MSR198

Use of Wearable-Generated Real-World Data to Objectively Identify Occupational Stressors

Ekaterina Mut[1], Jakob Hohn[2], Jenny Voigt[2], Celine Schreiber[3], Sophia Mareike Geisler[4], Pauline Sophia Pinta[4], Lorena Gleisner[4], Juliette-Michelle Burkhardt[1], Hamlet Kosakyan[6], Christian Hrach[1], Bogdan Franczyk[1,5], Ulf-Dietrich Braumann[1], Carsta Militzer-Horstmann[3,4]

Affiliations
[1] Institute for Applied Informatics (InfAI e.V.), Leipzig, Germany; [2] 4K ANALYTICS GmbH, Leipzig, Germany; [3] University of Leipzig, Health Economics and Management, Leipzig, Germany; [4] Scientific Institute for Health Economics and Health System Research (WIG2 GmbH), Leipzig, Germany; [5] University of Leipzig, Information Systems Institute, Leipzig, Germany ; [6] Appsfactory GmbH, Leipzig, Germany

Background

Understanding and enhancing job satisfaction, which leads to improved **work-related quality of life (WrQoL)**, is essential in today's world, where work-life boundaries are increasingly blurred. High **WrQoL** promotes a healthier, more productive workplace, benefiting both employees and organizations.

Going beyond traditional qualitative methods, our research focuses on **objective measurement** of WrQoL using **machine learning techniques.** This new and innovative approach ensures reliable, high-quality data and mitigates respondent bias. In a first step, a systematic literature search identified WrQoL indicators, with central indicators being integrated into a base model and linked to data sources (Figure 1). **Heart rate variability (HRV)**, an indicator of the body's ability to adapt heart rate to physical and mental demands, has been identified as a key predictor of physiological and emotional states that influence WrQoL. By utilizing machine learning models trained on aggregated multimodal data, including **wearable and human computer interaction (HCI)-generated real-world data**, we aim to improve the predictive accuracy of individual WrQoL scores, with the potential to enhance workplace conditions and employee well-being.

Objectives

This study aims to identify a meaningful combination of HRV-related indices (HRVIs) to assess physiological responses to occupational stressors, such as high workload. Using AI-based approaches, we seek to predict early signs of emotional distress or burnout.

Methods

Sensor selection
- Utilization of Polar optical sensors (OH1, Verity Sense, 360) for measuring HRV through photoplethysmographic (PPG) signals providing pulse-to-pulse intervals (PPI) due to their high accuracy and the ability to provide continuous live data streaming (Table S1).

PPI data acquisition
- Connection to iOS platform: utilizing the Polar SDK for continuous PPI acquisition.
- Timestamp for each PPI measurement by MI-LQ app and local storage on the mobile phone with the following structure:
  - Unix timestamp
  - PPI value
  - PPI error estimate
  - Skin contact status
  - Measurement validity status

HRV indices
- Calculation of seven relevant HRV indices based on 5-minute intervals (Table 1).
- *k*-means clustering: Partitioning of various HRV indices into a pre-defined number of clusters to describe individual physiological states (Figure 2).

Collection of work-related data
- HCI data: Continuous assessment of mouse and keyboard dynamics with specific software.
- Valence and arousal assessment: Regular non-verbal valence and arousal[1] (Figure S1) assessment using the Self-Assessment Manikins (SAM)[2] (Figure S2) via a developed self-assessment software to annotate HCI and HRV data.

Data analysis
- HRV cluster distribution during working hours, days and weeks to detect particularly stressful periods at work.
- Investigating correlations between HCI, valence and arousal data, and HRV indices.

Machine Learning
- Individual trained Long short-term memory (LSTM) models: Deep learning sequential networks that can learn hidden patterns in temporal sequences to retain information from previous time points.
- Analysing temporal changes in HRV and HCI data, trained on 5-minute data blocks ⟶ prediction of hourly valence and arousal levels.

Pilot study (planned)
- Study design: Cross-institutional study with 30 office employees over 4 weeks.
- Baseline data: Collecting individual and professional information to contextualize HRV data.

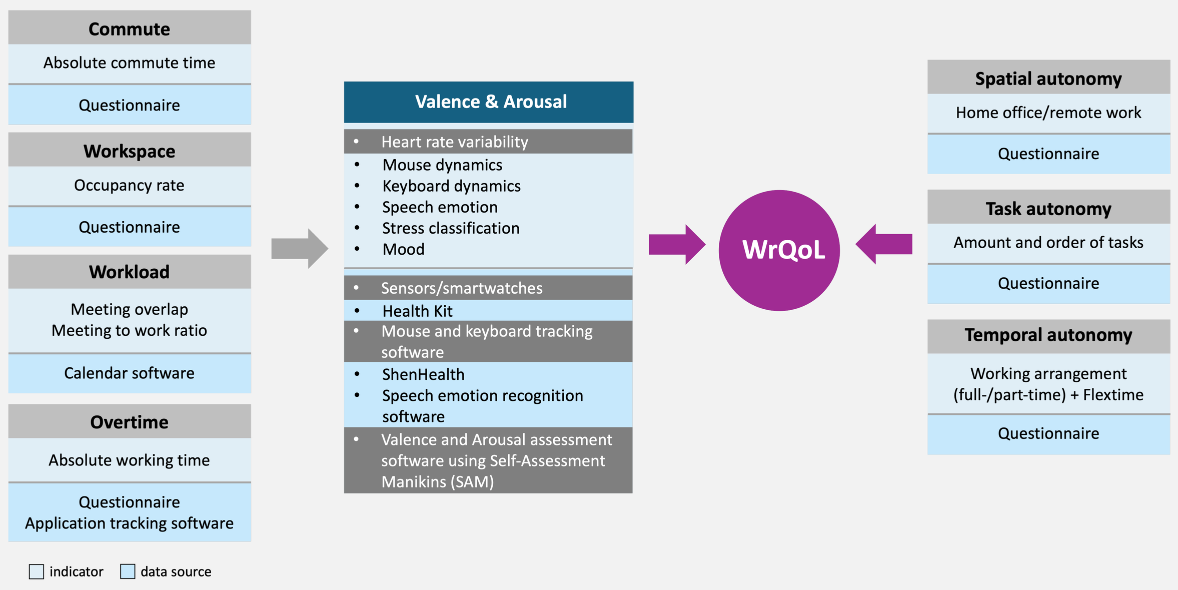

Figure 1. Base model of WrQoL indicators and their potential for objective measurement with various digital approaches. The poster focuses on the dark grey highlighted indicators and data sources.

Table 1. Overview and definition of HRV indices according to Baevsky and Chernikova (2017)[3]

HRV indices	Definition
Mean HR	Mean heart rate, measured in beats per minute (bpm)
Mean NN interval (NNI)	Mean of PPG-based PPIs
Total Power	Overall variance of HRV across all frequency domains, indicating autonomic nervous system activity
RMSSD	Root Mean Square of Successive Differences of PPIs
Stress Index (SI)	Measure of the amplitude of the modal value in the PPI histogram, measures mental stress and sympathetic influence, assessing the autonomic nervous system's regulatory capacity
IVR (Index of Vegetative Regime)	Ratio of sympathetic to parasympathetic influences on heart rhythm
PAPR (Parameter of Appropriateness of Heart Regulation)	Ratio of the amplitude of the modal value in the PPI histogram to respective modal value

Results

1. Identification of physiological states
- *k*-means clustering was used with Baevsky stress index (SI)[3] and mean heart rate (mean_hr) as key variables (Figure 2).
- Later working hours often showed relatively lower heart rates with higher SI, possibly indicating fatigue. The other states are present in varying proportions during regular working hours (Figure 3), making it possible to identify a dominant or typical state for each hour.
- The strain of a person is affected by the individual's activity at work, as indicated by different SI means comparing periods of HCI and periods of non-HCI (Figure 4). In addition, the SI rates and fluctuations during HCI and non-HCI activities differ between individuals, suggesting that personal characteristics such as temperament and individual strategies for coping with stress have a large influence on SI values. This indicates the importance of using personalized models with distinct baselines for different individuals to accurately capture their physiological states and responses to workload and activity.
- Particularly stressful days, characterized by significant fluctuations in physiological states during working hours, were identified by comparing the daily mean rates against a personalized baseline (Figure 5). Analysing data over entire weeks can reveal trends and patterns in these fluctuations.
- Certain software programs, such as web browser and spreadsheet (Excel), were associated with higher SI, suggesting increased mental effort, concentration, and possible excitement during use (Figure S3). In contrast, increased RMSSD values suggest that using specific programs stimulates parasympathetic activity, facilitating recovery and relaxation (Figure S4).

2. Comparison of physiological states with emotional self-evaluation
- Several physiological states were identified that reflect the influence of workload and stress on HRV. Subjective measures of valence and arousal[1] (Figure S1) showed low correlation with the HRV indices (Figures S5) based on clusters identified in Figure 2.
- Higher SI levels do not necessarily indicate a worsened mood. They can reflect stress, positive excitement (eustress), mental effort or tension. This is why the content of the work is important—whether tasks are completed, and goals are achieved can influence how the person feels. Someone may feel physically tired yet still be happy and fulfilled.

3. Mouse and keyboard patterns
- Analysis of mouse and keyboard behaviours did not reveal clear correlations with physiological states.
- No typical behaviour patterns could be linked to workload or HRV variations.

4. Machine learning approaches
- Random Forest Classifier predicting valence across two classes—negative and positive mood—based on HRV data: Accuracy 73%, AUC 78%, F1-score 74%.
- However, the introduction of a five-class system adversely affected prediction accuracy, highlighting the challenges associated with increased complexity.
- Key features influencing valence prediction: i) hour of the day with feature importance of 9% and ii) week (4.6%), mean NNI (4.4%), mean HR (3.8%), and max HR (3.6%) of less feature importance.

5. Evaluation of WrQoL
- Clusters (Figure 2) can be used as sub-labels for WrQoL.
- Combining HRV data with information on job satisfaction and task variety is crucial for assessment of WrQoL.

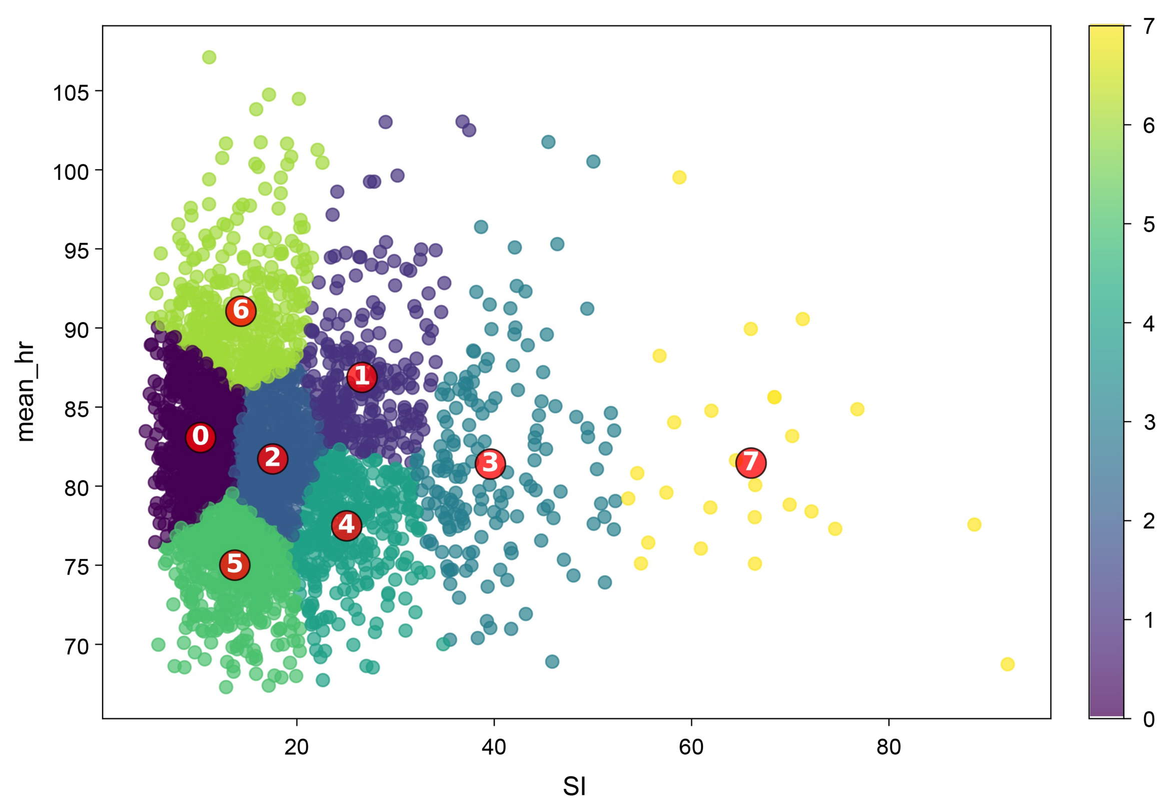

Figure 2. *k*-means clustering based on SI and mean_hr identifies different physiological states during working hours and human computer interaction. (n=1)

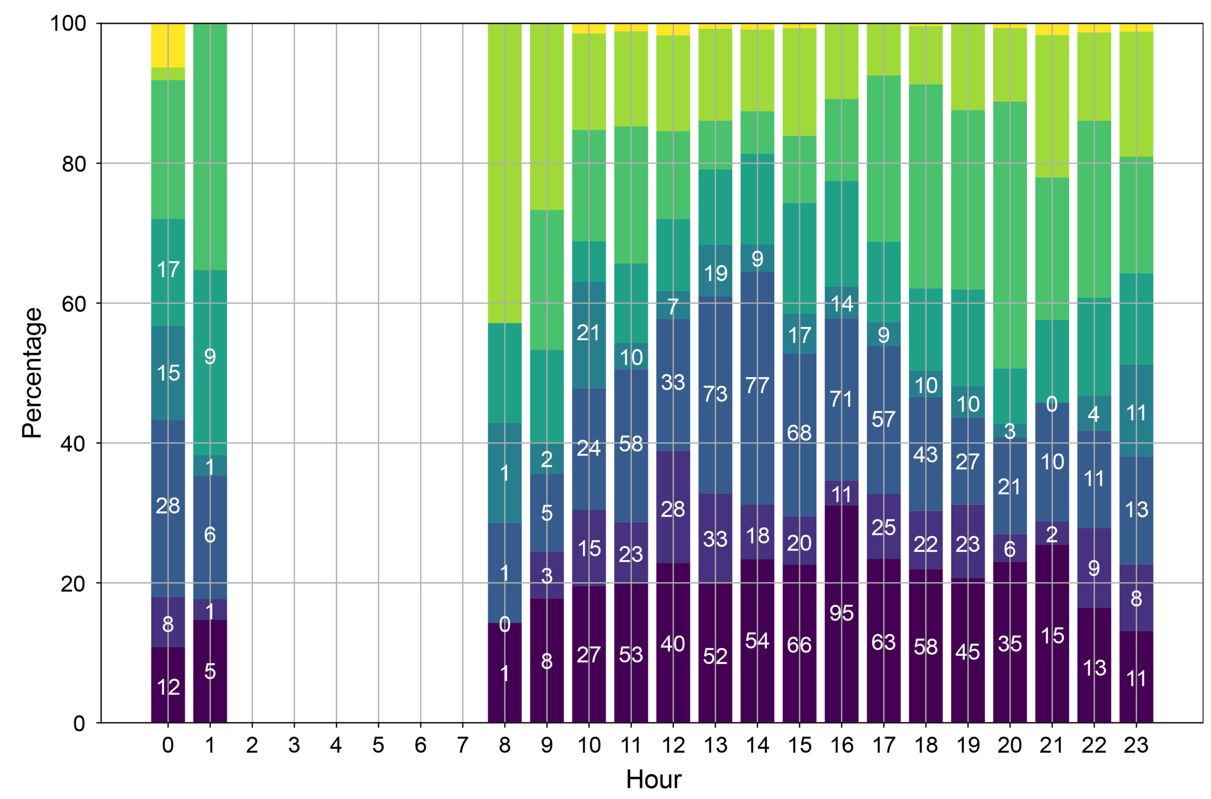

Figure 3. Percentage distribution of different physiological states identified using *k*-means clustering approach across working hours over the entire observation period of 6 months. (n=1)

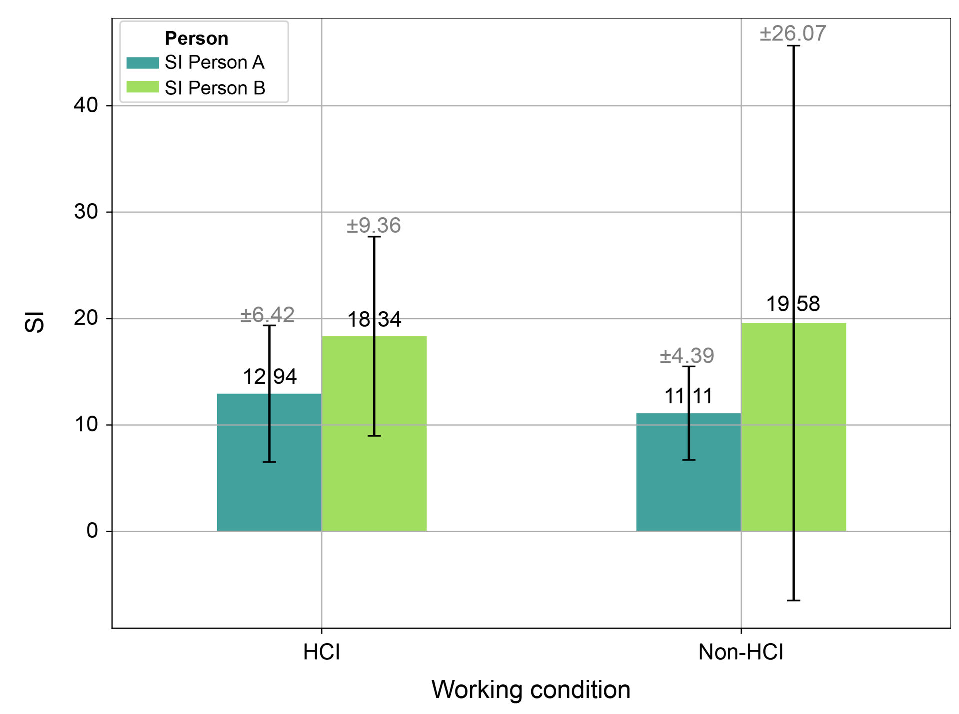

Figure 4. Comparison of SI rates and fluctuations among two individuals of varying temperaments and ages during HCI and non-HCI-based activities within working hours. Greater fluctuations in SI rates are observed across both types of activities for person B. These findings highlight the importance of using personalized models with distinct baselines for different individuals to accurately capture their physiological states and responses. (n=2)

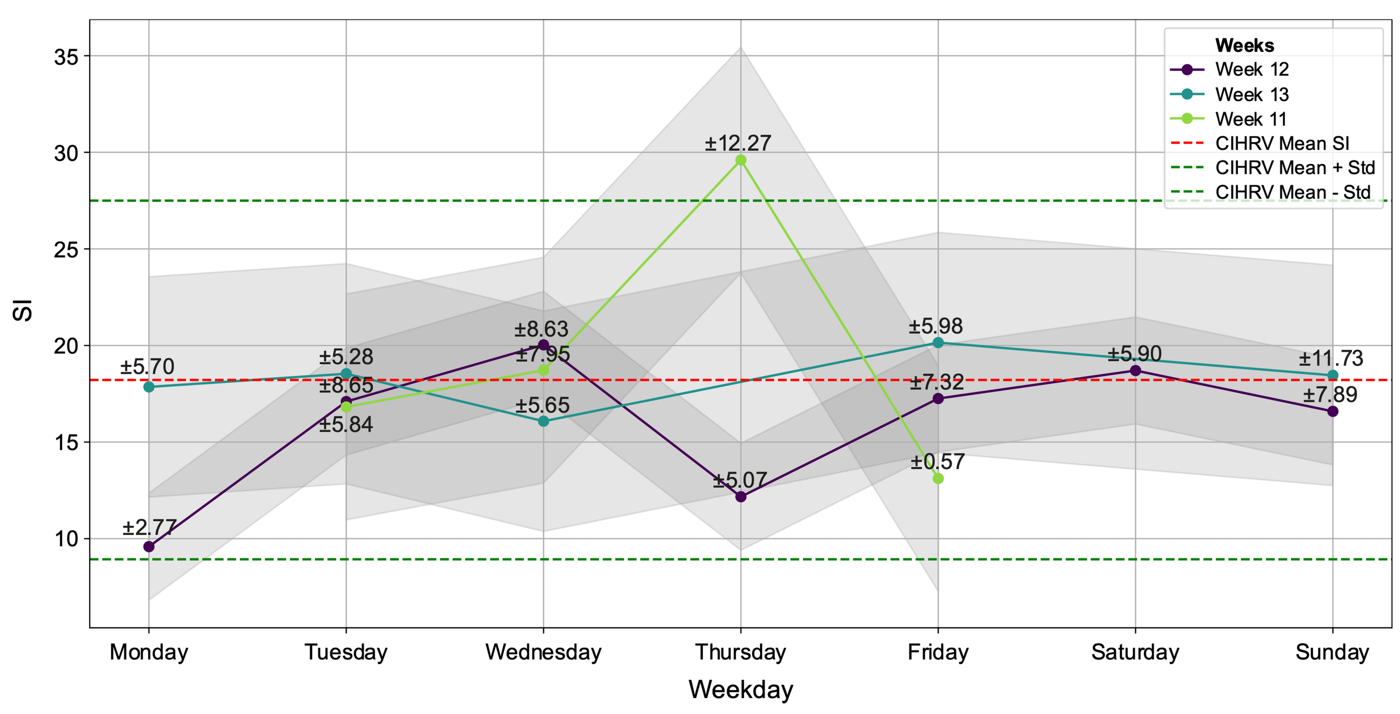

Figure 5. Visualization of the mean day SI values per week (solid line) for 3 weeks in comparison to the SI baseline (dashed lines), based on the overall mean and standard deviation of SI of 6 months. Comparison to the baseline allows identification of particularly stressful days (n=1)

Conclusions

1. **Different physiological states** could be identified during working hours through clustering of HRV indices, which are influenced by workload.
2. There was **no strong correlation between cluster indices** and subjective measures of **valence and arousal**.
3. Emotional states showed **time- and activity-dependent variations**, enabling the detection of the most stressful working periods.
4. Mouse and keyboard dynamics did not correlate with physiological states.
5. Cluster indices are potential sub-labels for WrQoL
6. Personalised models should be preferred used for machine learning approaches.

Our research is a first attempt to objectively assess physiological and emotional states that affect WrQoL, among other indicators, by integrating data from analyses of HRV, mouse and keyboard interaction, and valence and arousal. However, more contextual information on activity and workload, as well as data from more individuals, is needed to understand why physiological states differ and the extent to which individuals differ in their ranges.

Outlook

- Combination of data sources and roll-out of daily questionnaires to gain labeled context information on activity and workload.
- Automize data transfer and scoring processes.
- Expansion to Android as second operating system.
- Transfer to other work settings.
- Pilot study with approximately 30 office workers to generate a training dataset for the machine learning model.

Limitations

- **Brand-specific devices:** The reliance on specific manufacturers and devices (e.g., Polar sensors and Apple (iOS)) limits the generalizability of results and restricts integration with other fitness trackers, potentially reducing user accessibility.
- **Availability of HRV data:** Many common wearables only provide heart rate data and not PPI data needed to calculate HRV indices.
- **Data quality:** HRV measurement accuracy can be affected by motion artifacts and poor skin contact. A lack of standardization among manufacturers may lead to unreliable data. Heart rate data derived from echocardiogram provides better quality than PPG-derived heart rate data, which is 15% less accurate.
- **Transferability:** The focus on the office environment limits the applicability of findings to other work contexts, such as construction sites or laboratories.
- **Sample representativeness:** Limited diversity in the sample may restrict generalizability, as different occupational groups may show varying physiological and emotional responses to stress.

Abbreviations: HCI: human computer interaction, HR: heart rate, HRV: heart rate variability, HRVI: heart rate variability indices, IVR: index of vegetative regime, LSTM: long short-term memory, NNI: NN-interval, PPG: photoplethysmography, PAPR: parameter of appropriateness of heart regulation, PPI: pulse-to-pulse interval, RMSSD: root mean square of successive differences, SAM: self-assessment manikins, SI: stress index, Std: standard deviation, WrQoL: work-related quality of life

References: Please, refer to the supplemental material displayed on the digital poster (digital handout).

Scan me, to get more information about the project

Co-funded by the European Union

This project is co-financed from tax revenues on the basis of the budget adopted by the Saxon State Parliament.