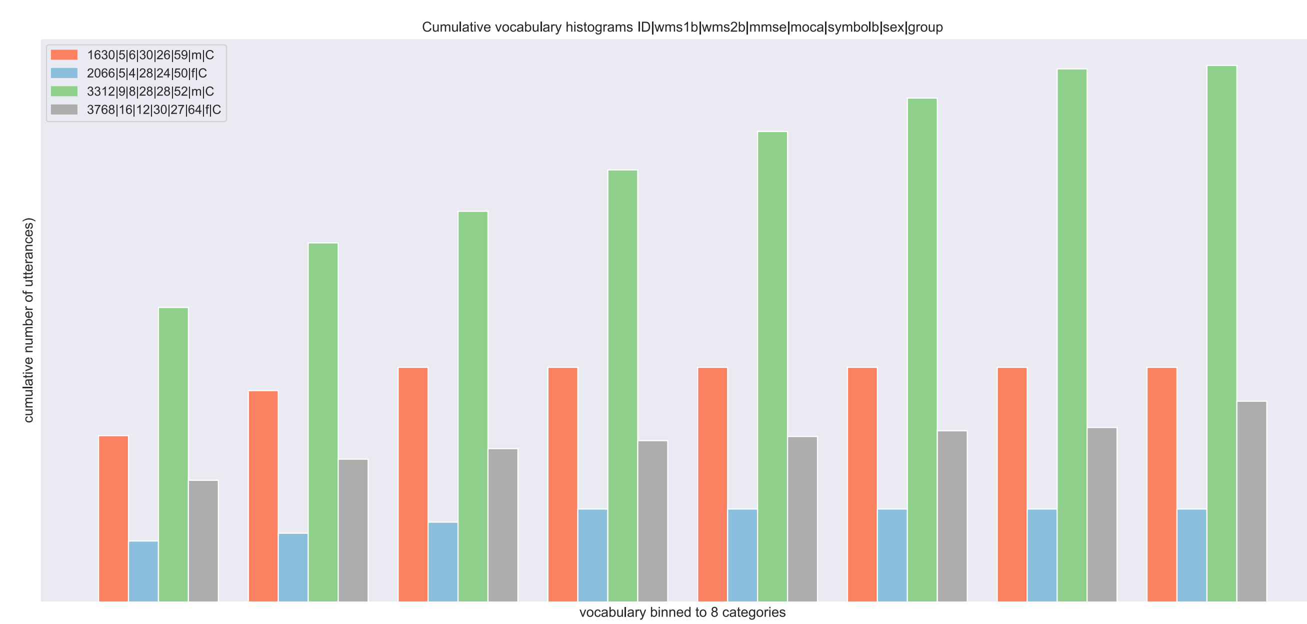
13C



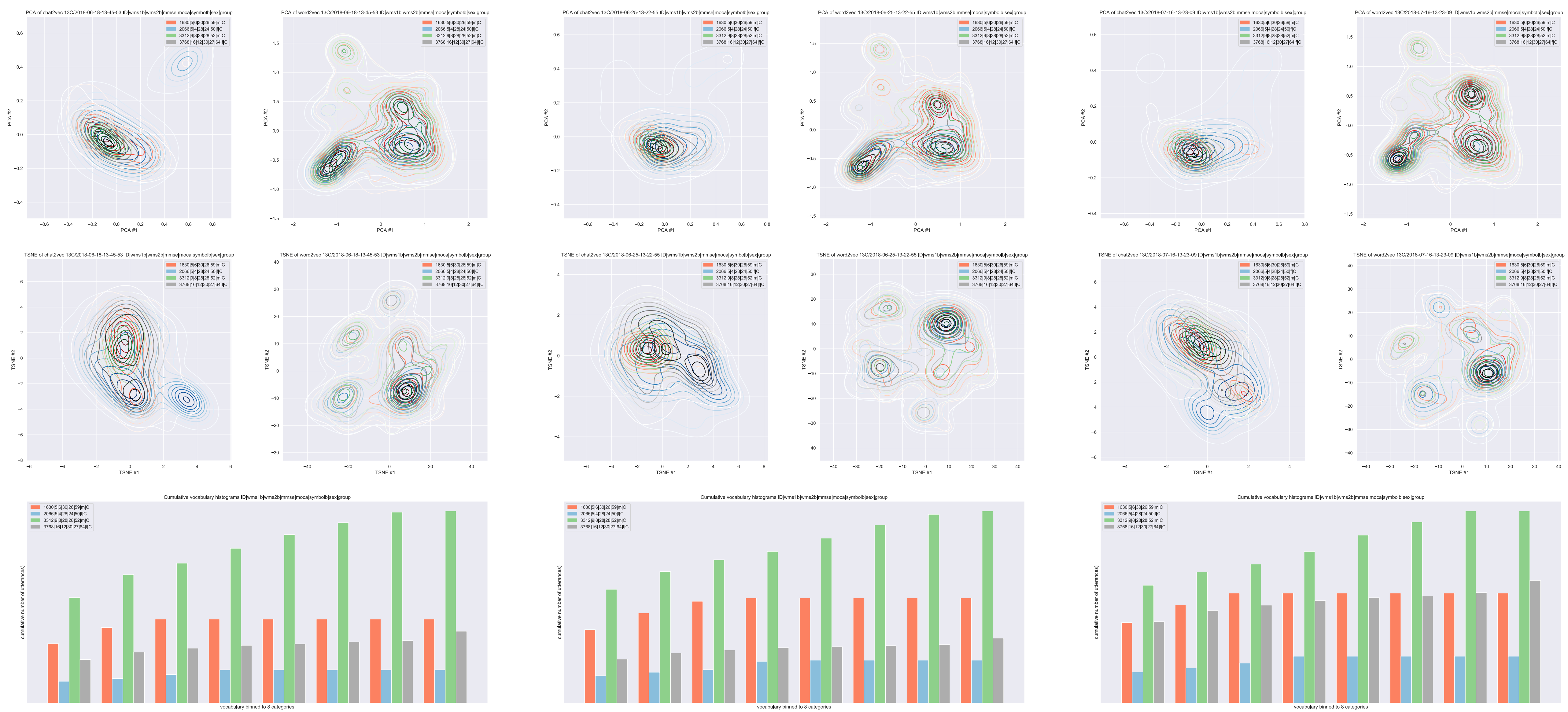




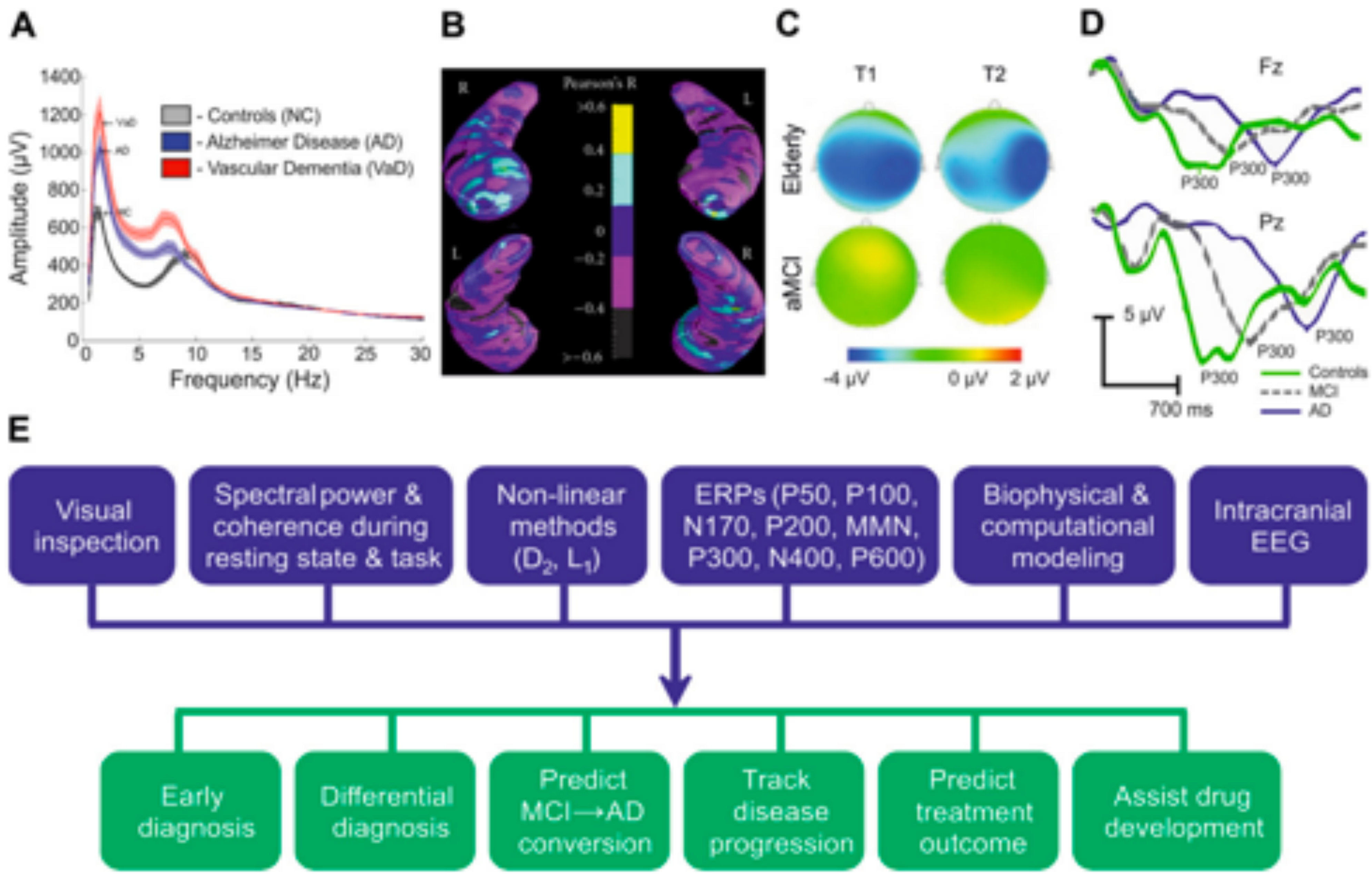




PRELIMINARY RESULTS IN "BAG-OF-WORDS/CHATS" ANALYSIS APPROACH: 13C



BRAIN FACTS



Horvath, A., Szucs, A., Csukly, G., Sakovics, A., Stefanics, G. and Kamondi, A., 2018. EEG and ERP biomarkers of Alzheimer’s disease: a critical review. *Front Biosci (Landmark Ed)*, *23*, pp.183-220.

BRAIN FACTS

ERP component	Amplitude in AD	Latency in AD	Amplitude in MCI	Amplitude in preclinical AD	Effect of cholinergic treatment
Early ERPs	IC	IC	DNA	DNA	DNA
N170	↓ / IC	↑	DNA	DNA	DNA
N200	↓	↑↑	↑	↑	Less influenced by cholinergic treatment
P300	↓↓↓	↑↑↑	↑↑↑	↑↑	Response to cholinergic treatment
N400	↓↓	↑↑	↑↑	↑	DNA
P600	↓↓	↑↑	↑↑	↑	DNA

The number of arrows indicates the number of studies reporting concordant results: ↑ = 1–5 studies ↑↑ = 5–10 studies ↑↑↑ >10 studies. DNA = Data not available; IC = inconclusive results.

EEG band	Spectral power and phase coherence in AD	Spectral power and phase coherence in MCI	Spectral power and phase coherence in preclinical AD
Gamma	IC	IC	DNA
Beta	↓↓	↓↓↓	↓↓
Alpha	↓↓↓	↓↓↓	↓↓↓
Theta	↑↑	↑↑	↑
Delta	↑↑↑	↑↑	↑↑↑

The number of arrows indicates the number of studies reporting concordant results: ↑ = 1–5 studies ↑↑ = 5–10 studies ↑↑↑ >10 studies. DNA = data not available; IC = inconclusive results.

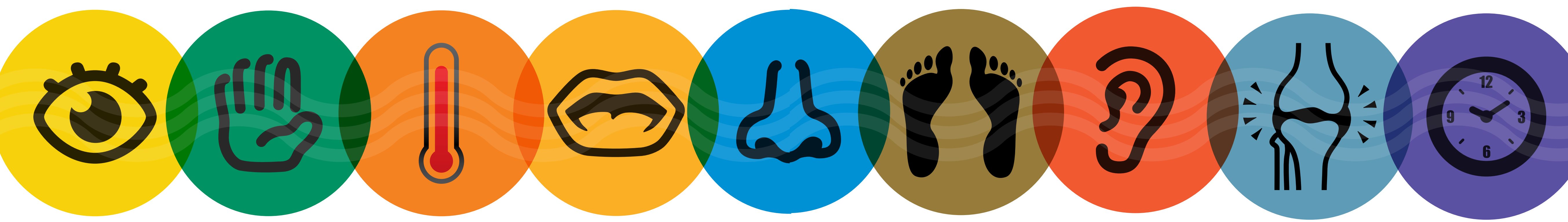
EEG feature	Alzheimer's disease	Frontotemporal dementia	Diffuse Lewy-body dementia	Depression-related cognitive decline
Delta-power	↑	x	↑	x
Theta-power	↑	x	↑	x
Alpha-power	↓	↓	↓	↑
Beta-power	↓	↓	↓	↑
Complexity	↓	↓	↓	↓
Connectivity	↓	↓	↓	↑
Epileptic activity	↑	↑	DNA	DNA

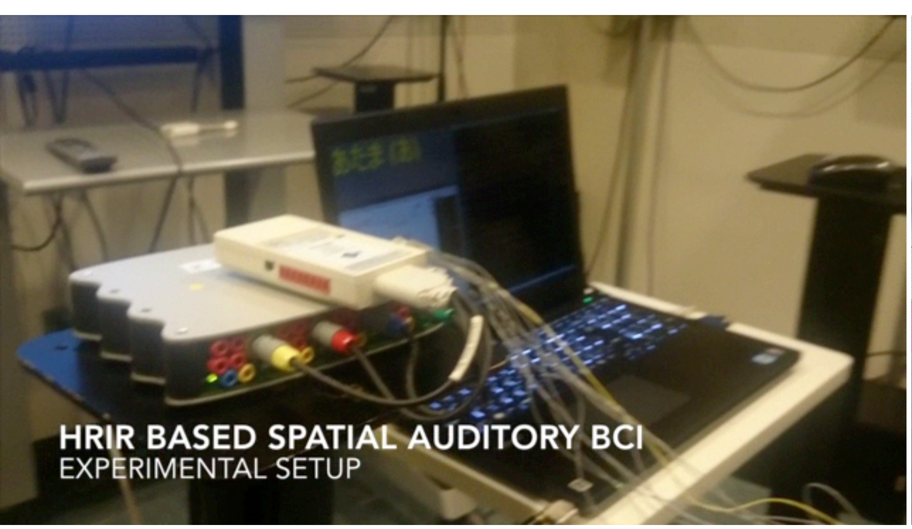
↑: increase, ↓: decrease, DNA: data not available. Across studies there are clear findings suggesting that most common forms of neurocognitive disorders show different EEG characteristics.

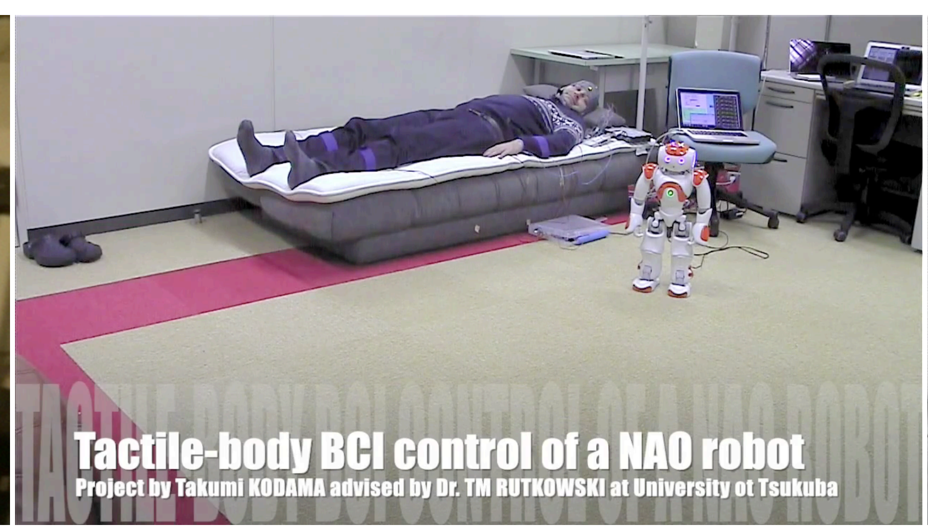
Horvath, A., Szucs, A., Csukly, G., Sakovics, A., Stefanics, G. and Kamondi, A., 2018. EEG and ERP biomarkers of Alzheimer’s disease: a critical review. Front Biosci (Landmark Ed), 23, pp.183-220.

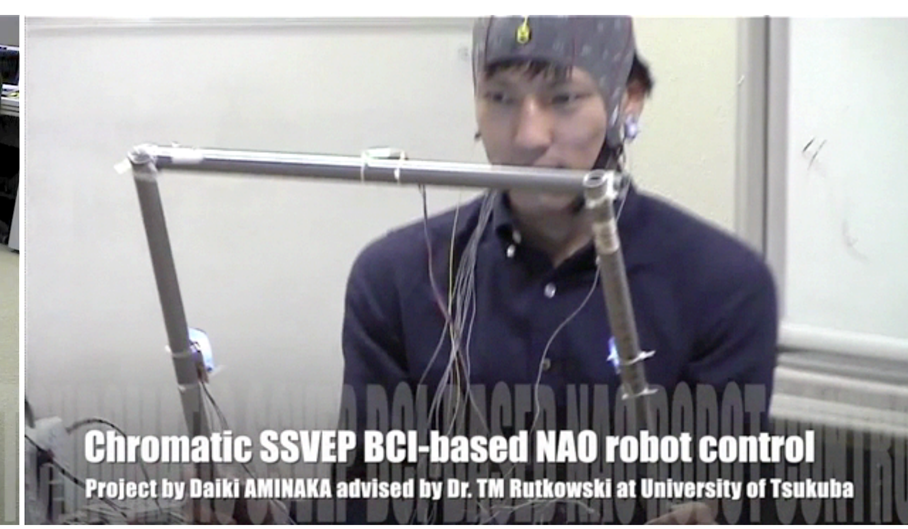
EEG + FNIRS BY G.TEC

MULTI–SENSORY STIMULATION ENVIRONMENTS FOR BCI AND NEUROTECHNOLOGY APPLICATIONS



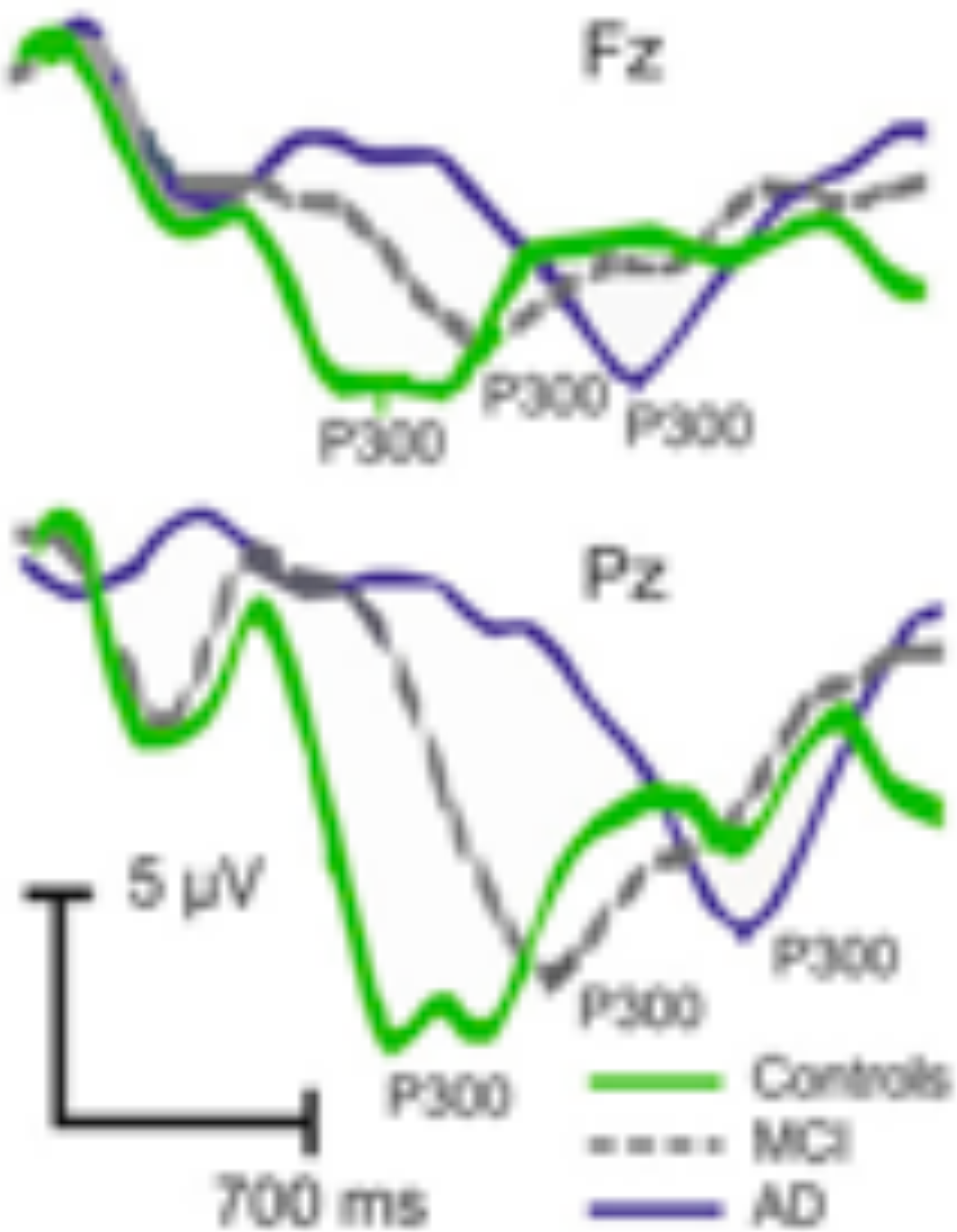


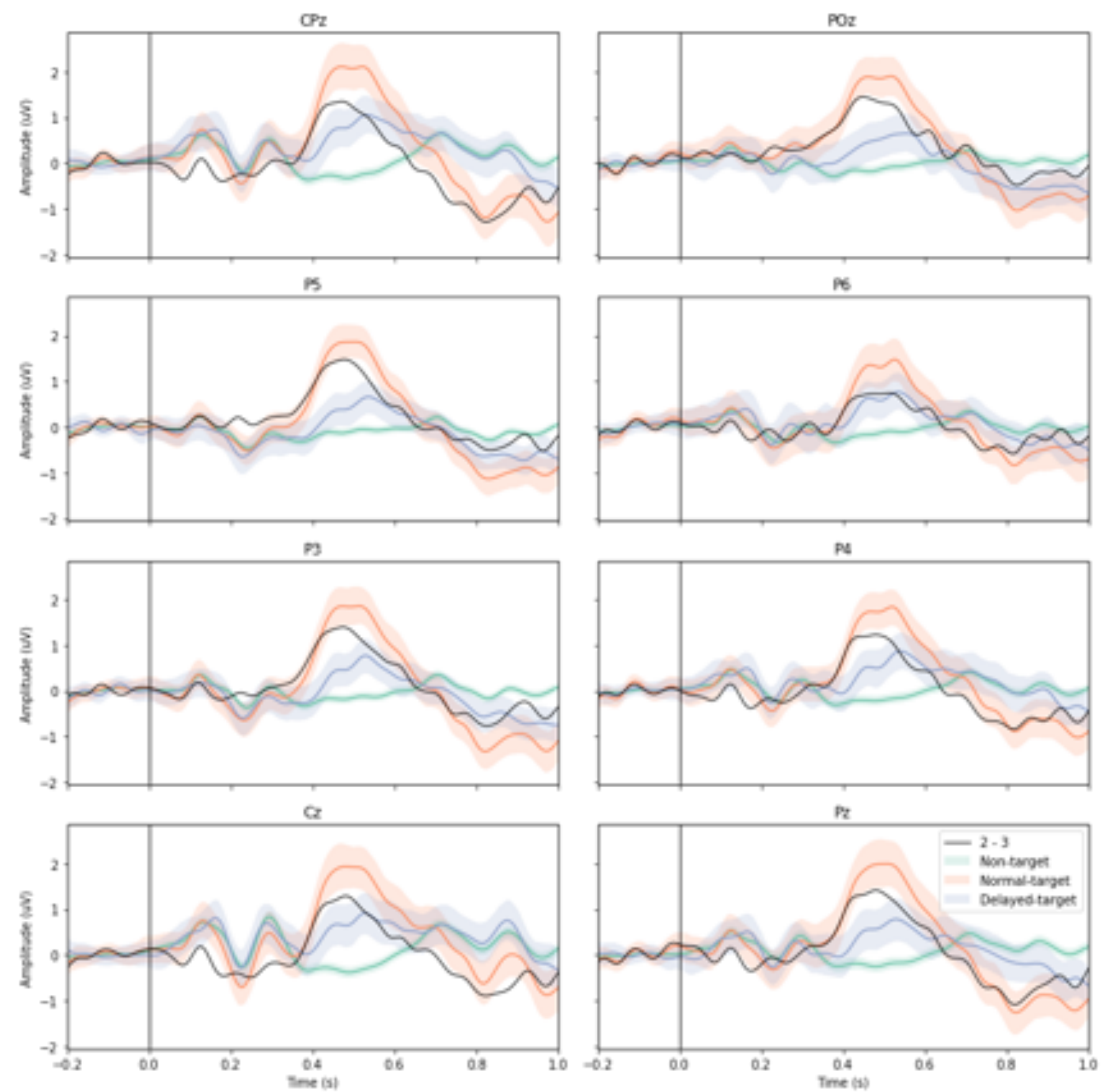




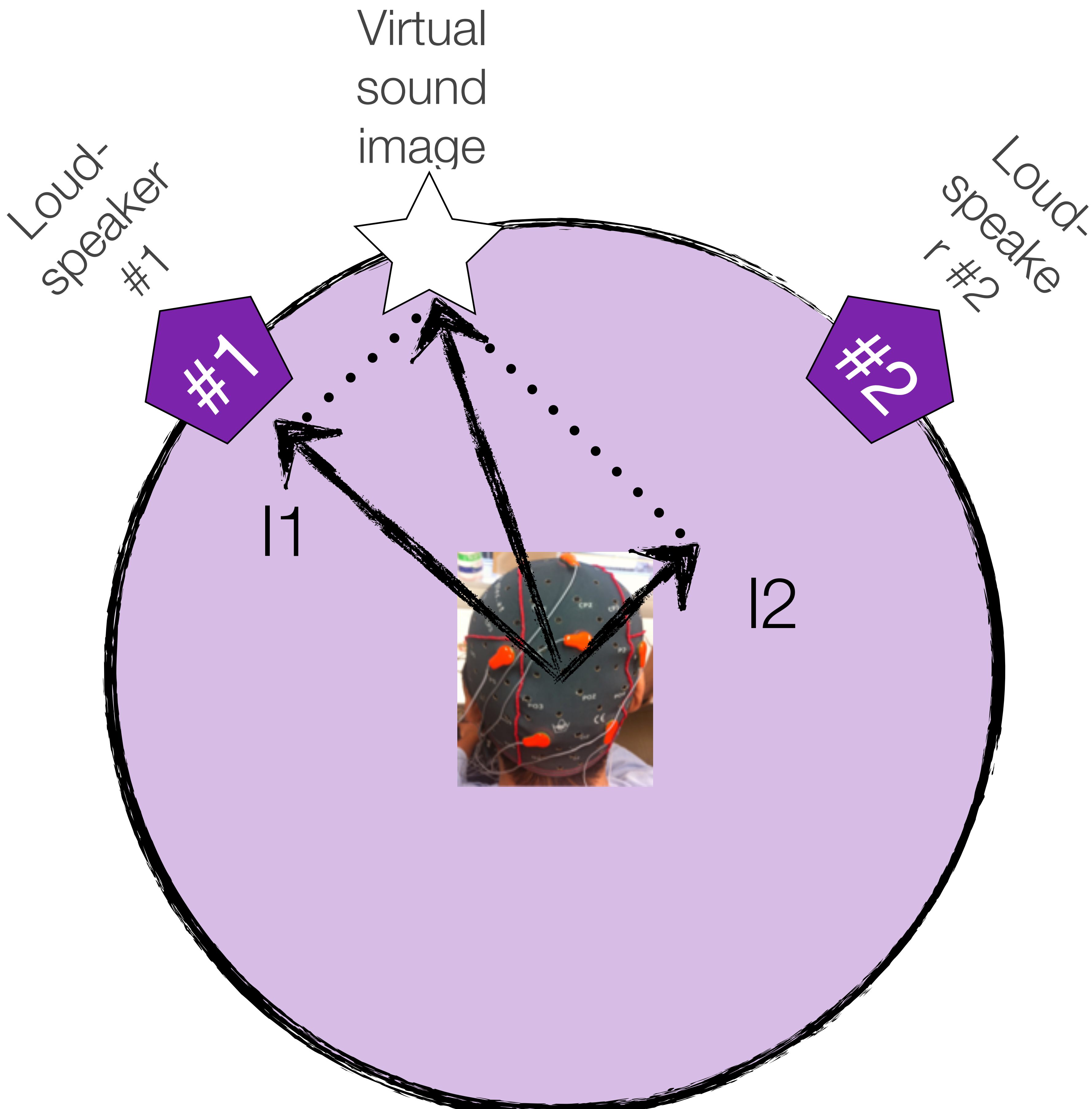
DEMENTIA & ERP

HORVATH, A., SZUCS, A., CSUKLY, G., SAKOVICS, A., STEFANICS, G. AND KAMONDI, A., 2018. EEG AND ERP BIOMARKERS OF ALZHEIMER'S DISEASE: A CRITICAL REVIEW. FRONT BIOSCI (LANDMARK ED), 23, PP.183–220.



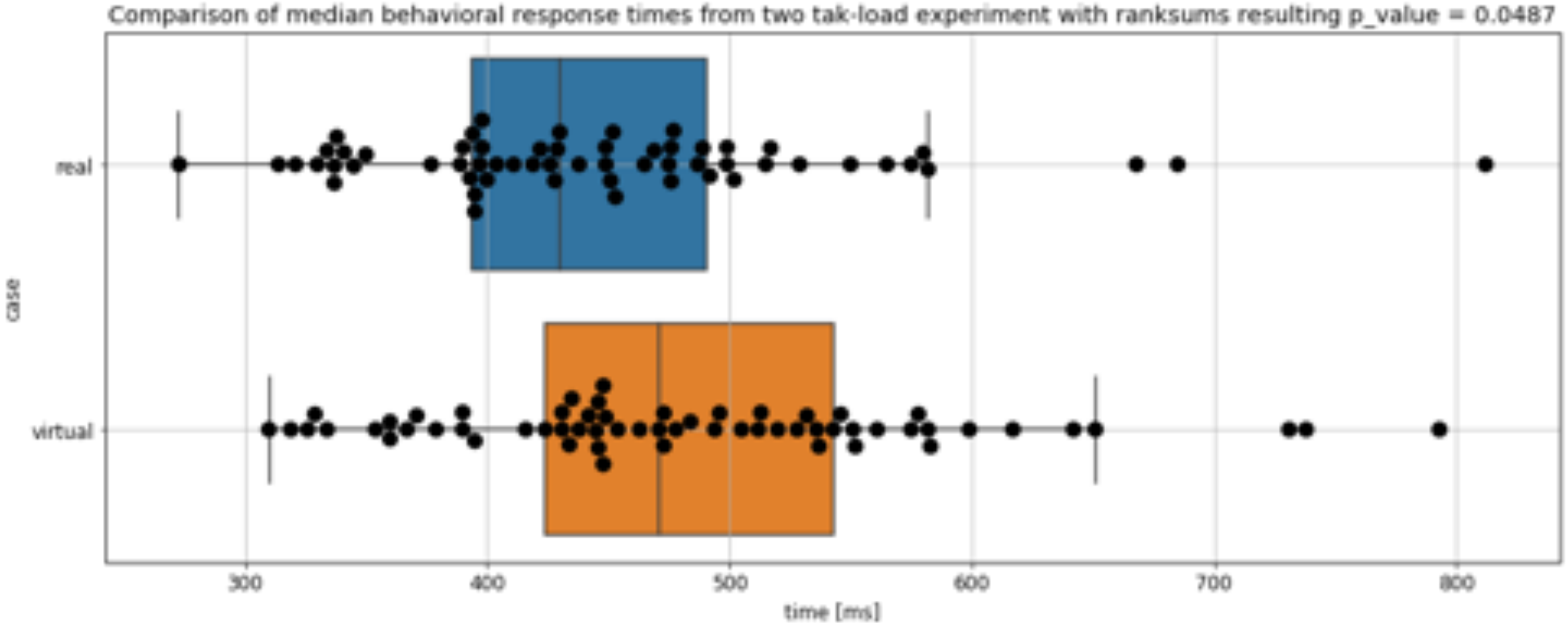


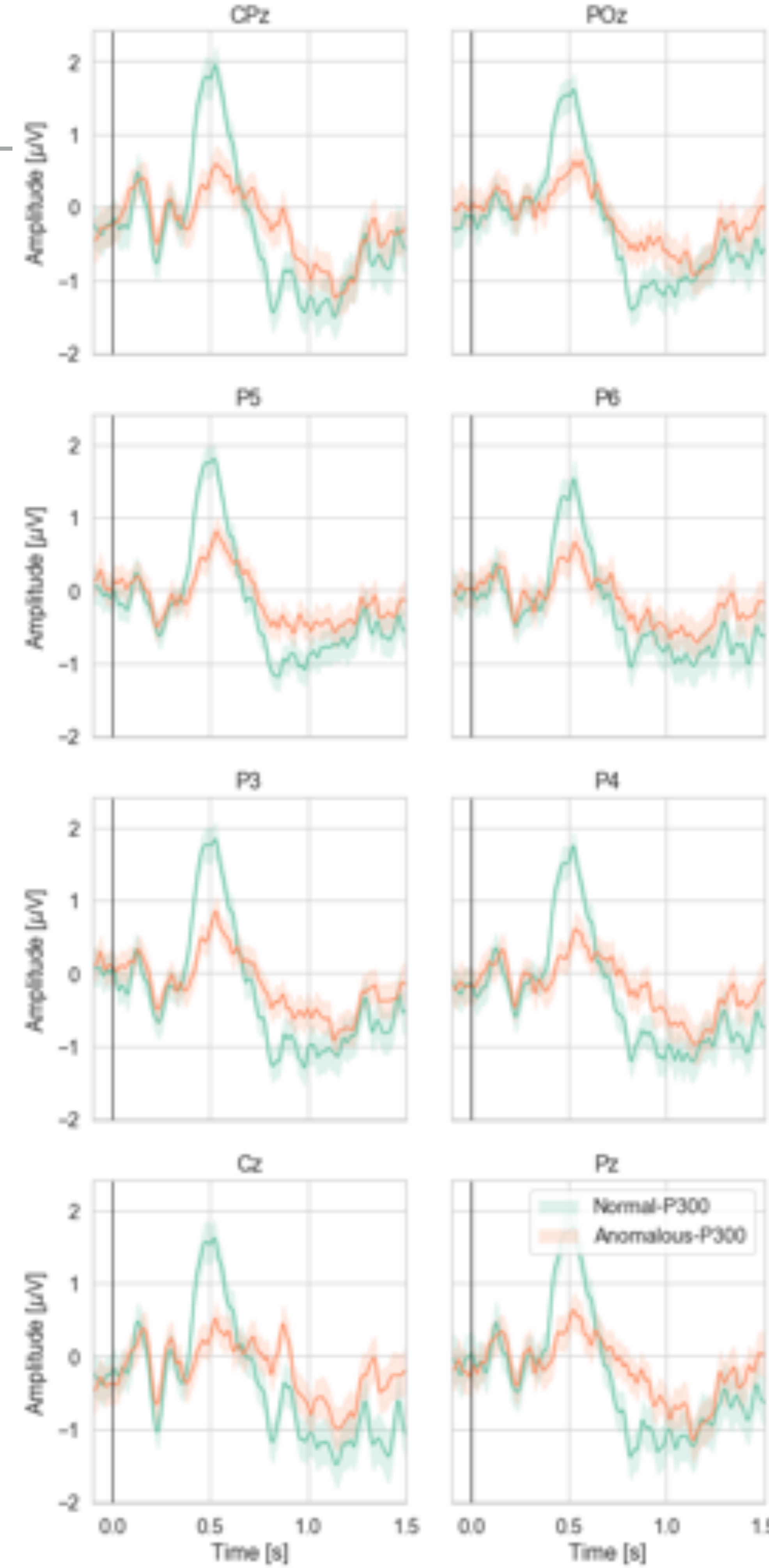
SPATIAL AUDITORY BCI WITH REAL & VIRTUAL SOUND SOURCES



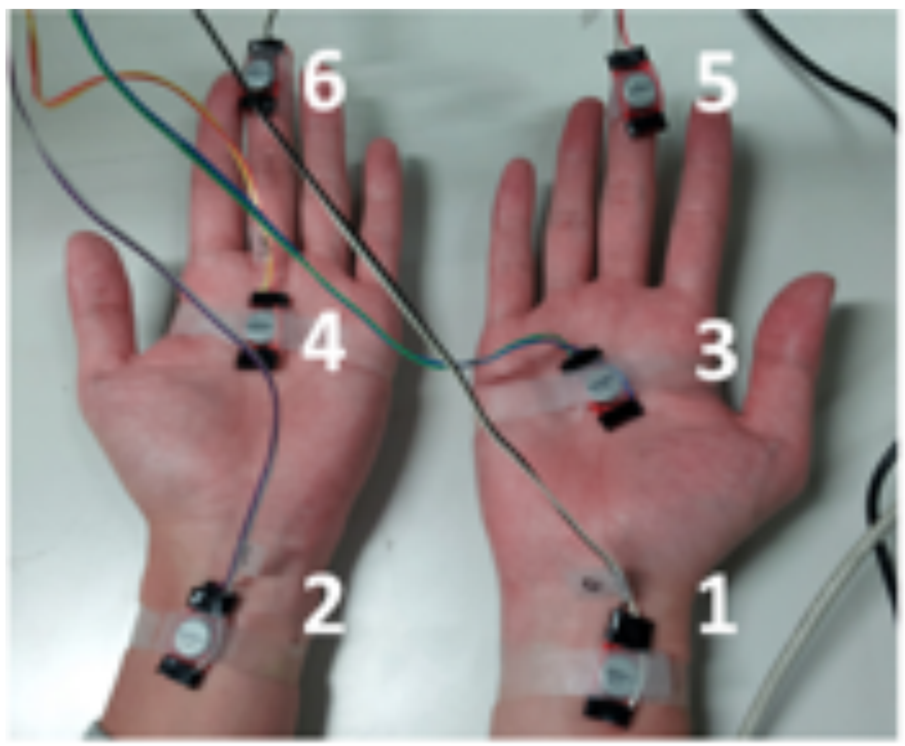
SPATIAL REAL VS. VIRTUAL SOUND BEHAVIORAL & ERP RESPONSES

Comparison of median behavioral response times from two tak-load experiment with ranksums resulting p_value = 0.0487

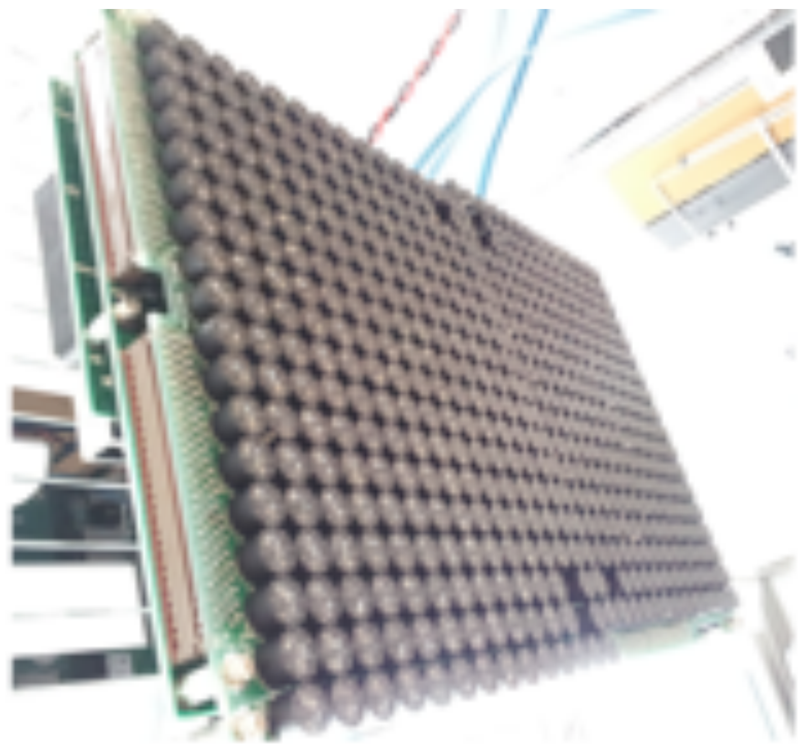




EXPERIMENTAL SETUP



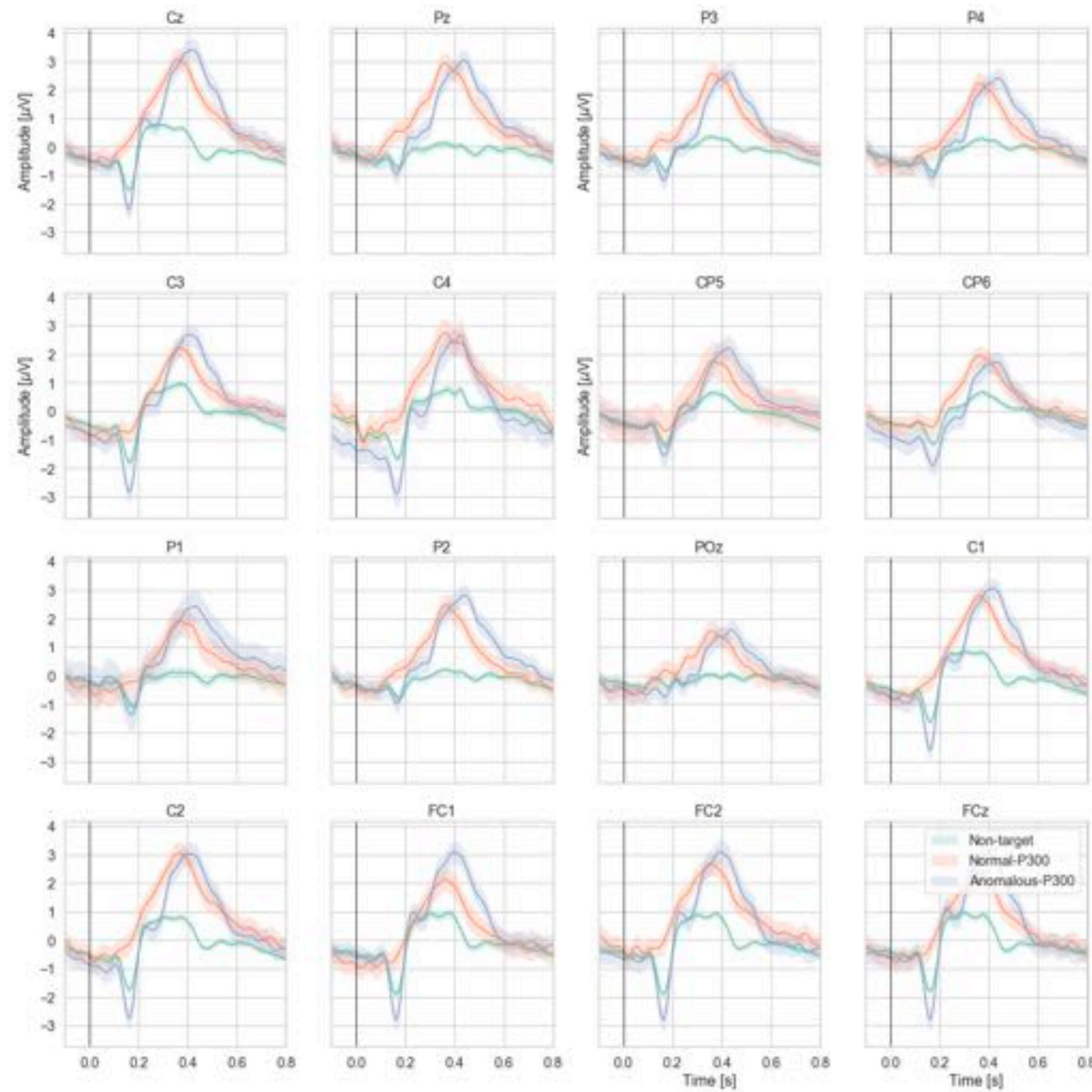
User palms with attached vibrotactile transducers used in vtBCI experiments. The stimulus locations where represented by digits to be spelled.

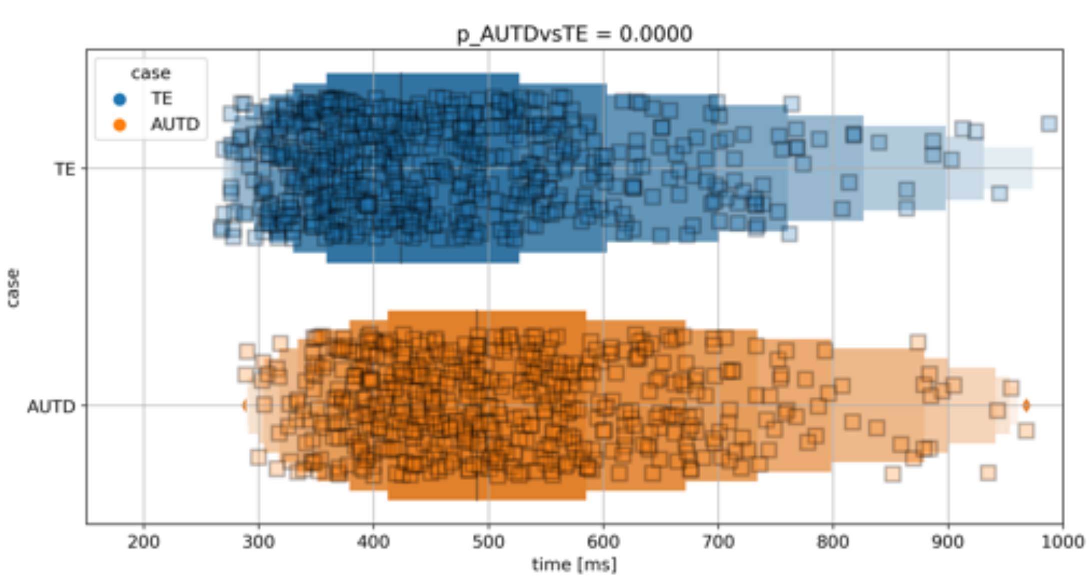


The AUTD array with ultrasonic transducers used to create the contact-less tactile pressure sensation.

The user during autdBCI experiment with the both palms placed under the AUTD array with ultrasonic transducers.

SPATIAL SKIN-CONTACT VS. AIRBORNE TACTILE STIMULUS BEHAVIORAL & ERP RESPONSES





RIEMANNIAN GEOMETRY CLASSIFIER

In a classifier training phase, a geometric mean covariance matrix $\mathbf{C}_c$, characterizing each ERP class c, is computed, using all EEG channels as inputs. In order to measure a distance of a newly captured ERP to the class–characterizing mean matrix, the RG techniques are used. A geodesic traversing two points $\mathbf{C}_i$ and $\mathbf{C}_j$ is the shortest path curve that connects them. It has a minimum length.
A Riemannian distance between covariance matrices is obtained as follows,

$$\delta_R = \left|\left|ln(\mathbf{C}_i^{-1}\mathbf{C}_j)\right|\right|_F = \sqrt{\sum_n [ln(w_n)]^2},$$

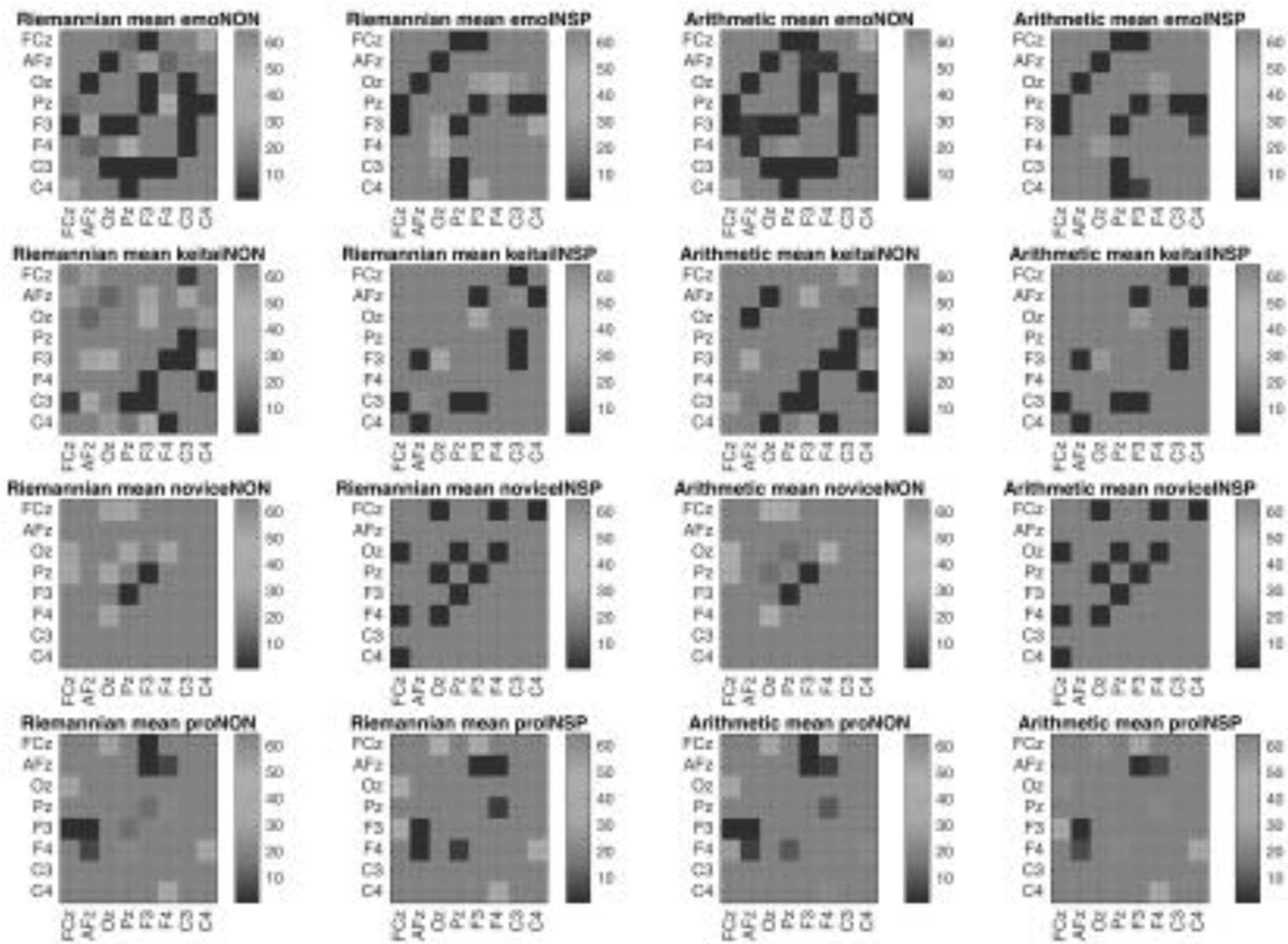
where the symbol $||\cdot||$ denotes a Frobenius norm and $w_1,\ldots,w_n$ the eigenvalues of $\mathbf{C}_i^{-1}\mathbf{C}_j$, respectively. The geometric mean of L covariance matrices is computed as,

$$D(\mathbf{C}_1, \cdots, \mathbf{C}_l) = \arg\min_{\mathbf{C}} \sum_{l=1}^{L} \delta_R^2(\mathbf{C}, \mathbf{C}_l)$$

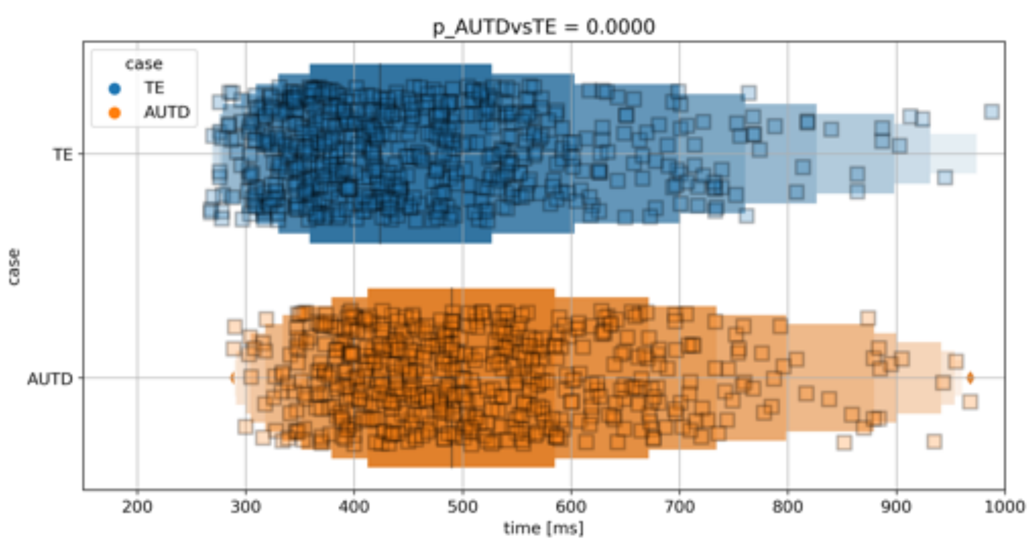
The shortest curve between two points (covariance matrices) on the manifold, which is a geodesic, is defined using the Riemannian metric as,

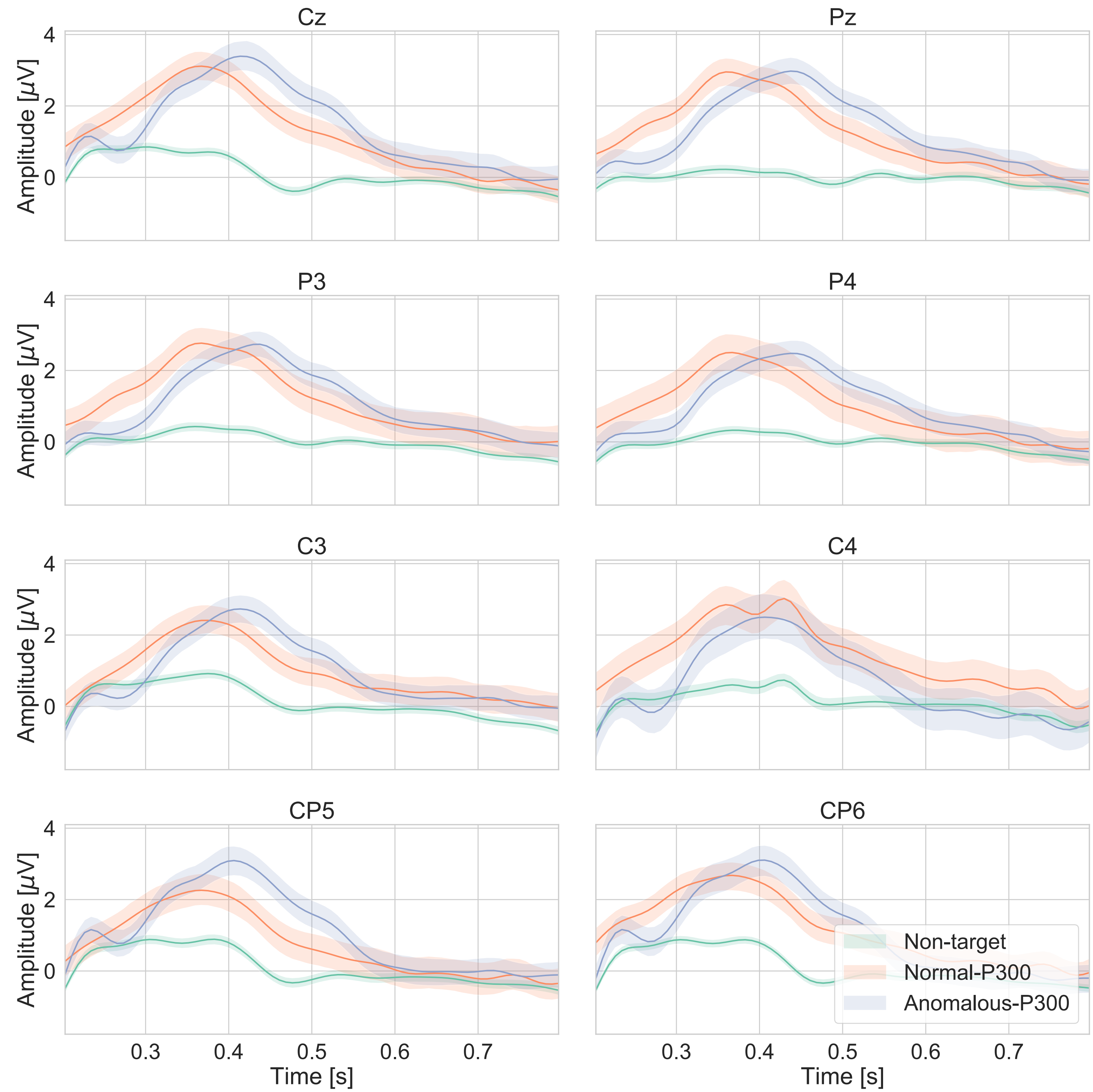
$$\Gamma(\mathbf{C}_i, \mathbf{C}_j, \tau) = \mathbf{C}_i^{\frac{1}{2}} \left(\mathbf{C}_i^{-\frac{1}{2}}\mathbf{C}_j\mathbf{C}_i^{-\frac{1}{2}}\right)^{\tau} \mathbf{C}_i^{\frac{1}{2}}, \text{ where } \tau \in \{0, 1\} \text{ is a scalar.}$$

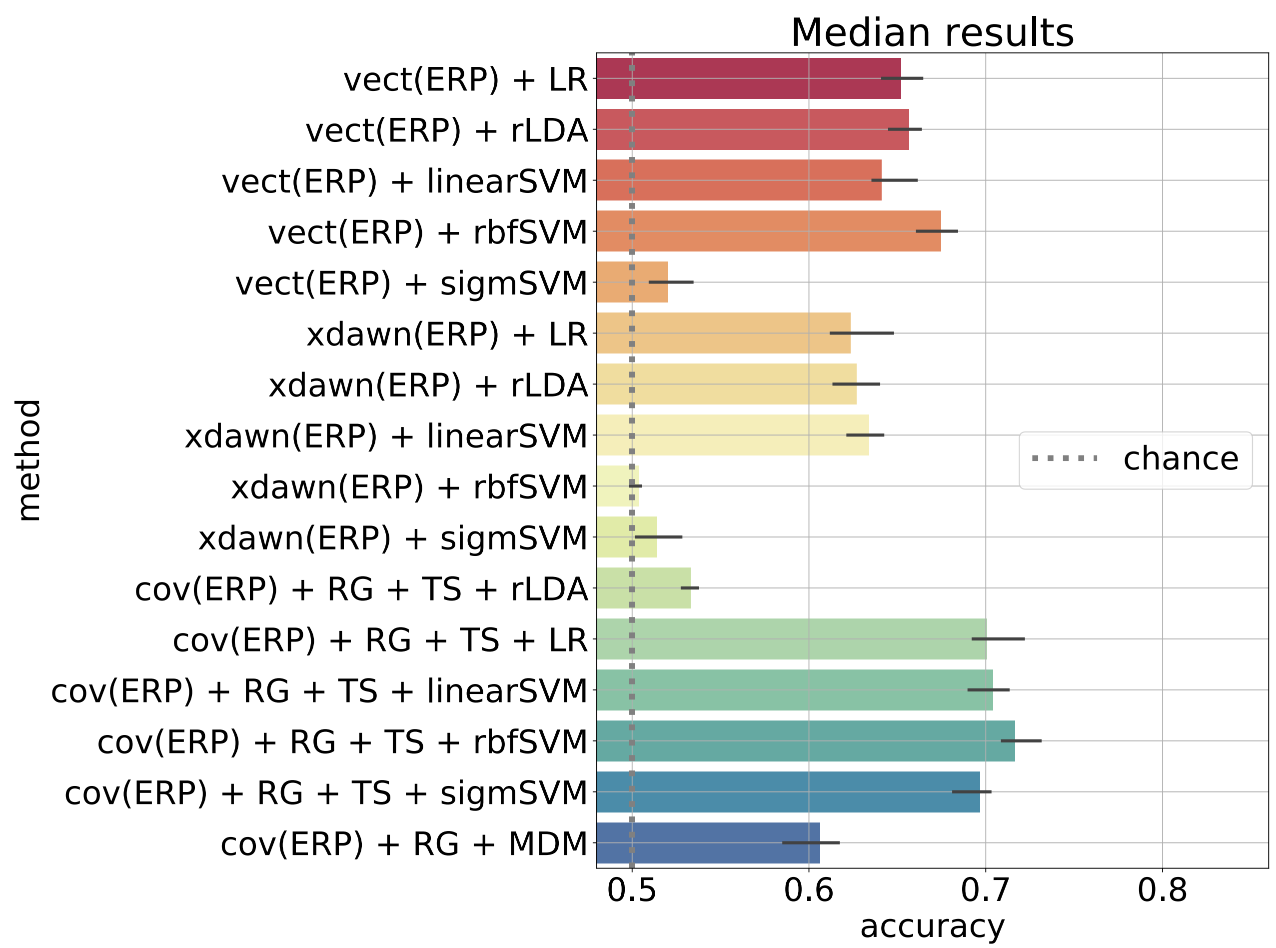
EXAMPLE COVARIANCE MATRICES OF VARIOUS EEG PATTERNS



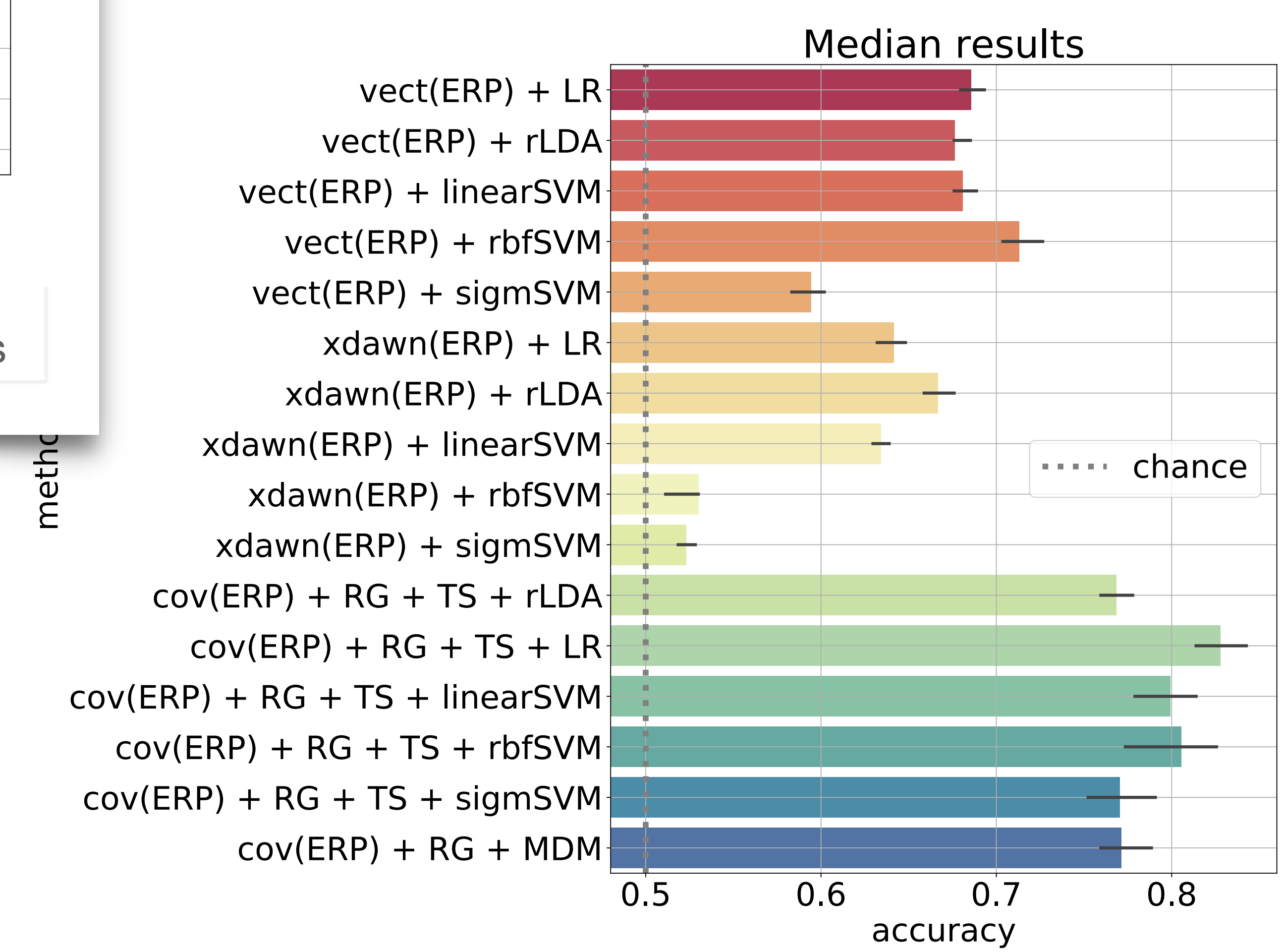
TACTILE (TIME-SHIFTED-P300) ERP RESPONSES







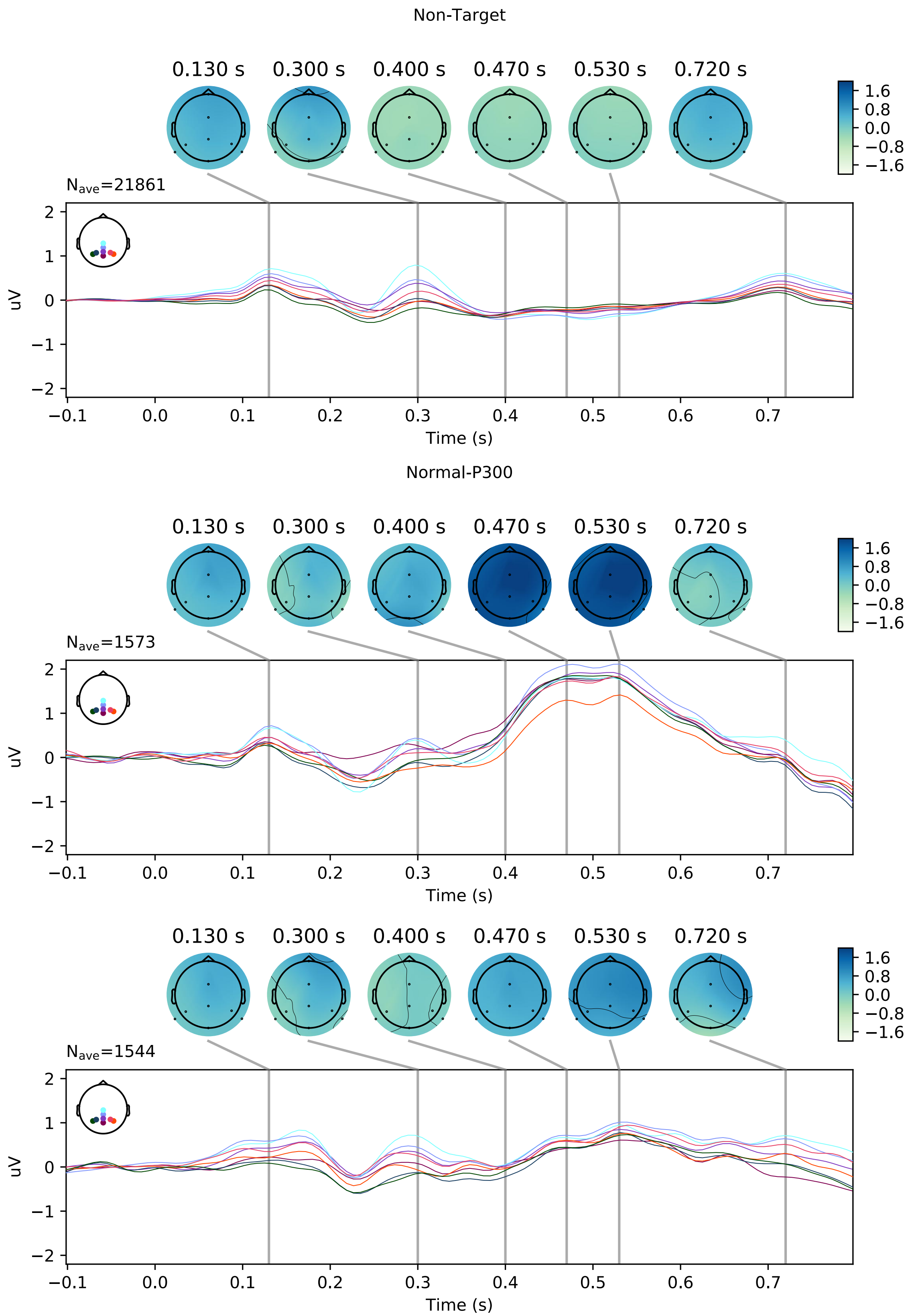
Results: Median accuracies of all subject dataset for training and testing using tactile and AUTD (time shifted) P300 responses

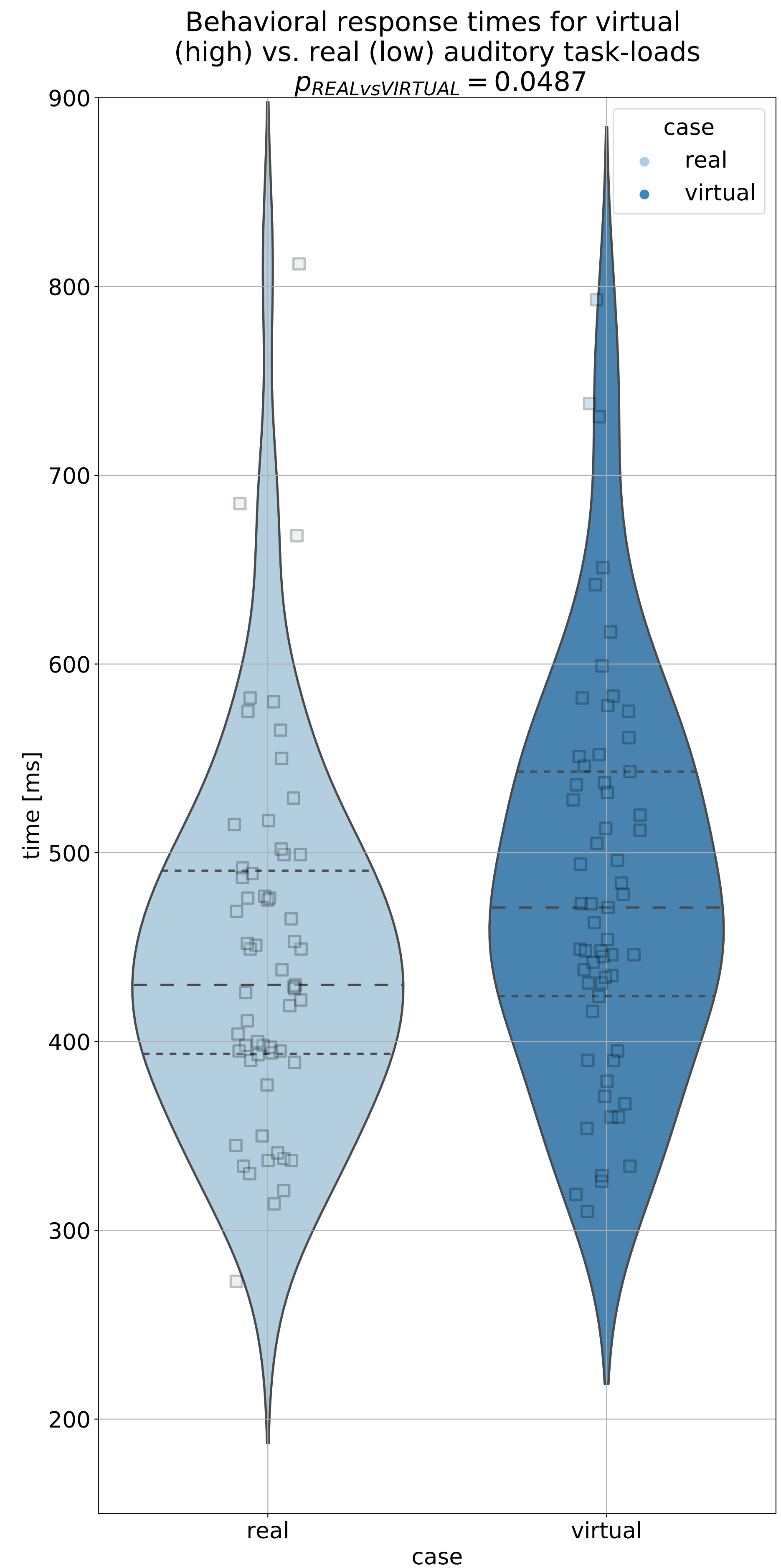


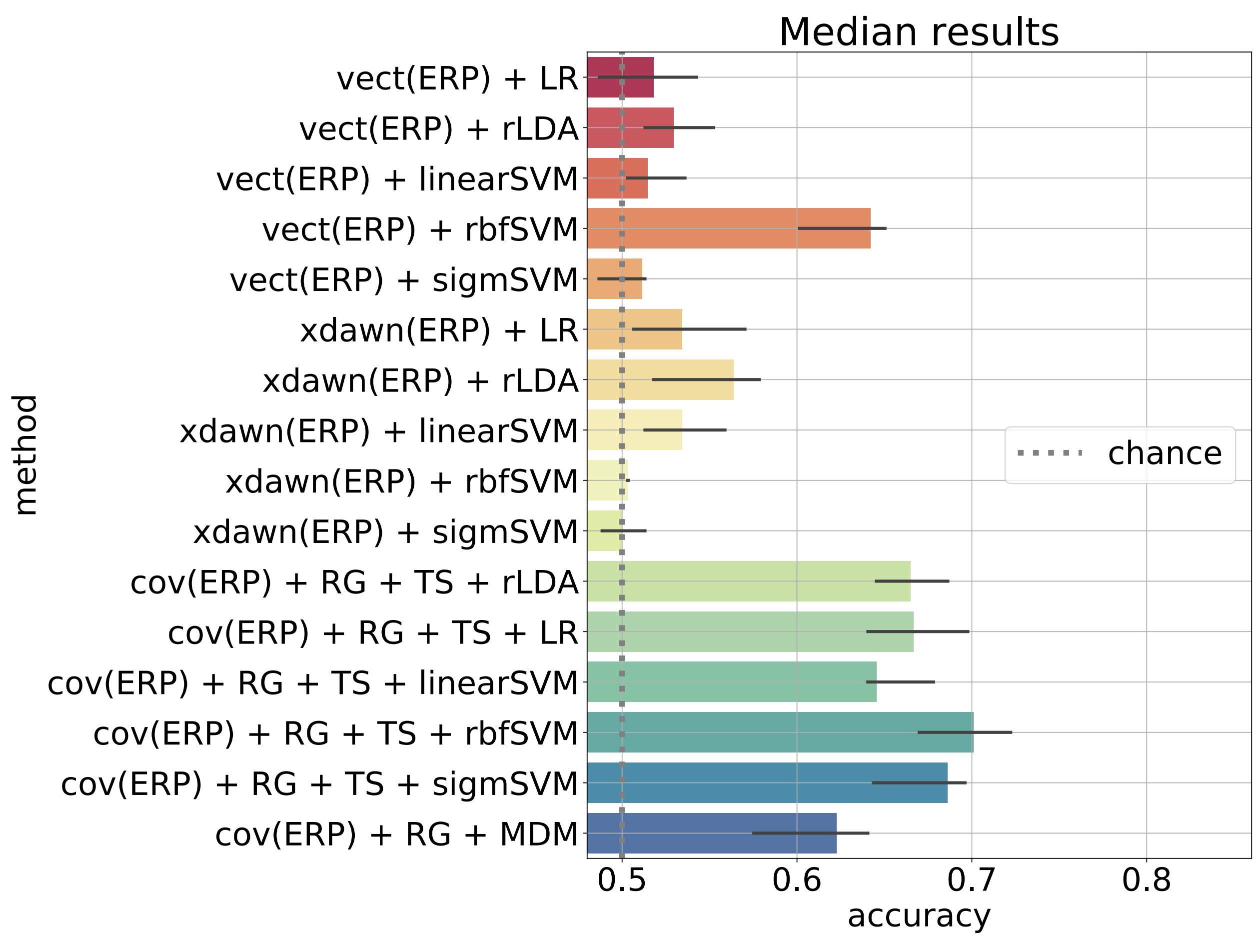
Results: Median accuracies of single subject dataset for training and testing using tactile and AUTD (time shifted) P300 responses

RESULTS

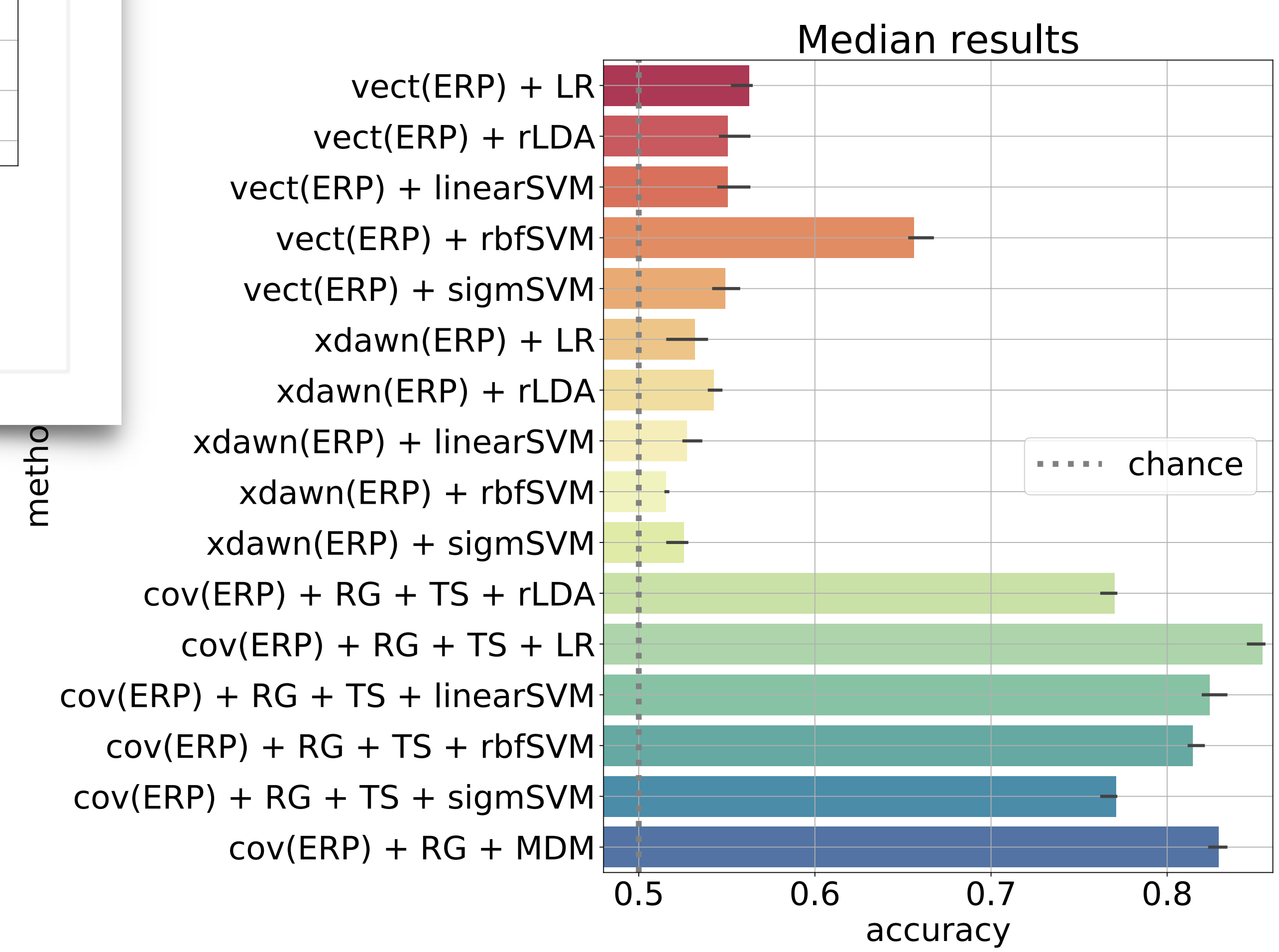
ERP SCALP TOPOGRAPHIES





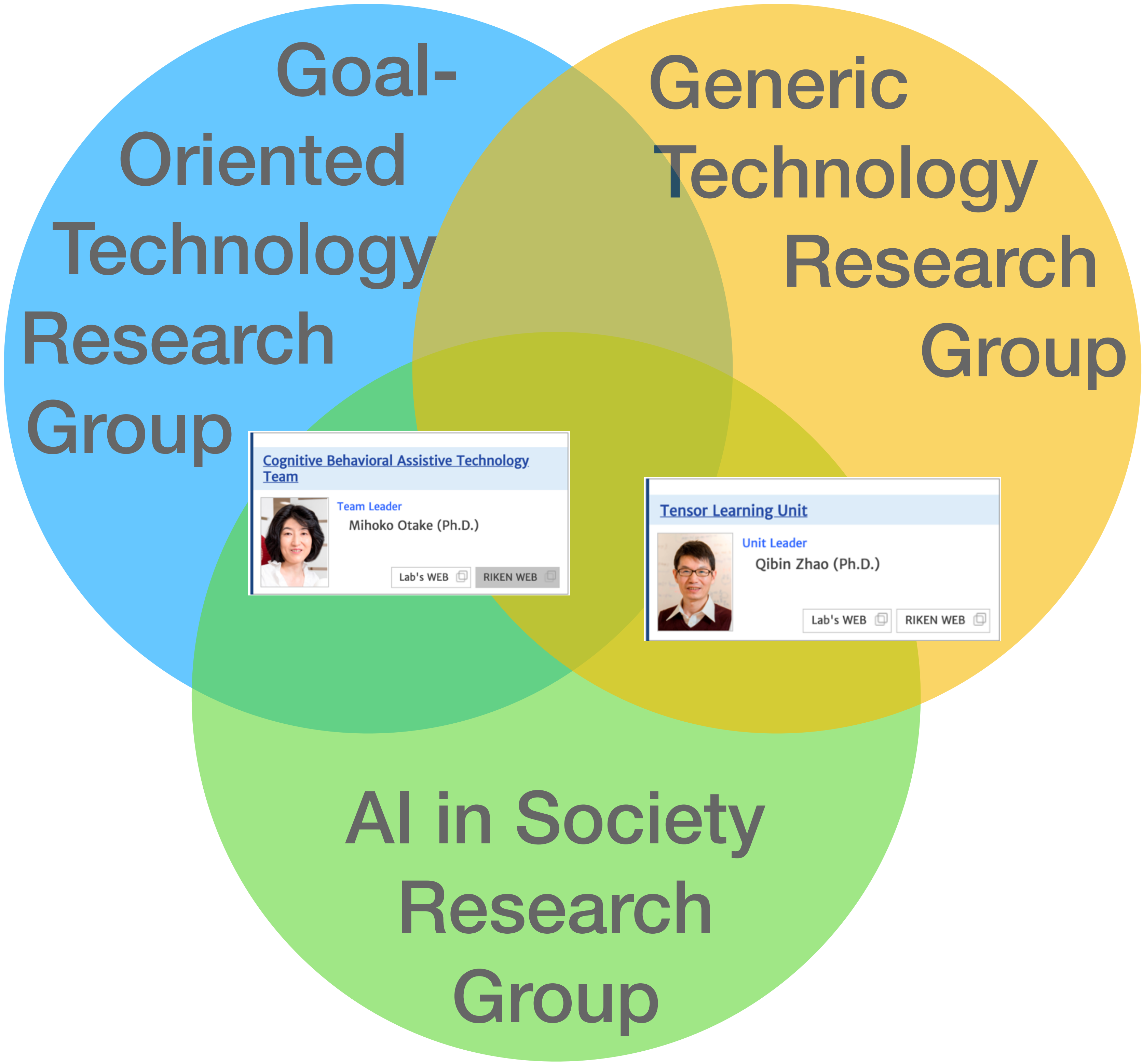


Results: Median accuracies of all subject dataset for training and testing using auditory (amplitude modulated) P300 responses



Results: Median accuracies of single subject dataset for training and testing using auditory (amplitude modulated) P300 responses

INTERNAL COLLABORATIONS OF CBAT TEAM WITHIN AIP



TENSOR–TRAIN–LAYER–BASED NEURAL NETWORK CLASSIFIERS

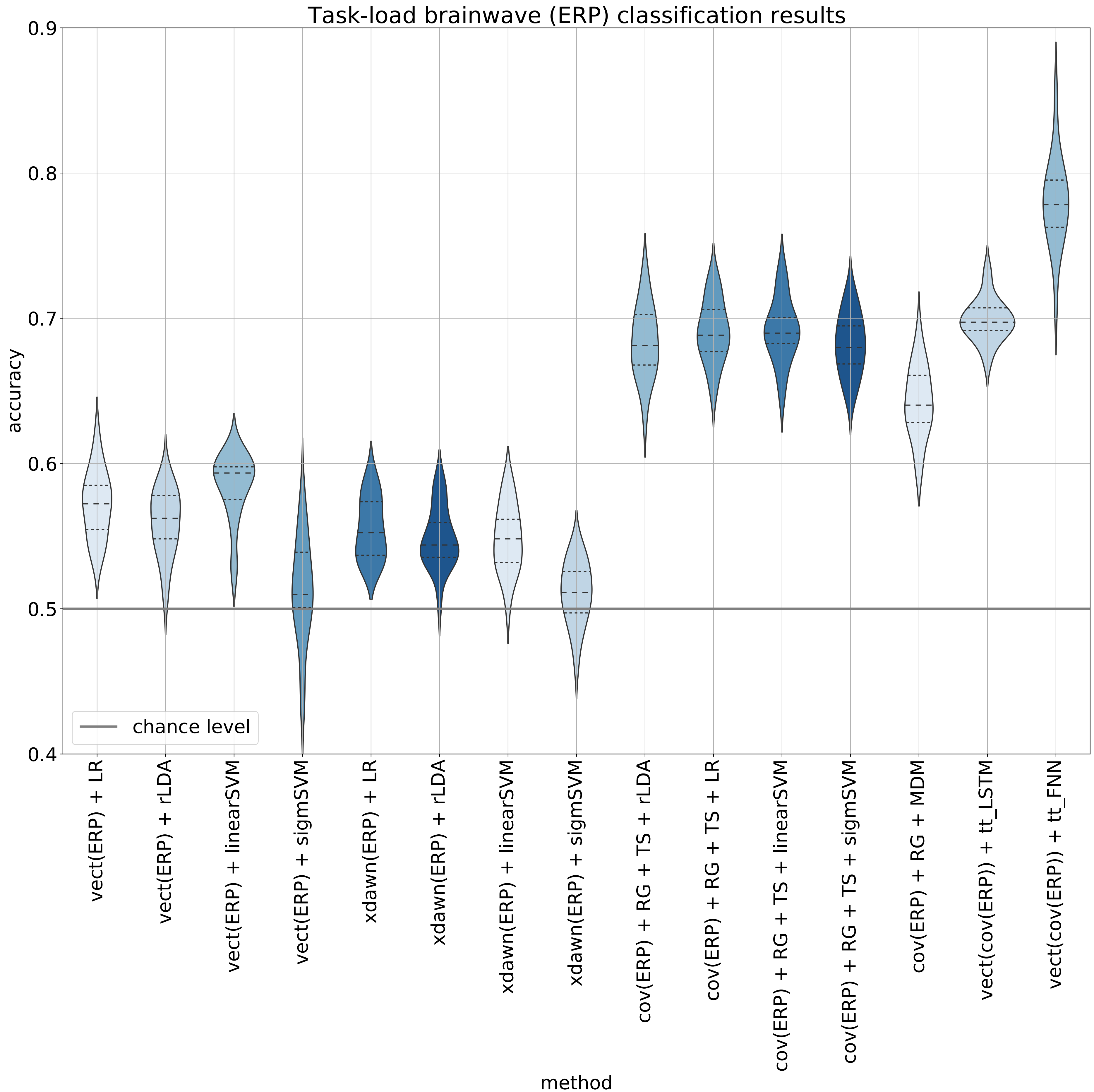
Deep neural networks recently started to gain attention in EEG community especially for the end-to-end processing setups. For our approach, we are interested in compacting machine learning architectures with possible application to wearable devices with limited computing powers. Tensor-train factorization (TTF) approaches have demonstrated an advantage of scaling multidimensional matrices to an arbitrary number of dimensions with a subsequent possibility to reshape a fully connected neural network layer into a tensor and then factorize it. The methodology as mentioned earlier is applied to compress large weight matrices of deep neural models within an entire end-to-end training. Using a trick from it is possible to tensor-train-factorize a weight matrix **W** of a fully connected neural network as a *d*-dimensional double-indexed tensor represented as

$$\widehat{\mathbf{W}}\left((i_1, j_1), (i_2, j_2), \ldots, (i_d, j_d)\right) = \mathcal{G}_1^*(i_1, j_1)\mathcal{G}_2^*(i_2, j_2)\ldots\mathcal{G}_d^*(i_d, j_d),$$

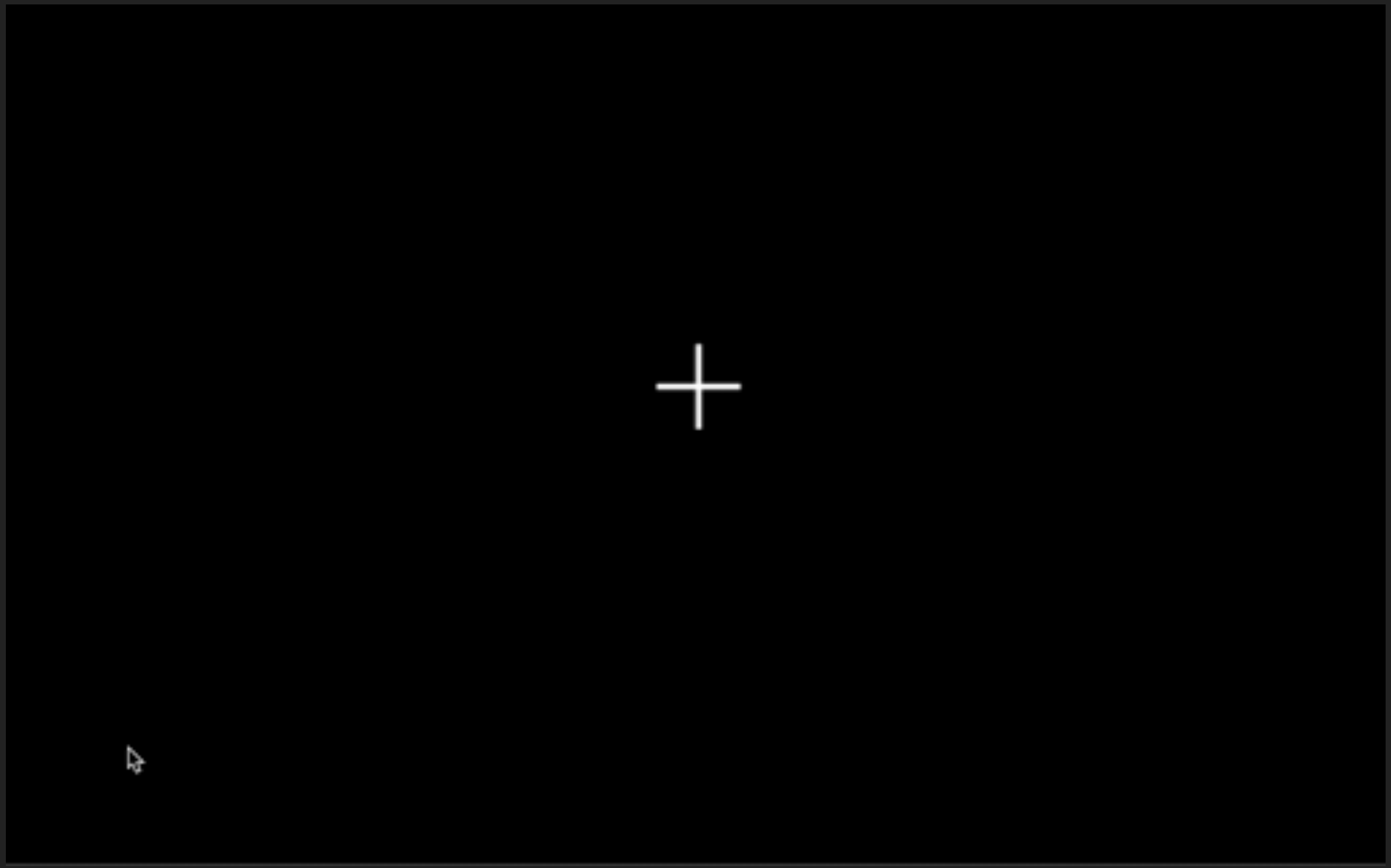
$$\text{where } \mathcal{G}_k^* \in \mathbb{R}^{m_k \times n_k \times r_{k-1} \times r_k}$$

are the so-called core tensors uniquely represented by indices (i_k, j_k). It is possible to compress the weight matrix **W** in a form of TT-formated low-rank tensors, approximately reconstructing the original one. We consider a two–layered neural network with *64* hidden units in each layer and replace both fully-connected layers by the TT-layers with all the TT-ranks in the network set to *4* and activation functions to ReLU. For an output of a fully connected final layer with sotfmax activations we choose two units. In order to prevent overfitting we train and test (using "unseen" during the training samples) the model with *50%/50%* balanced datasets. We also evaluate a recurrent neural network with LSTM units as a classifier with TT-layers applied to gating units. In this case the network has again *64* input LSTM units and TT-ranks set also to *4*.

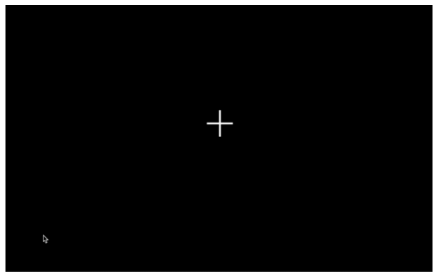
COMPARISON OF SHALLOW AND DEEP CLASSIFIERS



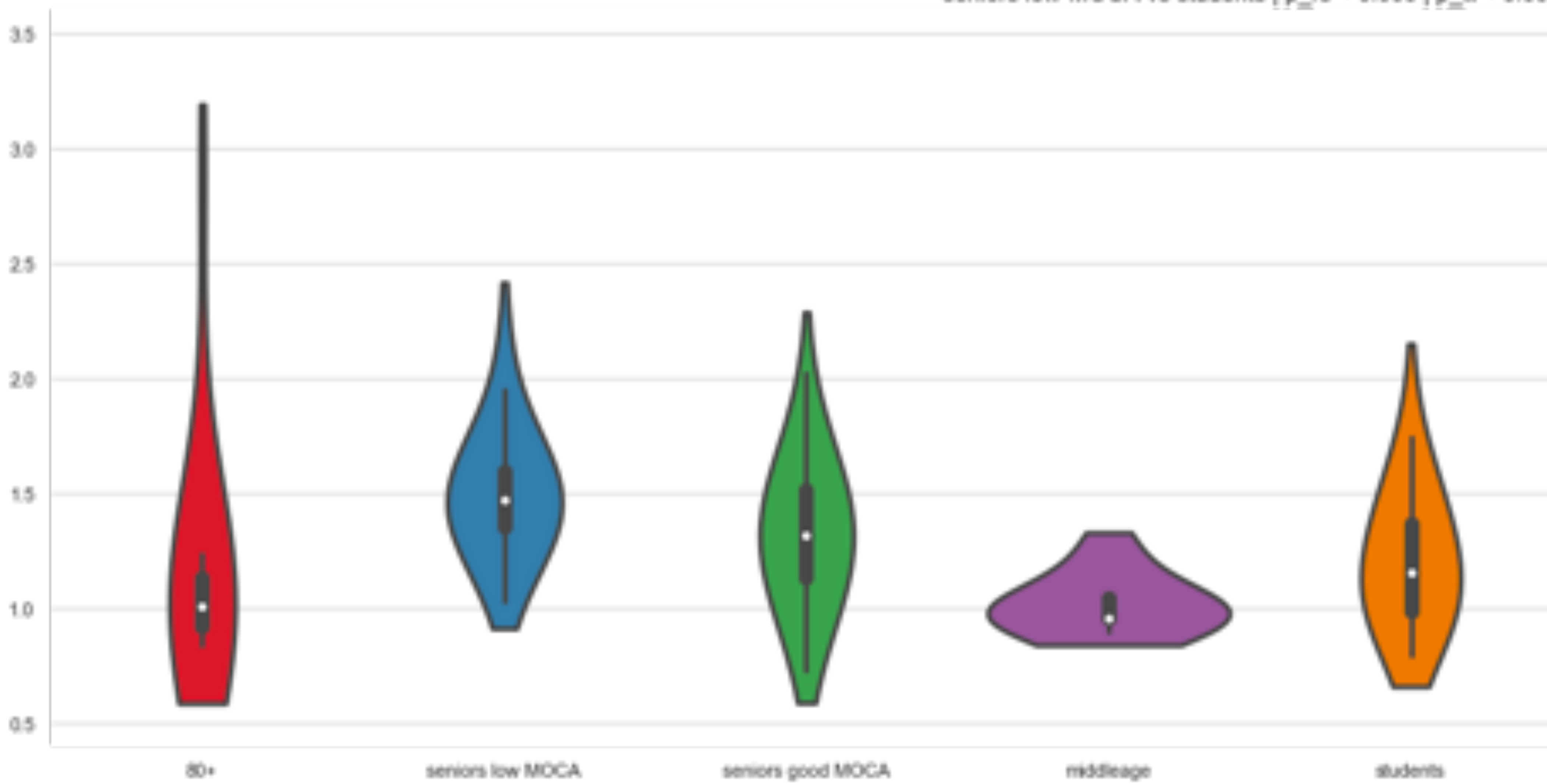
ODDBALL/P300-STYLE EXPERIMENT



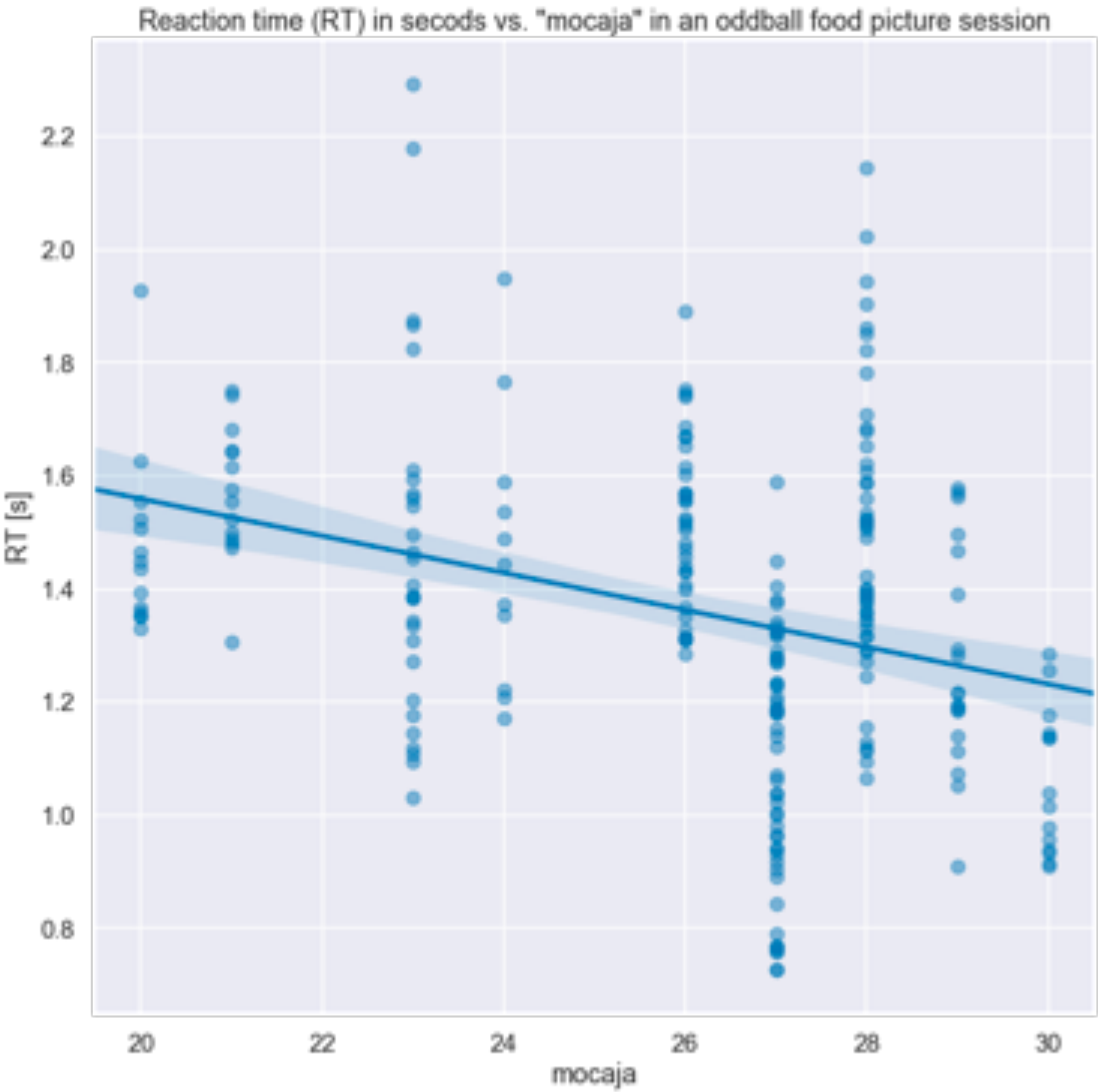
REACTION TIMES (HCI STYLE) FOR "FOOD RECOGNITION" SETUP



students vs. 80+ | p_rs < 0.047 | p_tt < 0.497
middleage vs. 80+ | p_rs < 0.055 | p_tt < 0.066
seniors low MOCA vs. 80+ | p_rs < 0.000 | p_tt < 0.000
seniors good MOCA vs. 80+ | p_rs < 0.001 | p_tt < 0.043
middleage vs. students | p_rs < 0.055 | p_tt < 0.066
middleage vs. good-MOCA | p_rs < 0.003 | p_tt < 0.006
middleage vs. low-MOCA | p_rs < 0.000 | p_tt < 0.000
seniors good-MOCA vs low-MOCA | p_rs < 0.000 | p_tt < 0.000
seniors good-MOCA vs students | p_rs < 0.001 | p_tt < 0.002
seniors low-MOCA vs students | p_rs < 0.000 | p_tt < 0.000



REACTION TIMES IN AN ODDBALL EXPERIMENT



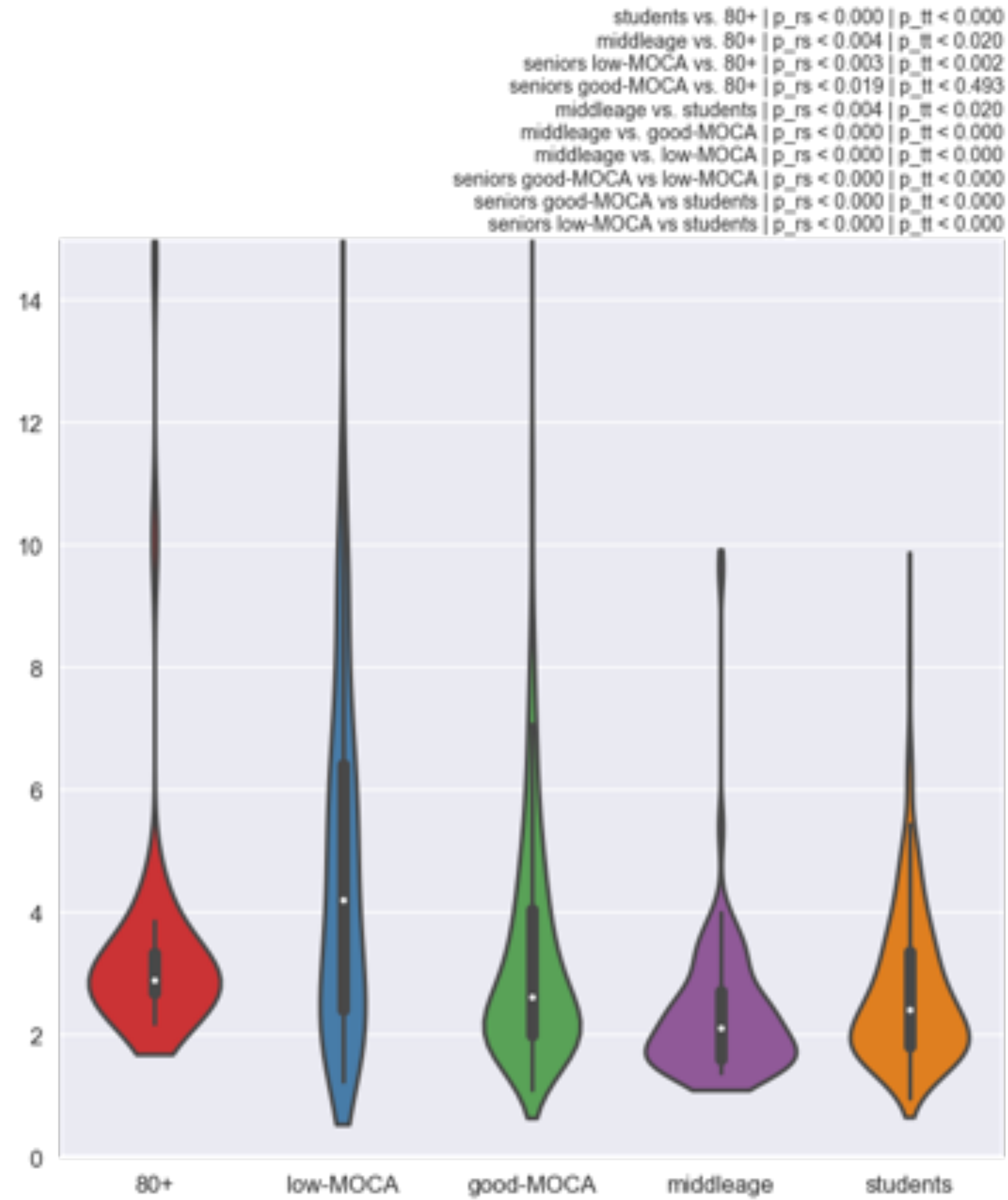
OLS Regression Results

Dep. Variable:	RT	R-squared:	0.284
Model:	OLS	Adj. R-squared:	0.262
Method:	Least Squares	F-statistic:	13.03
Date:	Mon, 17 Jun 2019	Prob (F-statistic):	4.17e-14
Time:	09:56:17	Log-Likelihood:	4.0481
No. Observations:	238	AIC:	7.904
Df Residuals:	230	BIC:	35.68
Df Model:	7		
Covariance Type:	nonrobust		

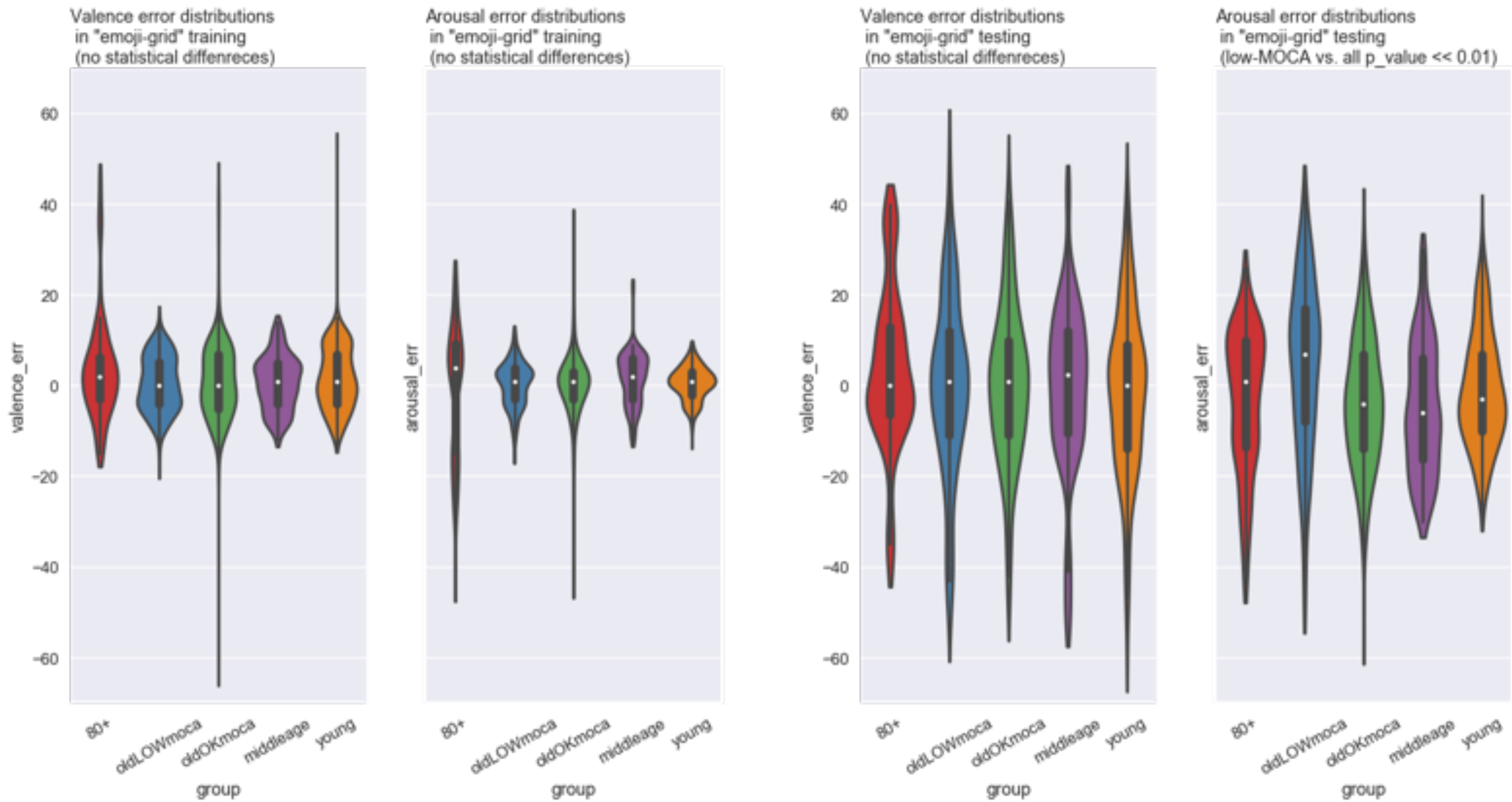
	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	2.6862	0.257	10.453	0.000	2.180	3.192
mocaja	−0.1610	0.033	−4.887	0.000	−0.226	−0.096
mocajca	0.1358	0.034	3.965	0.000	0.068	0.203
symbola	−0.0013	0.002	−0.862	0.389	−0.004	0.002
wms1a	−0.0411	0.011	−3.745	0.000	−0.063	−0.019
wms2a	0.0486	0.010	4.797	0.000	0.029	0.069
digit_f_ta	−0.0571	0.013	−4.236	0.000	−0.084	−0.031
mmseta	0.0015	0.011	0.143	0.887	−0.019	0.022

Omnibus:	16.410	Durbin-Watson:	1.170
Prob(Omnibus):	0.000	Jarque-Bera (JB):	18.395
Skew:	0.581	Prob(JB):	0.000101
Kurtosis:	3.710	Cond. No.	1.33e+03

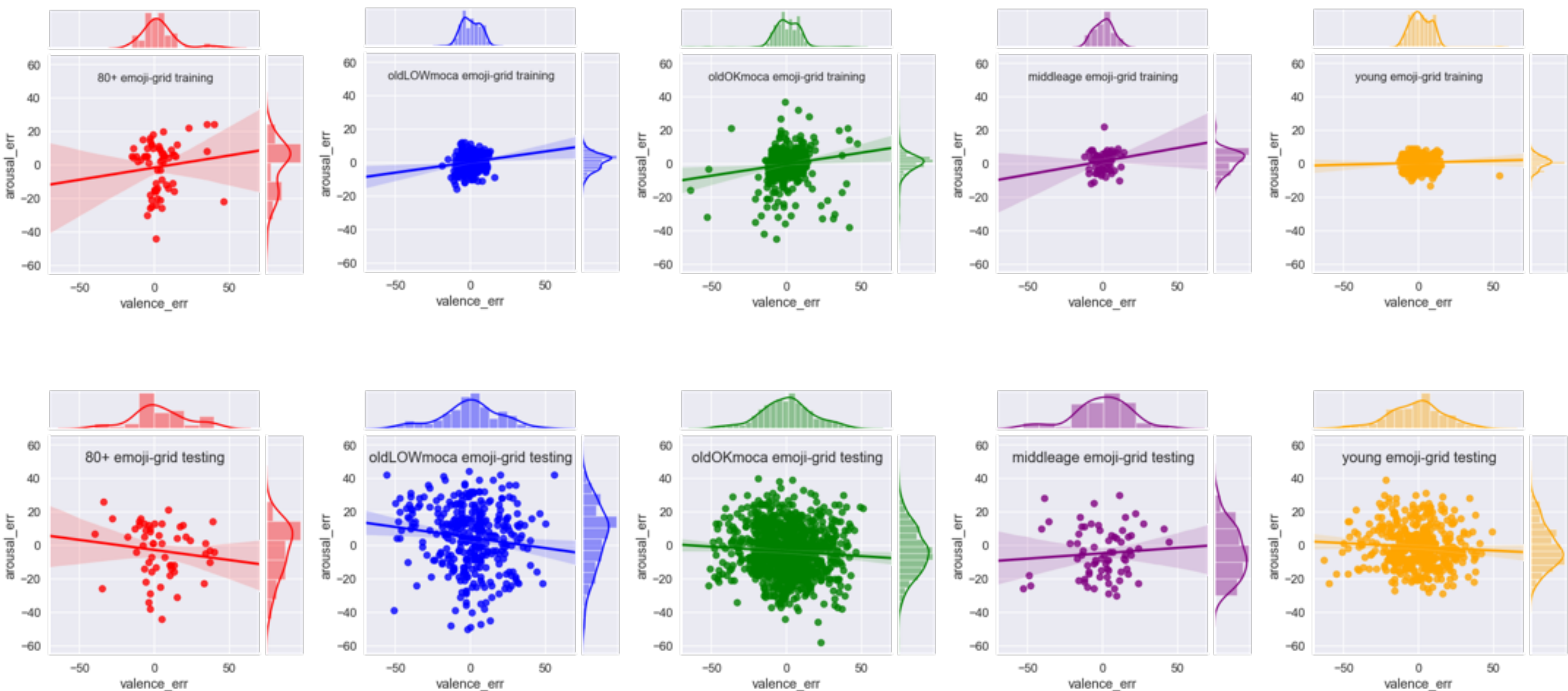
REACTION TIMES (RT)



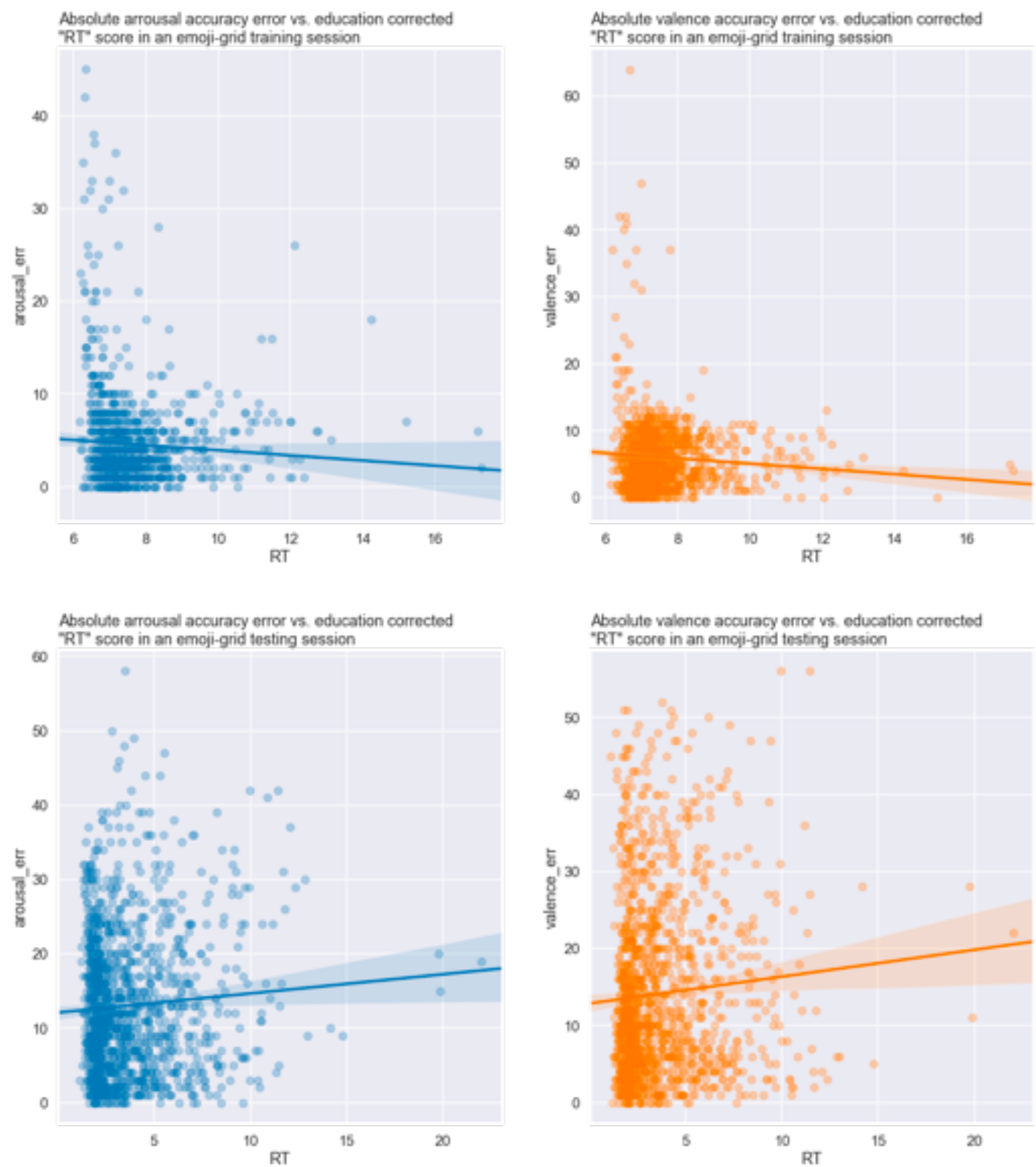
EMOJI–GRID ACCURACY TRAINING VS. TESTING RESULTS



EMOJI-GRID ACCURACY TRAINING VS. TESTING DISTRIBUTIONS

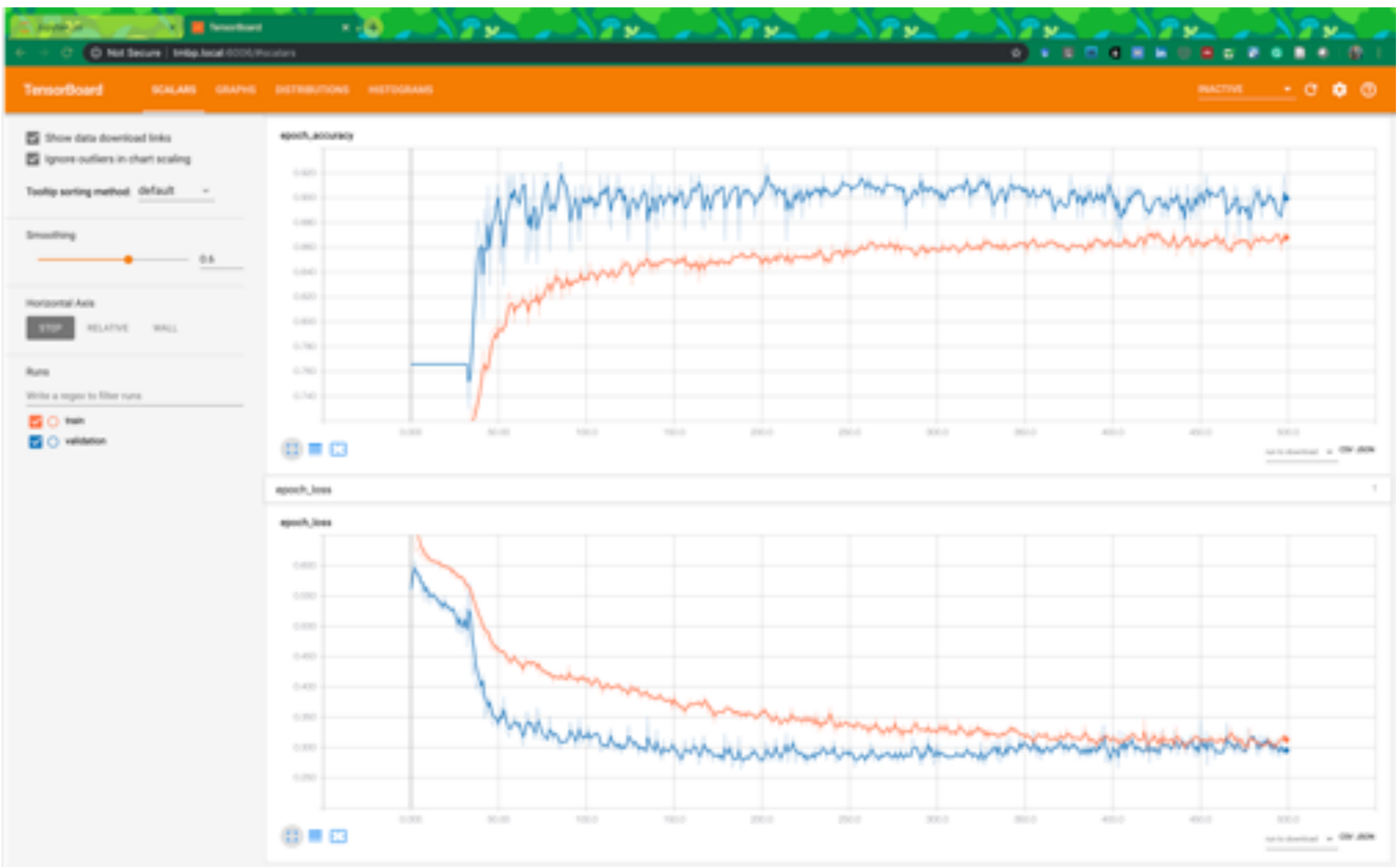


EMOJI-GRID ACCURACY TRAINING VS. TESTING FOR "MOCA-SCORED" SENIORS ONLY



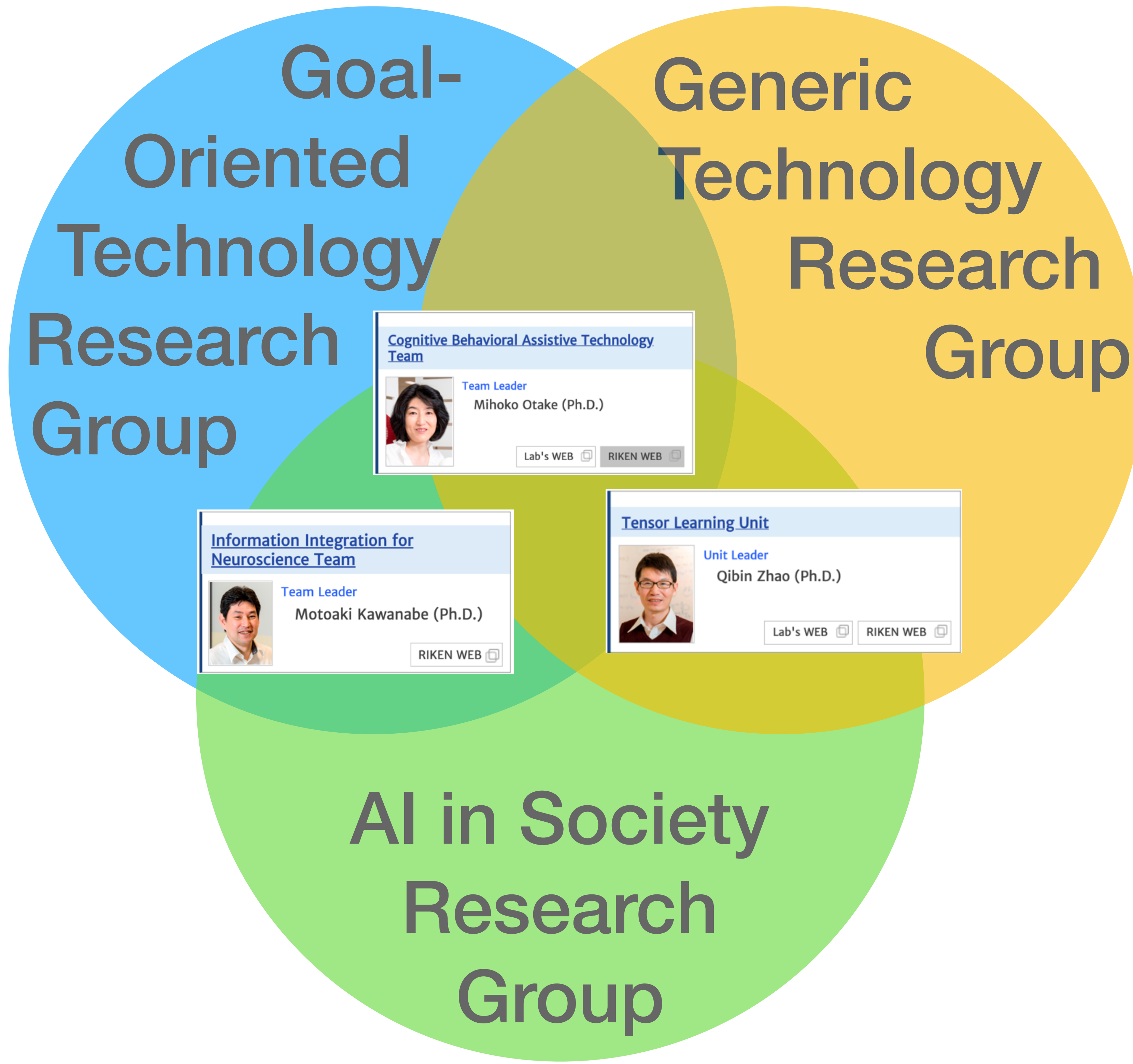
TT_FNN CLASSIFICATION RESULTS WITH RT (DEMENTIA VS. HEALTHY AT MOCA=25 THRESHOLD)

TEST SET ACCURACY ~90%

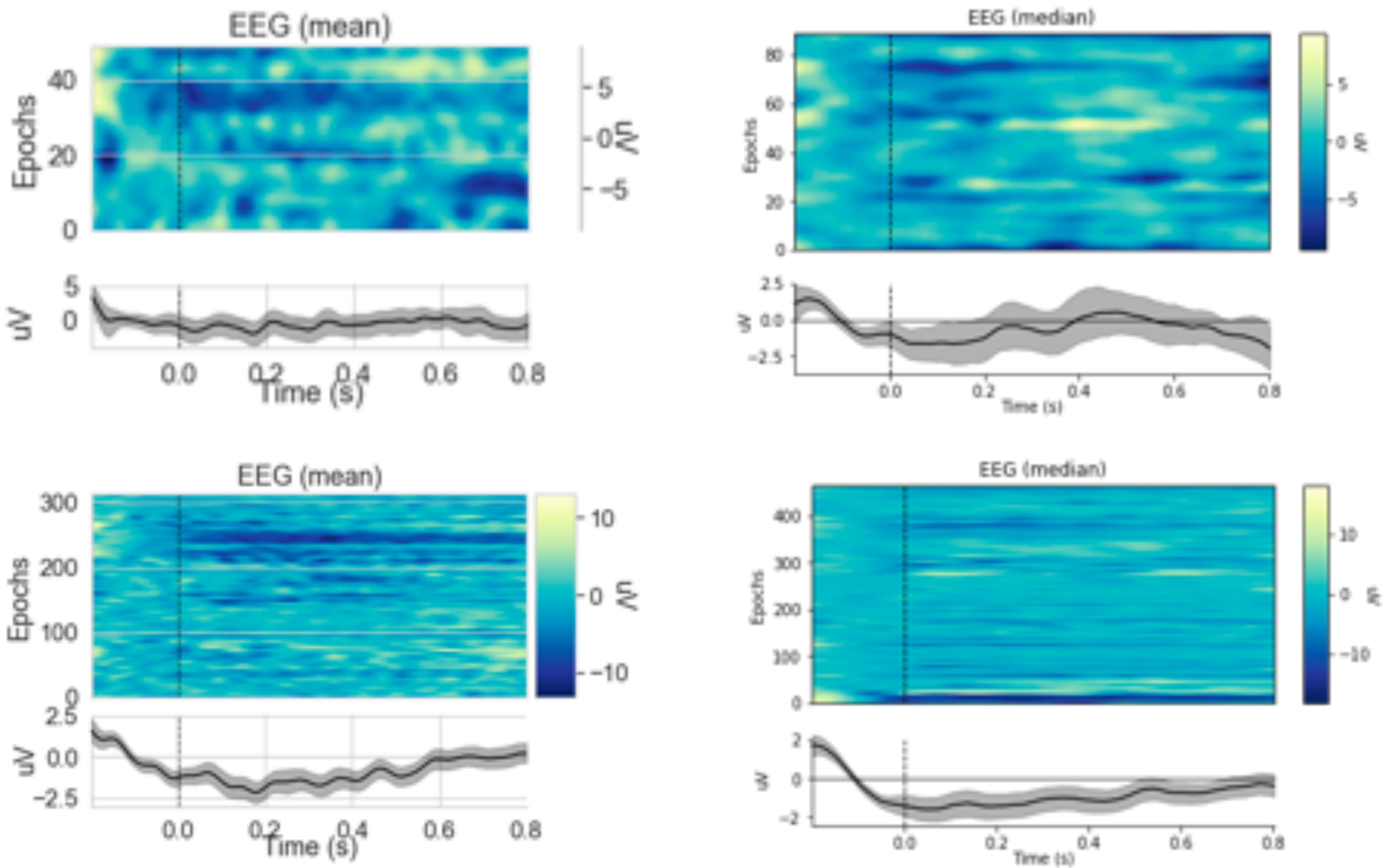


TT_FNN CLASSIFICATION WITHOUT RT AND ACCURACY DROPS TO ~69%

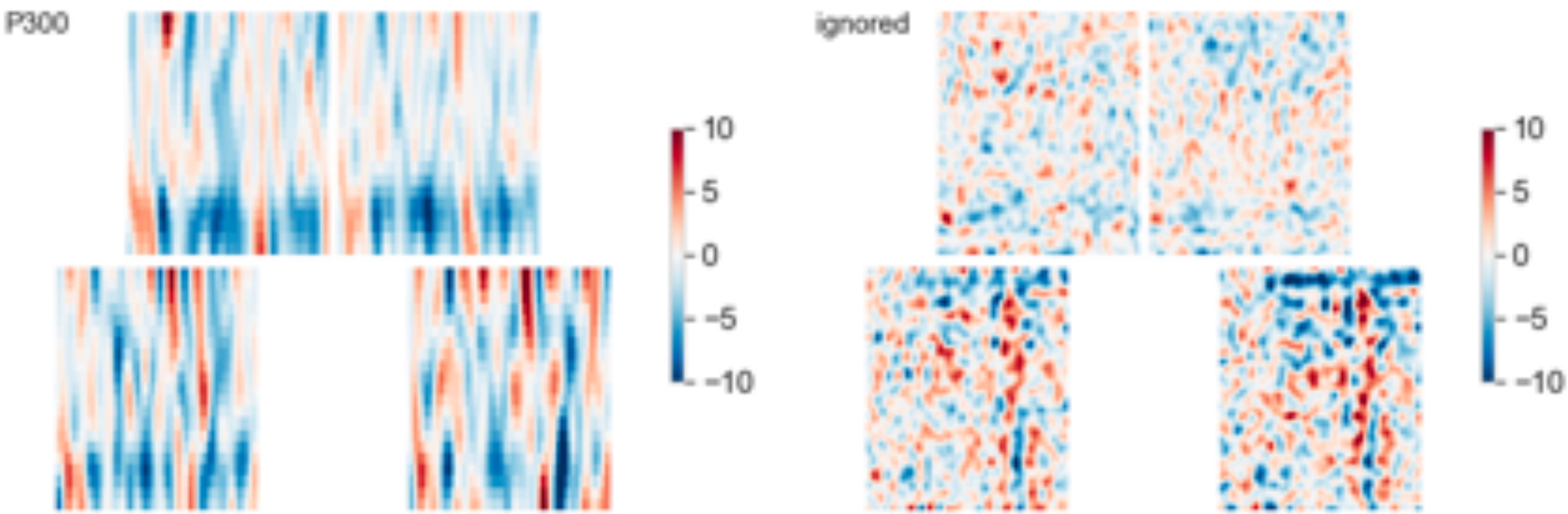
INTERNAL COLLABORATIONS OF CBAT TEAM WITHIN AIP



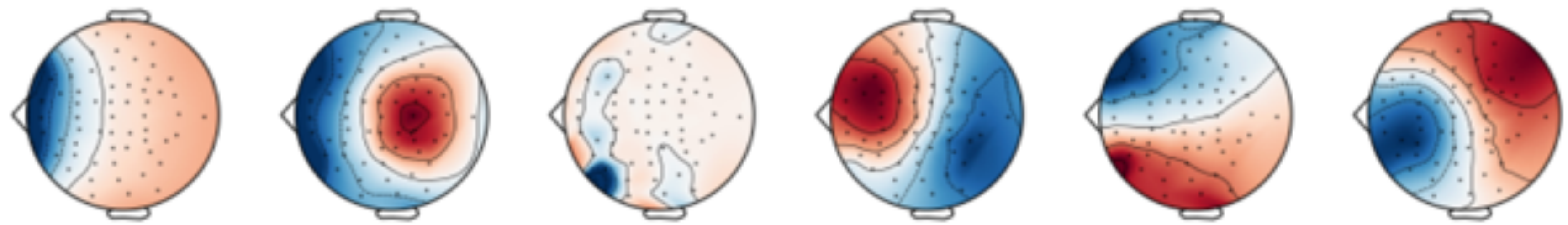
BEYOND ERP AND TOWARD MICRO STATES



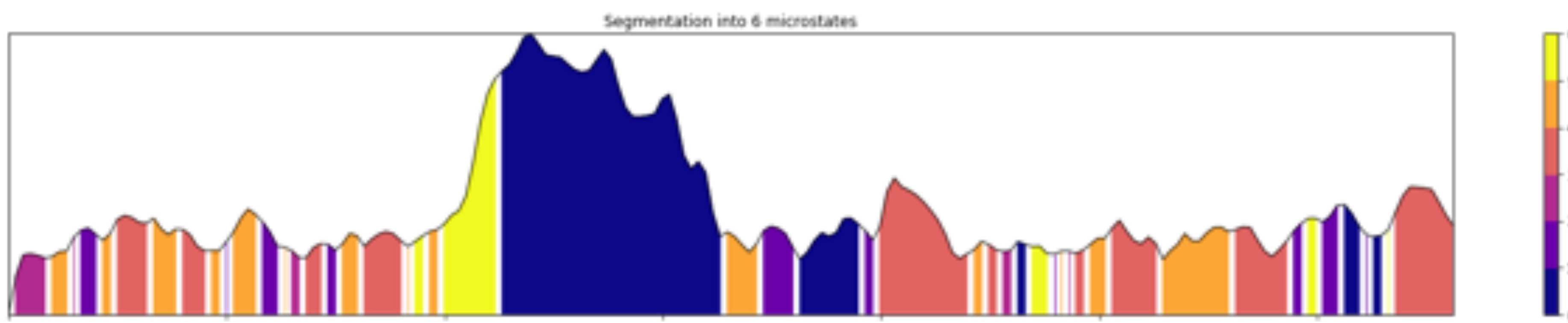
BEYOND ERP AND TOWARD MICRO STATES

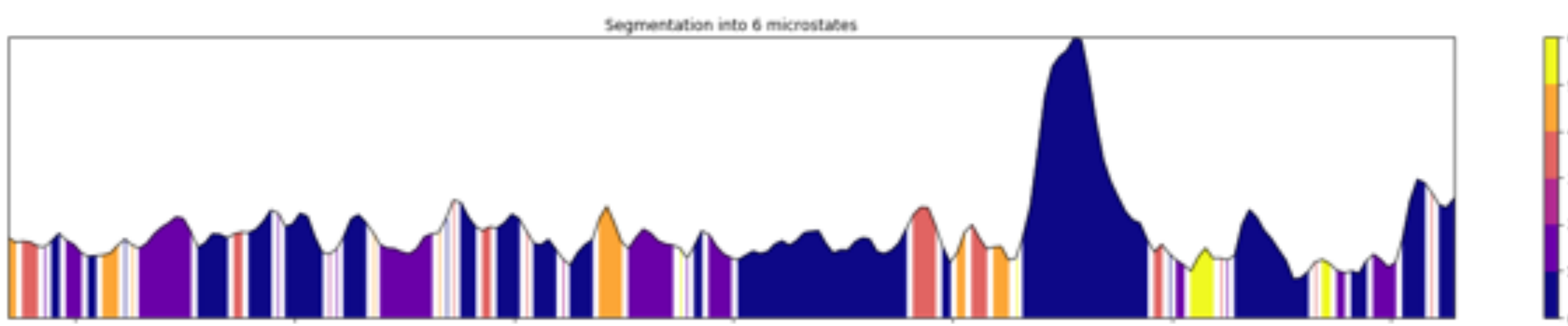


EEG-MICRO-STATES & AUTO-ENCODERS

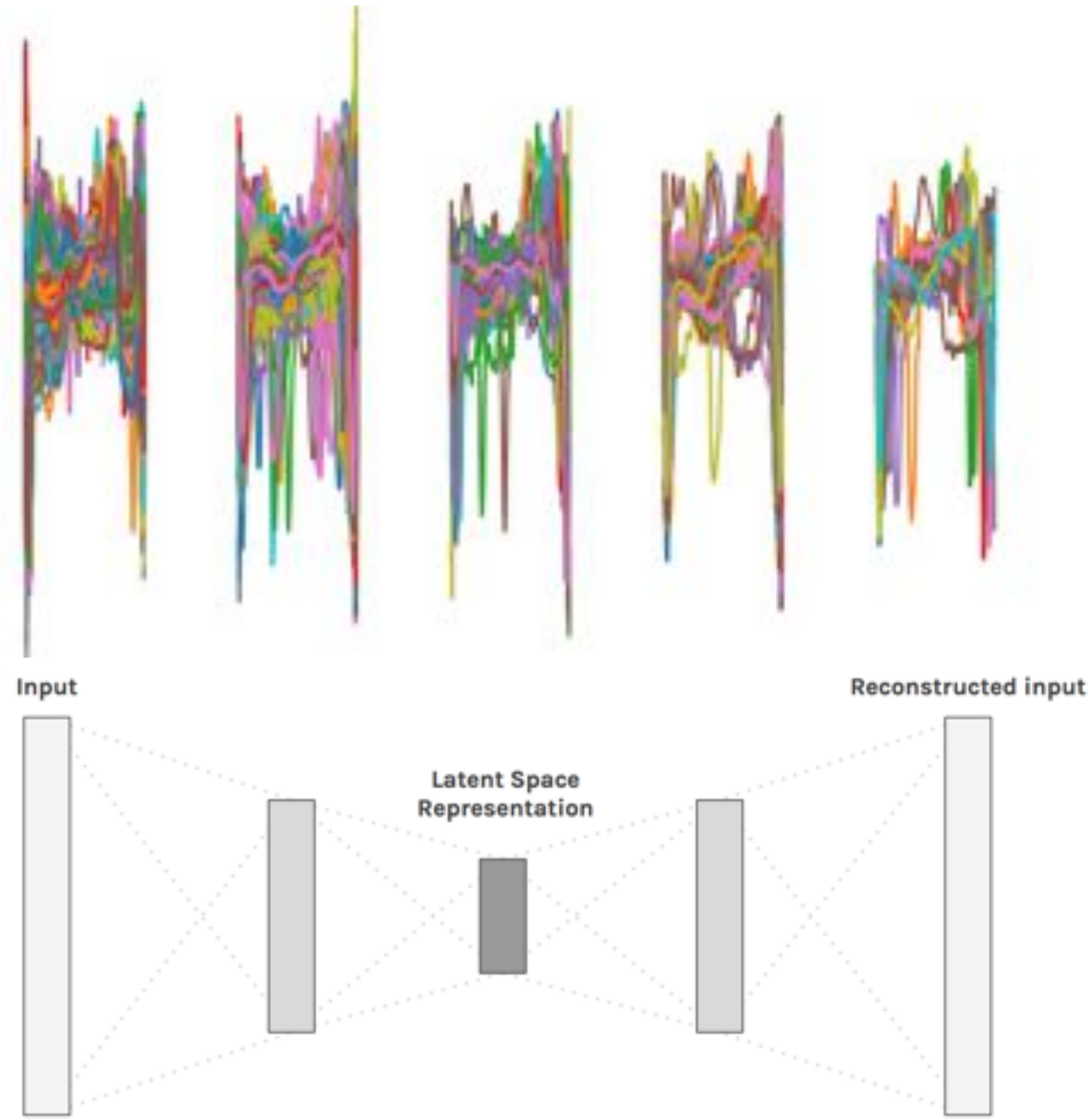


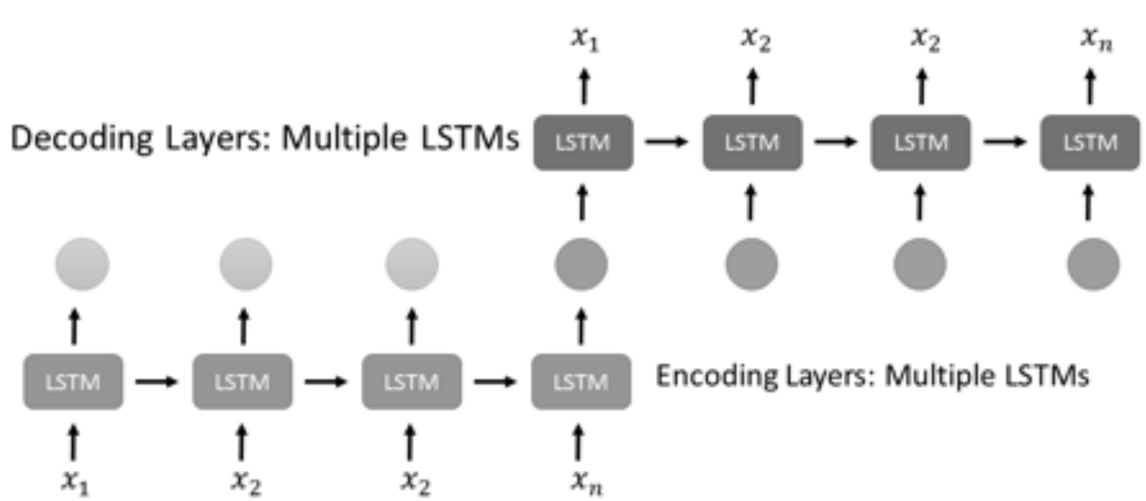
1 2 3 4 5 6

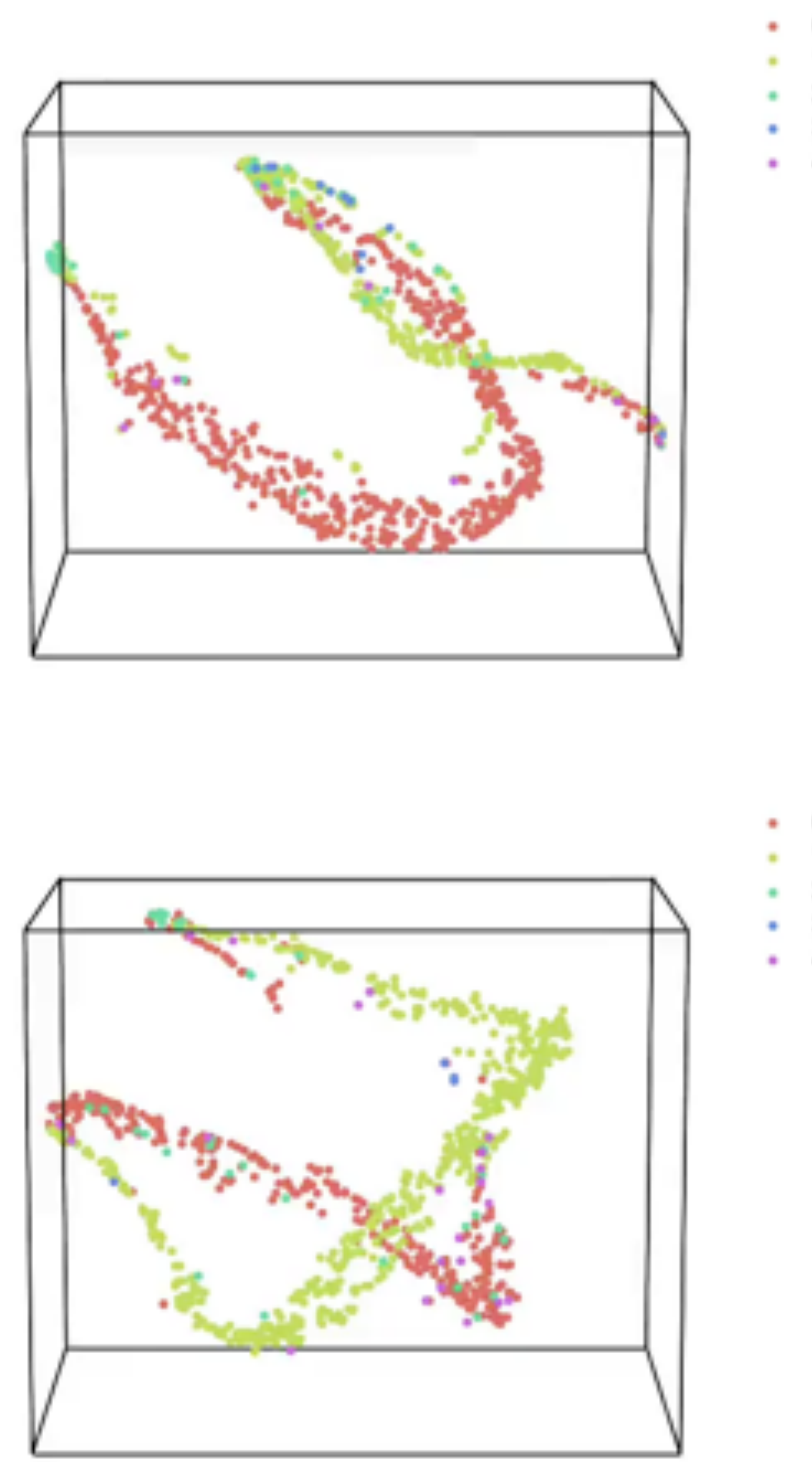




PRELIMINARY RESULTS OF DEEP LEARNING DATA-DRIVEN SYMPTOM DISCOVERY







CONCLUSIONS

- The discussed pilot results of the brainwave- and behavioral-based dementia biomarker elucidation offer an advance in a search for new imaging methods for elderly leading to a life improvement and healthcare cost minimization
- In a subsequent step of our approach, we will focus on the collection of more data in order to train better deep learning models for further improvements of the proposed approach