DigitalWorld 2024 Congress

The Sixteenth International Conference on eHealth, Telemedicine, and Social Medicine

eTELEMED 2024

Barcelona, Spain, May 26-30, 2024

Presenter: **Christoph Anders**

Load Induction then Simultaneous Relaxation: Insights from Multi-Modal Time-Series Data measured with Low-Cost Wearable Sensors

Christoph Anders[1], Sai Siddhant Gadamsetti[1], Nico Steckhan[1], Bert Arnrich[1]

[1]Digital Health - Connected Healthcare, University of Potsdam, Hasso Plattner Institute

Contact email: christoph.anders@hpi.de

Hasso Plattner Institut

Christoph Anders

Christoph Anders received the master's degree in IT Systems Engineering from the Hasso Plattner Institute, Germany, in 2021 with his master thesis titled '*Experimental evaluation of data preprocessing methods for time series classification on brain activity data*'. He is currently a doctoral student majoring in Digital Health at the Connected Healthcare chair at the Hasso Plattner Institute.

His research interests lie in the classification of mental states using wearables, taking research out of laboratory environments into uncontrolled environments, and enabling real-life applications to improve learning. Throughout his research, Christoph Anders has initiated and successfully completed multiple studies with human participants on mental states, from obtaining ethical clearances over data recordings and analyses to publications.

Web: https://hpi.de/arnrich/people/christoph-anders.html

Aims and contributions of our publication

In our paper we aimed at:

- analyzing group-wide processes (and) to evaluate mental workload and stress
- investigating the reliability of wearable sensor systems
- validating our proposed pipeline simultaneous recording of physiological data

Contributions of our study are fourfold:

1. development and validation of a study protocol ‘simultaneous data collection’ applicable for controlled and uncontrolled environments
2. development and publication of a python implementation for synchronous streaming of the Empatica E4 and Muse S wearables (Windows only)
3. collection, cleaning, and publication of a dataset of five sessions mental workload induction and subsequent relaxation for five pairs (25 hrs)
4. feature importance analysis revealed moderate to strong positive correlations between activity and features (statistical EEG features; heart rate variability)

Related Work

- Plenty interventions to reduce the stress levels of participants were investigated, such as:
  - Exposure to music and nature sounds [15]
  - Mind-body connection courses designed to reduce anxiety [16]
  - Combination of cognitive, somatic, dynamic, emotive and hands-on [17]

- Contradicting results w.r.t. artefacts: no substantial artefacts were present in mobile EEG readings, except for frontal recording sites [23] but significantly distorted recordings from electrodes at frontal, temporal, and ear positions were also reported [24]

- Measurements of synchronicity of event-related responses recorded with wearable sensors are rarely but effectively performed (i.e., with the Oddball paradigm) [13]

What are Mental States?

- “Stress can be defined as a state of worry or mental tension caused by a difficult situation.” (according to the WHO)

- “Mental workload may be viewed as the difference between the capacities of the information processing system that are required for task performance to satisfy performance expectations and the capacity available at any given time.” [1]

⇒ Stress and Mental Workload are connected

! If elevated over a prolonged time, these mental states can lead to adverse consequences, such as rising heart rate and blood pressure [2, 3], which in turn are connected with the risk of coronary heart disease and hypertension!

How to measure Mental States?

- Performance measures strongly highlight situations of mental overload
  - \+ Task performance, i.e. correct answers or time required to answer
  - \- Require previous knowledge or well-defined tasks

- Subjective measures are inaccurate for high vigilance and cognitive overload
  - \+ Well-studied questionnaires exist, such as NASA-TLX, PSS, etc.
  - \- Require time and active, truthful participation

- Physiological measures are objective and used by many experts
  - \+ Various signals can be recorded, such as Electroencephalography (EEG), Electrocardiogram (ECG), etc.
  - \- Require specialized equipment and trained staff

⇒ We decided to combine all of these measures to get a complete picture

Which measures were used in this study?

- Performance in pre-defined Cognitive Load Induction tasks
  - Response time, stimulus time, and accuracy

- Subjective answers to validated questionnaires
  - Brunel Mood Scale Questionnaire (BRUMS-Q), Stanford Sleepiness Scale (SSS), Visual Analogue Scale to Evaluate Fatigue Severity (VAS-F), as well as five-point Likert scales

- Objective measurements from well-studied measurement methods
  - EEG to discriminate task load, task type, and task difficulty [6], and ECG to quantify and predict stress reduction [8]

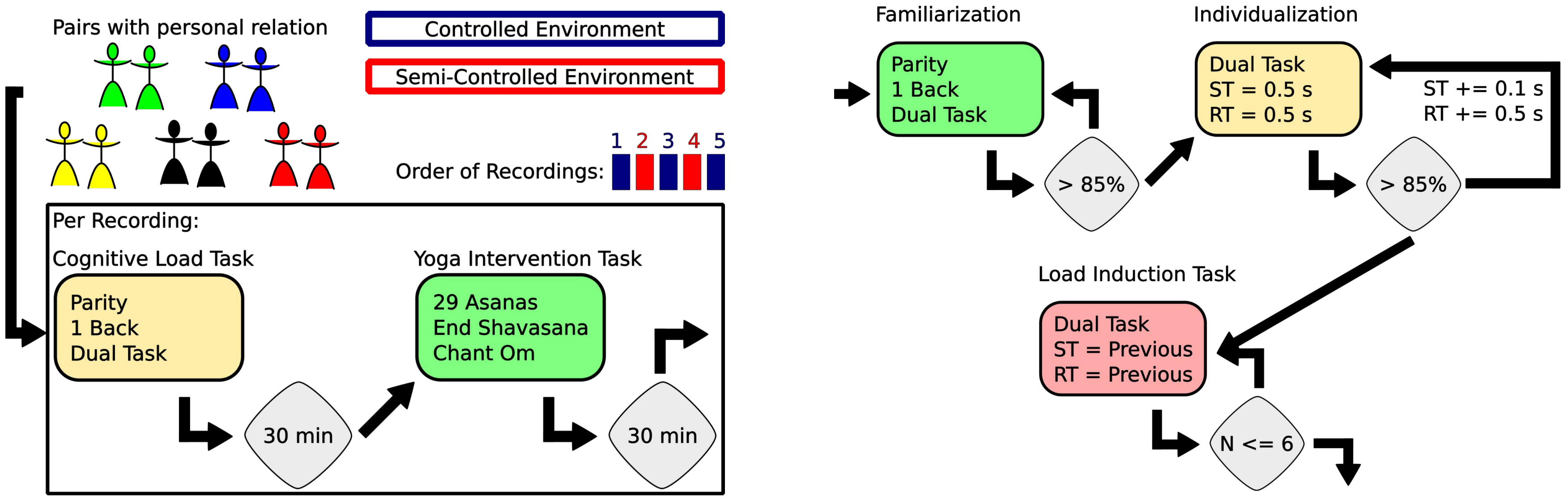
What works for one usually works for two

- For MW, it was found that by unconscious synchronization of brain activity across individuals, these individuals might utilize more mental resources than each individual alone would be able to! [9] This was studied in:
  - communication [10]
  - learning processes between teachers and students [11]
  - personal bond between individuals was found to be a modulator [12]

! Analysis of synchronization requires not only two data streams, but also the (constant) quantification of the offset between these two data sources!

⇒ We decided to develop a similarity analysis pipeline using the well-studied Oddball paradigm [13]

Study Design



Yoga Instructions Video: https://www.youtube.com/watch?v=6hZIzMpHl-c

Ethical approval by the Institutional Review Board (application number 69/2023)

Which wearables were used?

Empatica E4

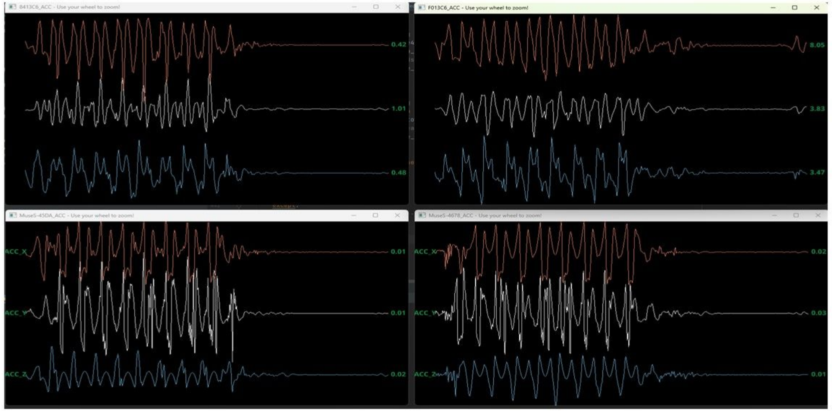
Photoplethysmography, Electrodermal Activity, and accelerometer sensors

Muse S

Four EEG sensors plus reference, PPG, gyroscope, and accelerometer sensors

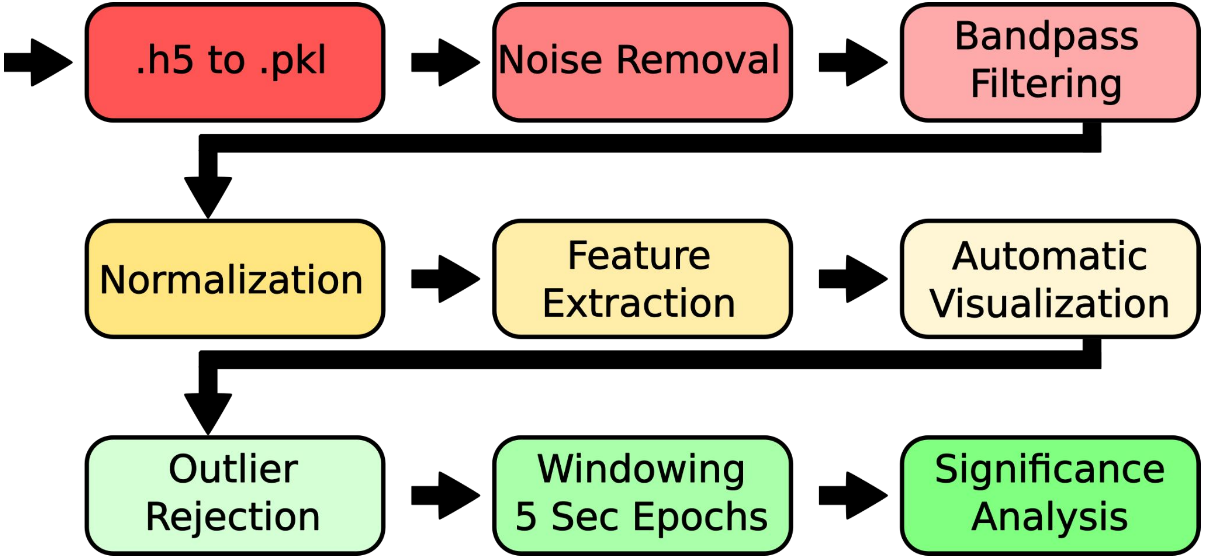
Data Collection Pipeline

- Implemented in Python 3.9 and building on top of PyLSL as well as the Empatica E4 streaming server
- For each wearable sensor, a separate BLED112 Bluetooth Dongle had to be utilized (i.e., in total four dongles were used for this setup)
- Run via a command line interface and allows real-time stream monitoring



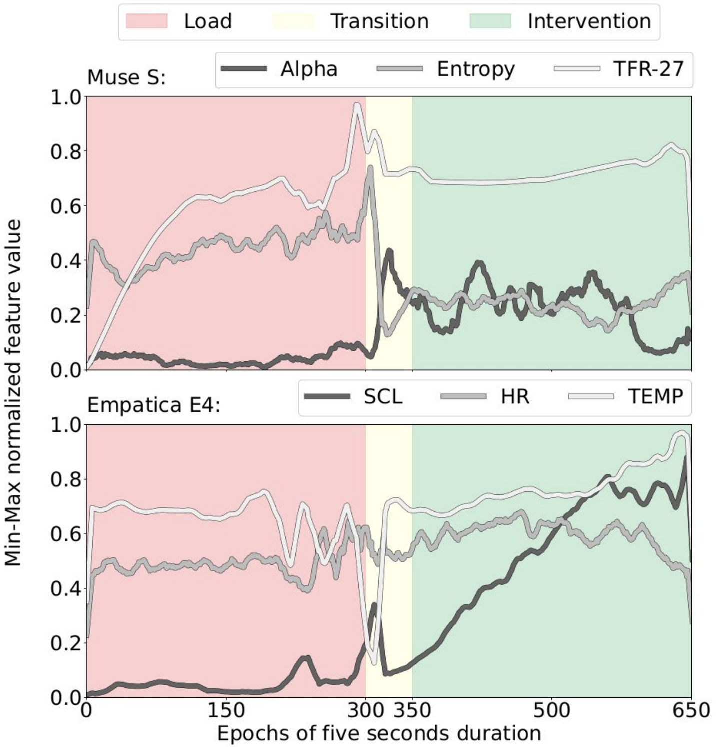
Code: https://github.com/siddhant61/StreamSense - contributions are welcome!

Data Preprocessing Pipeline

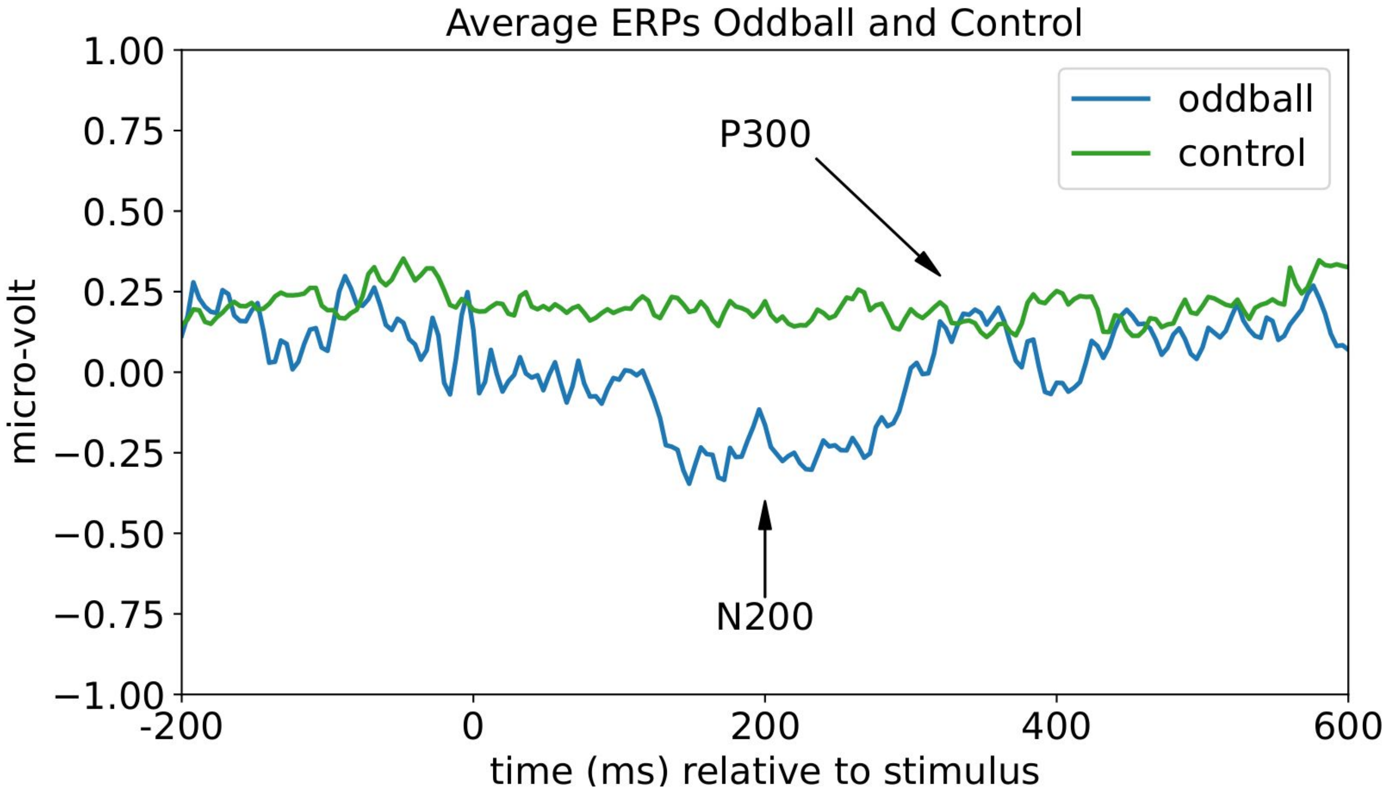


Code: https://github.com/siddhant61/ProSense - contributions are welcome!

Data Exploration



Event Related Potential (ERP) Analysis



Which Features were used?

Data was split in windows of five seconds duration, and extracted were:

<u>Muse Features</u>

at each electrode individually:

- power spectral densities
- band-powers
- band power ratios
- spectral entropy
- various statistical features such as the standard deviation, the variance, and skewness
- …

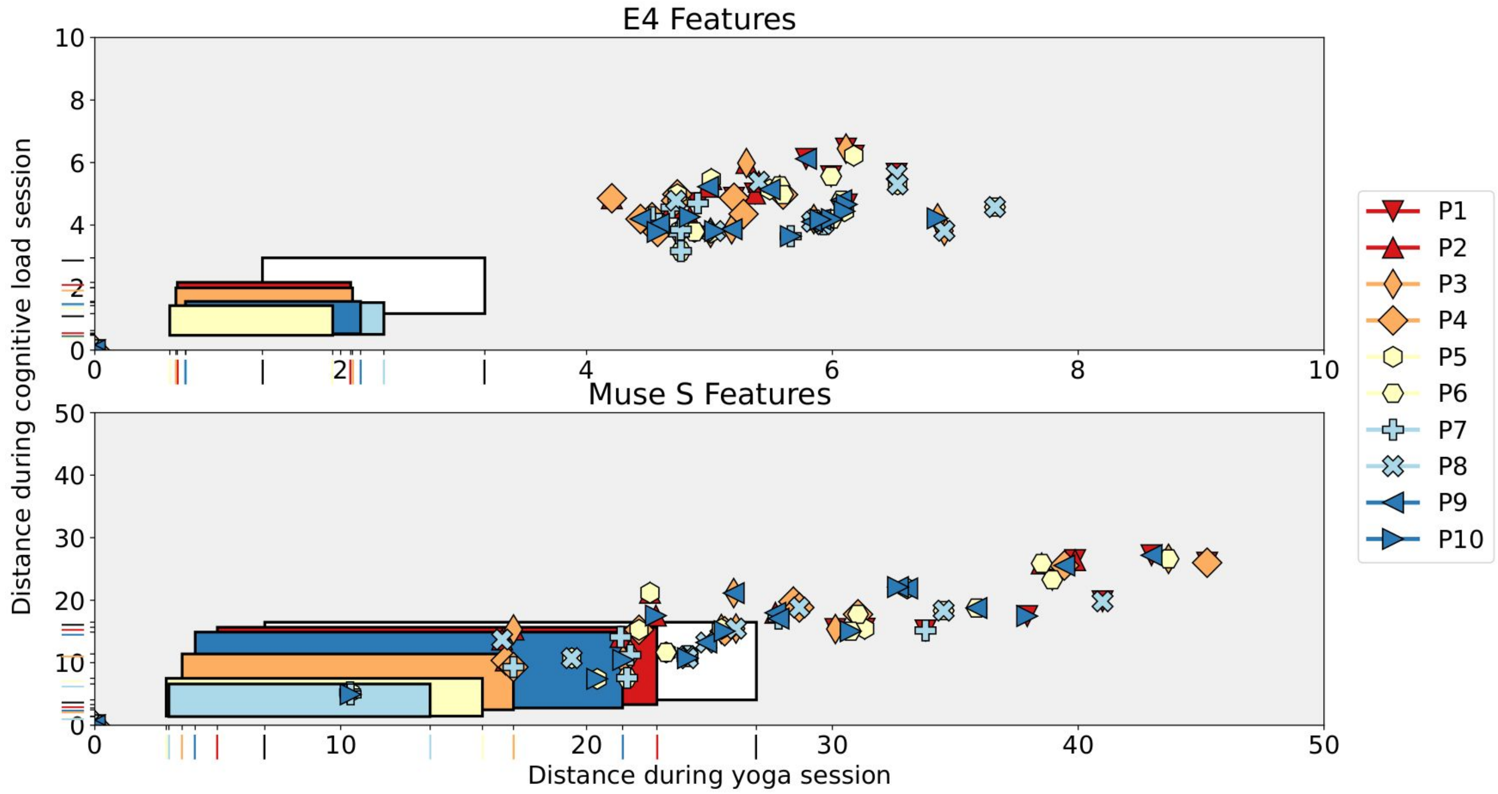
<u>Empatica Features:</u>

- Kurtosis, Skewness, Entropy, Min, Mean, and Max for Acceleration and BVP data
- Skin conductance level and Skin conductance response value for GSR
- Heart rate, and heart rate variability for PPG
- …

<u>All Features:</u>

All Muse and Empatica Features

Similarity of Normalized Dynamic Time Warping Distances



Machine Learning Hyperparameters (https://scikit-learn.org/stable/)

- Logistic Regression (LR)
  penalty (l1, l2, **None**)
  solver (**lbfgs**, liblinear, sag, saga)
- Decision Trees (DT)
  criterion (gini, **entropy**)
  splitter (**best**, random)
  max_depth (5, ..., 300, None); **145**
- Nearest Neighbors (NN)
  leaf_size (1, 2, ..., 50); **25**
  n_neighbors (1, 2, ..., 30); **21**
  p (**1**, 2)

- Support Vector Machine (SVM; Linear Support Vector Classifier)
  penalty (**l1**, l2)
  C (0.01, 0.1, 1,10, 100, **1000**)

- Due to class imbalance, one random resampling before all experiments; from 41:59 to 50:50

- Train-Validate-Test Split 60-20-20

- 5-Fold Nested Cross-Validation with inner HalvingGridSearchCV

Time Series Classification Performance

Model	Feature-Set	Generalized	Personalized
NN	All	88.80%	**90.01%**
NN	Muse	84.28%	**86.59%**
NN	E4	72.64%	**79.68%**
LR	All	80.12%	**82.13%**
LR	Muse	68.94%	**80.77%**
LR	E4	73.36%	**73.81%**
SVM	All	79.73%	**81.33%**
SVM	Muse	68.40%	**80.45%**
SVM	E4	73.32%	**73.98%**
DT	All	78.06%	**78.60%**
DT	Muse	71.62%	**79.51%**
DT	E4	68.42%	**75.75%**

Feature Importance

- Investigated using a correlation analysis performed after artefact removal

- Artefact removal: the dynamic Interquartile Range (IQR) method built on the standard deviation of each feature was utilized; Details are provided in the source code at https://github.com/siddhant61/ProSense

- Moderately correlated with the study phase (i.e., cognitive load induction or stress reduction):
  - Muse S data: 0.58 with p-value under 0.001: the statistical feature standard deviation within samples across all electrodes; 0.50 with p-value under 0.001: the power values (e.g., Alpha power) at AF7 only
  - Empatica E4 data: 0.51 with p-value under 0.001: the heart rate variability

Limitations

- Ongoing development ⇒ 3 out of 25 recordings have different artefacts
- Uncontrolled environments: bluetooth channel saturation led to 3 recordings showing a significant amount of artefacts
- Due to the change from one yoga asana to another pose, the second half of all recordings is partially distorted due to strong movement artefacts
- Data labelling during yoga was impractical. Consequently, the temporal resolution of self-assessed labels during the relaxation task is low
- ERP analysis worked only in few cases due to technical issues (TV screen refresh rate), and the absence of N200 during control tasks does not sufficiently support any conclusion on the capability of the developed pipeline to synchronously record data (the shaking-protocol did, however)
- The recordings were performed in winter, and some participants reported feeling a bit sick, consequently potentially affecting the comparability of temperature and GSR readings across participants and sessions

Future Work

- Changing the load induction and the stress reduction tasks for other tasks
- Subsequent data collection on individuals rather than small groups, extend the population size, and compare the findings on biomarkers
- Extending the analysis of the synchronicity of physiological responses before, during, and after the stress reduction intervention (e.g., yoga)
- Use a combination of participant-given, pre-defined, and objective labels
- Develop recommendations for organizations’ policies on stress management through interdisciplinary work, including a cost-benefit breakdown for a) individuals, b) organizations, and c) societies

Conclusion

- Physiological data of five groups of two participants were recorded, following a five-appointment study design
- Biomarkers of cognitive demands, such as the Alpha Power and the Heart Rate Variability can be used in combination with machine learning to timely and accurately classify individuals’ mental states
- Resulting (personalized) classifiers be integrated into eHealth platforms and offer monitoring or tailored advice on interventions based on the individuals’ stress patterns
- Dataset available upon request from christoph.anders@hpi.de

References

[1] D. Gopher and E. Donchin, “Workload: An examination of the concept,” in Handbook of perception and human performance, Vol. 2: Cognitive processes and performance. John Wiley & Sons, pp. 1–49, retrieved: 4, 2024.

Available: https://psycnet.apa.org/record/1986-98619-019

[2] S. G. Hart and L. E. Staveland, “Development of NASA-TLX(task load index): Results of empirical and theoretical research,”in Advances in Psychology, ser. Human Mental Workload. North-Holland, vol. 52, pp. 139–183, retrieved: 4, 2024.

Available: https://www.sciencedirect.com/science/article/pii/S0166411508623869

[3] S. Delliaux, A. Delaforge, J.-C. Deharo, and G. Chaumet, “Mentalworkload alters heart rate variability, lowering non-linear dynamics,”vol. 10, retrieved: 4, 2024.

Available: https://www.frontiersin.org/article/10.3389/fphys.2019.00565

[6] R. L. Charles and J. Nixon, “Measuring mental workload usingphysiological measures: A systematic review,” vol. 74, pp. 221–232,retrieved: 4, 2024.

Available: https://www.sciencedirect.com/science/article/pii/S0003687018303430

[8] A. Tonacci et al., “Can machine learning predict stress reductionbased on wearable sensors’ data following relaxation at workplace?a pilot study,” vol. 8, no. 4, p. 448, number: 4

Available: https://www.mdpi.com/2227-9717/8/4/448

References (ctd.)

[9] Y. Pan, X. Cheng, and Y. Hu, “Three heads are better than one:cooperative learning brains wire together when a consensus is reached,”vol. 33, no. 4, pp. 1155–1169, retrieved: 4, 2024.

Available:https://doi.org/10.1093/cercor/bhac127

[10] A. Kuhlen, C. Allefeld, and J.-D. Haynes, “Content-specific coordinationof listeners’ to speakers’ EEG during communication,” vol. 6, retrieved:4, 2024.

Available: https://www.frontiersin.org/articles/10.3389/fnhum.2012.00266

[11] S. Dikker et al., “Brain-to-brain synchrony tracks real-world dynamicgroup interactions in the classroom,” vol. 27, no. 9, pp. 1375–1380,retrieved: 4, 2024.

Available: https://linkinghub.elsevier.com/retrieve/pii/S0960982217304116

[12] D. Bevilacqua et al., “Brain-to-brain synchrony and learning outcomesvary by student–teacher dynamics: Evidence from a real-worldclassroom electroencephalography study,” vol. 31, no. 3, pp. 401–411,retrieved: 4, 2024.

Available: https://doi.org/10.1162/jocn_a_01274

[13] O. E. Krigolson et al., “Using muse: Rapid mobile assessmentof brain performance,” vol. 15, retrieved: 4, 2024.

Available: https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2021.634147

References (ctd.)

[15] E. Largo-Wight, B. K. O'Hara, and W. W. Chen, "The efficacy of a brief nature sound intervention on muscle tension, pulse rate, and self-reported stress: Nature contact micro-break in an office or waiting room," vol. 10, no. 1, pp. 45–51, retrieved: 4, 2024.

Available: https://doi.org/10.1177/1937586715619741

[16] C. Finkelstein, A. Brownstein, C. Scott, and Y.-L. Lan, "Anxiety and stress reduction in medical education: an intervention," vol. 41, no. 3, pp. 258–264, retrieved: 4, 2024.

Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2929.2007.02685.x

[17] S. Sallon, D. Katz-Eisner, H. Yaffe, and T. Bdolah-Abram, "Caring for the caregivers: Results of an extended, five-component stress-reduction intervention for hospital staff," vol. 43, no. 1, pp. 47–60, retrieved: 4, 2024.

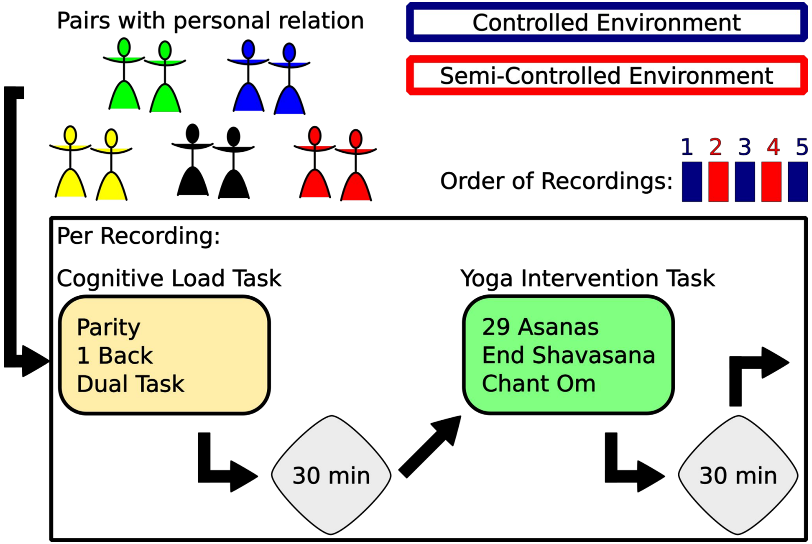
Available: https://doi.org/10.1080/08964289.2015.1053426

[23] D. Hagemann and E. Naumann, "The effects of ocular artifacts on (lateralized) broadband power in the EEG," vol. 112, no. 2, pp. 215–231, retrieved: 4, 2024.

Available: https://doi.org/10.1016/s1388-2457(00)00541-1

[24] S. L. Kappel, D. Looney, D. P. Mandic, and P. Kidmose, "Physiological artifacts in scalp EEG and ear-EEG," vol. 16, no. 1, p. 103, retrieved: 4, 2024.

Available: https://doi.org/10.1186/s12938-017-0391-2



Data Collection Pipeline

Data Processing Pipeline

Website

Load Induction then Simultaneous Relaxation: Insights from Multi-Modal Time-Series Data measured with Low-Cost Wearable Sensors

Christoph Anders[1], Sai Siddhant Gadamsetti[1], Nico Steckhan[1], Bert Arnrich[1]

[1]Digital Health - Connected Healthcare, University of Potsdam, Hasso Plattner Institute

Contact email: christoph.anders@hpi.de