

Continuous measurement of stress levels in naturalistic settings using heart rate variability: An experience-sampling study driving a machine learning approach

Pietro Cipresso^{1,2}, Silvia Serino³, Francesca Borghesi^{1,2}, Gennaro Tartarisco⁴, Giuseppe Riva^{2,3}, Giovanni Pioggia⁴, Andrea Gaggioli^{2,3}

¹ Department of Psychology, University of Turin, Turin, Italy

² Applied Technology for Neuro-Psychology Lab, Istituto Auxologico Italiano, Milan, Italy

³ Università Cattolica del Sacro Cuore, Milan, Italy

⁴ Institute for Biomedical Research and Innovation (IRIB), National Research Council of Italy (CNR), Messina, Italy

ABSTRACT

Developing automatic methods to measure psychological stress in everyday life has become an important research challenge. Here, we describe the design and implementation of a personalized mobile system for the detection of psychological stress episodes based on Heart-Rate Variability (HRV) indices. The system's architecture consists of three main modules: a mobile acquisition module; an analysis-decision module; and a visualization-reporting module. Once the stress level is calculated by the mobile system, the visualization-reporting module of the mobile application displays the current stress level of the user. We carried out an experience-sampling study, involving 15 participants, monitored longitudinally, for a total of 561 ECG analyzed, to select the HRV features which best correlate with self-reported stress levels. Drawing on these results, a personalized classification system is able to automatically detect stress events from those HRV features, after a training phase in which the system learns from the subjective responses given by the user. Finally, the performance of the classification task was evaluated on the empirical dataset using the leave one out cross-validation process. Preliminary findings suggest that incorporating self-reported psychological data in the system's knowledge base allows for a more accurate and personalized definition of the stress response measured by HRV indices.

Section: RESEARCH PAPER

Keywords: Psychological stress; psychophysiology; psychometrics; signal processing; assessment; experience sampling methods; heart rate variability

Citation: Pietro Cipresso, Silvia Serino, Francesca Borghesi, Gennaro Tartarisco, Giuseppe Riva, Giovanni Pioggia, Andrea Gaggioli, Continuous measurement of stress levels in naturalistic settings using heart rate variability: An experience-sampling study driving a machine learning approach, Acta IMEKO, vol. 10, no. 4, article 36, December 2021, identifier: IMEKO-ACTA-10 (2021)-04-36

Section Editor: Carlo Carobbi, University of Florence, Gian Marco Revel, Università Politecnica delle Marche and Nicola Giaquinto, Politecnico di Bari, Italy

Received October 11, 2021; **In final form** December 20, 2021; **Published** December 2021

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the European funded project 'Interstress-Interreality in the management and treatment of stress-related disorders', Grant Number: FP7-247685.

Corresponding author: Pietro Cipresso, e-mail: p.cipresso@auxologico.it

1. INTRODUCTION

It is well known that long-term exposure to stress can lead to immunodepression and dysregulation of the immune response, thus significantly enhancing the risk of contracting a disease or altering its course. However, increased symptomatology is not only associated with severe stressors (infrequent major life events), but also with minor daily stressors (i.e. "micro-stressors") that are ignored or poorly managed [1]–[4]. Defining effective techniques to measure daily stressful episodes in

ecological conditions has thus been identified as an important research objective. To address this challenge, several research groups have started investigating the use of wearable sensors solutions to infer stress from continuous biosignal measurements [5] (for a review, see [6]). Such systems integrate sensors together with on-body signal conditioning and pre-elaboration, as well as the management of the energy consumption and wireless communication systems. Although preliminary testing of these systems has yielded encouraging results [7], a major limitation of current solutions is that they mostly rely on complex sensor architectures and use labelling

methods that are often based on evaluation of human coders [8]. Other authors have proposed the measurement of Heart Rate Variability (HRV) [9]–[14] analysis as a potentially effective approach for monitoring stress in mobile settings [15]–[17]. Actually, HRV indices can be used to estimate activity of autonomous nervous system (ANS) in relation to affective and cognitive states, including mental stress [18]–[20]. However, real-time recognition of stress from HRV measures requires appropriate strategies to i) detect HRV changes using minimally-invasive ECG equipment; ii) relate these changes to mental stress levels; and iii) control the potential confounding effects of physical activity. In the following, we describe how we addressed these issues in designing and implementing a personalized mobile system for automatic recognition of psychological stress based on HRV indices. The original contribution of the proposed method is that, to our best knowledge, this is the first approach that integrates the detection of HRV features with the ground-truth of subjective perception of stressful events.

2. MEASURING PSYCHOLOGICAL STRESS IN NATURALISTIC ENVIRONMENTS

According to Cohen et al. [21], stress is a biopsychosocial phenomenon in which “environmental demands tax or exceed the adaptive capacity of an organism, resulting in psychological and biological changes that may place a person at risk for disease” (p. 3). This conceptualization suggests that in measuring stress, it is not only necessary to consider environmental demands, but also appraisals of such demands, as well as physiological systems that come into play. Consistent with this definition, two main approaches have been introduced to assess psychological stress in naturalistic conditions: the first is based on self-reporting of participants’ subjective experiences and perception of stressful events; the second approach is based on sensing physiological signals associated with the stress response. In the next, we provide a description of these procedures, along with a discussion of their strengths and limitations.

2.1 Subjective psychological measures

The experience sampling method (ESM), also known as ecological momentary assessment (EMA), is a naturalistic observation technique that allows capturing participants’ thoughts, feelings, and behaviours at multiple times across a range of situations as they occur in the natural environment [22]. In a typical ESM study, participants are asked to fill out a form when prompted by an acoustic signal. Thanks to repeated sampling, a number of surveys are collected from each participant throughout the day, thus providing an ecologically-valid and highly detailed description of subjective quality of experience. Ecological validity is a strong requirement in psychometrics since this is a measure of how a task or test is able to predict behaviours in a real-world setting. ESM has been applied to study a wide range of behaviours and experiences, including daily stress [23], [24]. However, this procedure has high costs and places a significant burden on the participant, thus limiting its practical applicability as a stress monitoring technique [10], [25].

A less expensive and time-consuming approach for assessing experience and affect in everyday life is the day reconstruction method (DRM), developed by Kahneman and colleagues [26]. It involves the retrospective recall of the study period as a continuous sequence of episodes, which are rated on a series of affect scales. DRM reports have been validated against

experience sampling data, showing that this technique allows identifying changes in affect over the course of the day with almost the same accuracy than ESM [26], [27]. However, since DRM respondents are asked to reconstruct the previous day by completing a structured self-administered questionnaire, this method is potentially susceptible to recall biases. Furthermore, it has been suggested that using retrospective measures as proxies for actual experience may result in weaker or inconsistent results particularly when tested in connection to biologic pathways [28].

2.2 Objective physiological measures

An alternative strategy to assess stress in everyday situations is based on the analysis of physiological correlates of this experience. Psychological stressors are linked with the activation of two main neuro-physiological pathways, which are involved in the maintenance of homeostasis: the hypothalamic-pituitary-adrenocortical (HPA) axis and the sympathetic-adrenal medullary system (SAM). As concerns the first system, one of the most investigated biomarker is salivary cortisol, which, together with catecholamines, is one of the end products of HPA activation [20], [23], [29], [30]. However, due to significant between- and within individual variation in diurnal secretion of cortisol, the measurement of the magnitude of cortisol response is not an easy procedure, which requires the application of advanced statistical approaches such as multilevel models [31], [32]. With respect to the SAM system, HRV, defined as the variation over time of the period between consecutive heartbeats, is increasingly regarded as a potentially convenient and non-invasive marker of autonomic activation associated with psychological stressors [33]. The normal variability in heart rate (HR) is controlled by the balancing activation of the (acceleratory) sympathetic and of the (deceleratory) parasympathetic branches of the autonomic nervous system. However, under stressful events or contexts, there is a trend towards increased sympathetic control and reduced vagal tone, which is associated with decreased HRV [19]. On the other hand, higher HRV has been associated with the availability of context- and goal-based control of emotions [34]. Based on this preliminary evidence, several authors have been experimenting with wearable heart monitor for the identification of stress levels from HRV, in both healthy and clinical populations. For example, Kim and coll. [35] used HRV patterns to discriminate between subjects reporting high and low levels of stress during the day, with an overall accuracy of 66.1 %. In a similar study, Melillo et al. [16] compared within-subject variations of short-term HRV measures using short term ECG recording in students undergoing university examination. By applying Linear Discriminant Analysis on nonlinear features of HRV for automatic stress detection, these authors were able to obtain a total classification accuracy of 90 %. Kimhy et al. [17] investigated the relationship between stress and cardiac autonomic regulation in a sample of psychotic patients, using experience sampling in combination with cardiac monitoring. They found that momentary increases of stress were significantly associated with increase in sympathovagal balance and parasympathetic withdrawal. In addition to studies which have examined the association between HRV and stress during waking hours, other recent research has proposed the use of HRV patterns during sleep as supplement to the analysis of subjective assessments and voice messages collected during workday [36], with encouraging, albeit preliminary, results.

Table 1. Feature extraction from electrocardiogram (ECG).

Measure	Description
RR mean	Mean of all RR intervals
AVNN	Average of all NN intervals
SDNN	Standard deviation of all NN intervals
rMSSD	Square root of the mean of the squares of differences between adjacent NN intervals
NN50	Differences between adjacent NN intervals that are greater than 50 ms
TOTPWR	Total spectral power of all NN intervals up to 0.04 Hz
LF	LF Total spectral power of all NN intervals between 0.04 Hz and 0.15 Hz
HF	HF Total spectral power of all NN intervals between 0.15 Hz and 0.4 Hz
LFbyHF	LF/HF Ratio of low to high frequency power (Sympathovagal balance)

2.3 Towards an integrated approach for personalized stress recognition in mobile settings

As previously discussed, a fundamental issue in the measurement of stress in everyday life is that this response is idiosyncratic, because it depends on individual's perception of challenges and the skills which he/she can use to face those challenges. As a consequence, any approach aiming at inferring stress levels from "honest" physiological signals should not overlook the role played by the subjective appraisal of the situation. Furthermore, since HRV values are characterized by high interindividual variability, it is important that the system is tailored to the individual characteristics [37], [38]. One possible approach to develop adaptive systems for stress recognition has been suggested by Morris and Guilak [37]. The strategy proposed by these authors involves identifying subject's baseline and stress threshold in the lab by elicitation of sympathetic and parasympathetic responses, and then using this information to differentiate between stress and nonstress in daily life. A first attempt to implement this approach has been performed by Cinaz et al. [38]. These authors measured participants' sympathetic and parasympathetic responses during three different levels of mental workload (low, medium, and high) in a controlled laboratory setting. Then, they investigated whether the data collected in this calibration session were appropriate to discriminate corresponding workload levels occurred during office work. To this end, individual HRV responses of each workload level were used to train the models and test the trained models on the data collected while the subjects performed normal office-work, using a mobile ECG logger. Afterward, a multiple regression analysis was applied to model the relationship between relevant HRV features and the subjective ratings of perceived workload: resulting predictions were correct for six out of the seven subjects. In the present work, we propose an experience-sampling approach for incorporating subjective knowledge in the classification of psychological stress from HRV indices (Table 1). The methodology consists of three main steps. In the first, experimental phase, we carried out an experience-sampling study to select the HRV features which best correlate with self-reported stress levels (Section 3). Drawing on these results, a personalized classification system was developed which is able to automatically detect stress events from those HRV features, after a training phase in which the system learns from the subjective responses given by the user (Section 4). In the final step, the performance of the classification task was evaluated using the leave one out cross validation process (Section 5).

3. METHOD

The objectives of this experiment were two-fold: i) to select a subset of HRV features which best correlate with self-reported stress levels collected during everyday activities; ii) to select a subset of self-reported questions about perceived stress levels which can be used as ground truth to train the final system.

3.1 Participants

Participants were 15 healthy subjects (8 males and 7 females, mean age = 23.33 years, St. Dev.= 1.49), monitored longitudinally, for a total of 561 ECG analyzed, to select the HRV features which best correlate with self-reported stress levels. Participants were recruited through opportunistic sampling. Participants filled a questionnaire assessing factors that, in the opinion of the investigators, might interfere with the measures being assessed (i.e., caffeine consumption, smoking, alcohol consumption, exercise, hours of sleep, disease states, and medications). Written informed consent was obtained by all subjects matching inclusion criteria (age between 18 and 65 years, generally healthy, absence of major medical conditions, and completion of informed consent).

3.2 Materials

Data were collected through PsychLog [39], a mobile experience sampling platform designed for research in mental health, which allows simultaneous collection of psychological, physiological (ECG) and motion activity data. Psychological data are collected from surveys that can be simply customized by the experimenter. For the purpose of this study, we used the Italian adaptation of the ESM questionnaire applied by Jacobs et al. [40] for studying the immediate effects of stressors on mood. The survey includes open-ended and closed-ended questions investigating thoughts, current context (activity, persons present, and location), appraisals of the current situation, and mood. All self-assessments were rated on 7-point Likert scales. HR and activity data are acquired from a wireless electrocardiogram (Shimmer Research™) equipped with a three-axial accelerometer. The wearable sensor platform includes a board that allows the transduction, amplification and pre-processing of raw sensor signals, and a Bluetooth transmitter to wirelessly send the processed data. The unit is mounted on a soft-textile chest strap designed to seamlessly adapt to the user's body shape, bringing full freedom of movement. Sensed data are transmitted to the mobile phone Bluetooth receiver and gathered by the PsychLog computing module, which stores and process the signals for the extraction of relevant features.

3.3 Design and procedure

Participants received a short briefing about the objective of the experiment and filled the informed consent. Then, they were provided with the mobile phone with pre-installed PsychLog application, the wearable ECG and accelerometer sensor and a user manual including experimental instructions. The application was pre-programmed to collect data over 7 consecutive days, at random intervals during waking hours. At the end of the experiment, participants returned both the phone and the sensors to the laboratory staff. After, participants were debriefed, thanked for their participation, and dismissed (Figure 1).

3.4 Data analysis

Following the procedure suggested by Jacobs et al. [40], three different psychological stress measures were computed in order to identify the stressful qualities of daily life experiences.

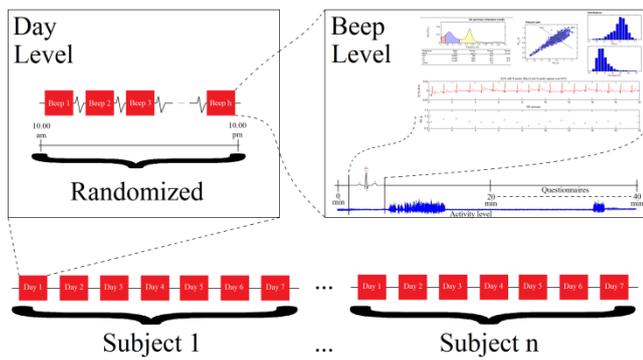


Figure 1. Schematic representation of experimental design.

Ongoing “Activity-Related Stress” (ARS) was defined as the mean score of the two items “I would rather be doing something else” and “this activity requires effort” (Cronbach’s alpha = 0.72). To evaluate social stress, participants rated the social context on two 7-point Likert scales “I don’t like the present company” and “I would rather be alone”; the “Social Stress Scale” (SS) resulted from the mean of these ratings (Cronbach’s alpha = 0.59). Finally, for “Event-Related Stress” (ERS), subjects reported the most important event that had happened since the previous beep, whether or not it was still ongoing. Subjects then rated this event on a 7-point bipolar scale (from 3 very unpleasant to 3 very pleasant, with 0 indicating a neutral event). All positive responses were recorded as 0, and the negative responses were recorded so that higher scores were associated with more unpleasant and potentially stressful events (0 neutral, 3 very unpleasant). In addition to those scales, an item (not included in the original survey by Jacobs et al. [40]) asked participants to rate the perceived level of stress on a 10-point Likert scale. This item was included as a global subjective measure of stress. Given the repeated sampling, Likert-type scales data were standardized (Mean = 0; St. Dev. = 1) on each participant’s weekly mean for every variable before performing the analyses. ESM data can be aggregated at the report level (the unit of analysis is the individual diary entry) or at the subject level (the unit of analysis is the participant). In the present study, most of the analyses were conducted using the subject-level aggregation, because this approach avoids problems related to unequal weights and produces more conservative significance tests [41].

3.5 Selection of psycho-physiological features

To analyse HRV features, the QRS peaks and RR interval time series recorded and saved on the PsychLog application were exported and further processed with the software Matlab (version 7.10) in order to compute a set of HRV indexes. To this end, the ECG signal was first elaborated for artifact correction, and then a fast Fourier transform was used to compute the power spectrum in the LF (0.04–0.15 Hz) and HF (0.15–0.40 Hz) bands [42]–[45]. To estimate the effect of HRV indexes (independent variable) on stress level (dependent variable) we applied hierarchical linear analysis, an alternative to multiple regression which is more suitable for our nested data. Actually, hierarchical structure of data makes traditional forms of analysis unsuitable, since within-subject data are collected at many points in time during each day, across several days. Moreover, traditional repeated-measures designs require the same number of observations for each subject and no missing data. Finally, hierarchical linear analysis allows to take into account further dependencies existing in the data.

4. RESULTS

Given the repeated sampling, Likert-type scales data were standardized (mean = 0; SD = 1) on each participant’s weekly mean for every variable before performing the analyses. ESM data can be aggregated at the report level (the unit of analysis is the individual diary entry) or at the subject level (the unit of analysis is the participant). In the present study, most of the analyses were conducted using the subject-level aggregation, because this approach avoids problems related to unequal weights and produces more conservative significance tests [31]. The following table provides the correlations between stress measures described before. As can be seen from Table 1, all scales measuring stress (STRESS, ARS, SS, and ERS) are significantly correlated between them. Out of 561 “beeps”, participants filled 541 reports (96 %), of which 456 were included in the analysis (84 %). A total of 561 ECG sampling were recorded (100 %), and 374 were included in the analysis (69 %). The following Table 2 provides the correlations between stress measures described before. In fact, as can be seen from Table 2, all scales measuring stress (STRESS, ARS, SS, and ERS) are significantly correlated between them. The hierarchical linear analysis, both aggregation levels (report-level and subject-level) were considered in the model. Results indicated a statistical significant hierarchical regression model for RMSSD (Beta -.5350813; St. Dev.: .2151596; $p < .013$), NN50 (Beta -1.152351; St. Dev.: .5322348; $p < .030$), LF / HF (Beta 1.176422; St. Dev.: .5386275; $p < .029$) (Table 3). The RMSSD method is preferred to NN50 because it has better statistical properties [42], [43]. Findings of this ESM experiment allowed to identify a subset of HRV features, which showed best correlations with self-reported psychological stress levels. In the next section, we describe how these psycho-physiological features were implemented in a

Table 2. Correlations among psychological self-reported measures.

		ZSTRESS	ZARS	ZSS	ERS
ZSTRESS	<i>r</i>	1	,312**	,215**	,213**
	<i>Sig.</i>		< .001	< .001	< .001
	<i>N</i>	540	534	528	456
ZARS	<i>r</i>	,312**	1	,393**	,146**
	<i>Sig.</i>	< .001		< .001	.002
	<i>N</i>	534	535	529	457
ZSS	<i>r</i>	,215**	,393**	1	,188**
	<i>Sig.</i>	< .001	< .001		< .001
	<i>N</i>	528	529	529	457

** Correlation is significant at the 0.01 level (2-tailed).

Table 3. Summary of hierarchical regression analysis for HRV variables predicting global perceived stress (Number of observations = 374).

Global Perceived Stress	B	SE B	Z	$p > Z $	95 % CI	
HR	0.51	0.22	2.37	0.02	0.09	0.94
RMSSD	-0.53	0.21	-2.49	0.01	-0.96	-0.11
NN50	-1.15	0.53	-2.17	0.03	-2.20	-0.11
LF	0.62	.031	2.02	0.04	0.02	1.23
LF/HF	1.18	0.54	2.18	0.03	0.12	2.23

personalized stress monitoring system, which was designed to learn from individual's subjective assessments of stressful situations and use this knowledge to detect stress events. The personalized stress monitoring system includes three main components: a) a mobile acquisition/feedback module (for the collection of psycho-physiological and activity data); b) a remote analysis-decision module (for the analysis and classification of stress levels), c) a mobile visualization-reporting module (for the reporting of detected stress events).

4.1 Mobile acquisition/feedback module

The mobile acquisition/feedback component consists of two elements: a wireless electronic module coupled with a commercial chest band for collecting ECG and motion data; and a smartphone application for the collection of psychological data, the transmission of data to the analysis-decision module and the visualization of stress events detected by the system.

4.1.1 ECG and motion activity data acquisition

The electronic acquisition platform, produced by Shimmer Research™ allows the transduction, amplification and pre-processing of raw ECG signals, and the transmission of elaborated data via Bluetooth to the smartphone. The unit is mounted on a soft textile chest strap (model miCoach™ by Adidas) designed to seamlessly adapt to the user's body shape, bringing full freedom of movement. A smartphone application, running on Android operative system, was developed for preliminary elaboration of sensor data and remote data base (DB) archiving. The application processes the ECG and accelerometer signals for the extraction of three relevant parameters: HR, activity index and RR intervals for further HRV analysis provided by the analysis-decision module. The wearable electronic board collects raw sensor data with on-body signal conditioning. The ECG signal is sampled at 256 Hz and sent to the smartphone with the tri-axial acceleration data (a_x , a_y , a_z). The smartphone application pre-processes user's physiological signal through a stepwise filtering stage aimed at removing typical ECG artifacts and interferences. In particular, baseline wander due to body movements and respiration artifacts is removed using a cubic spline 3rd order interpolation between the fiducial isoelectric points of the ECG [46]. The power line interference and muscular noise are removed using an infinite impulse response (IIR) notch filter at 50 Hz and an IIR low pass filter at 40 Hz. Then, the Pan-Tompkins method is applied [47] to detect the QRS complex and to extract HR and the time series sequence of non-uniform R-R intervals. Since variation of the ECG parameters is significantly affected by the activity performed by the user [48], [49], signal magnitude area [50] is also extracted from the three-axis accelerometer signals in order to measure motion activity levels.

4.1.2 Psychological data acquisition

The acquisition of psychological data is managed by an electronic survey, which is displayed at random times during the day on the application's screen. The survey includes a subset of Likert-type items measured selected from the ESM by Jacobs et al. [41] described in section 4.1.2. Only the ESM items which highest correlation with HRV features were included in the final survey: this choice was made in order to reduce as much as possible the burden on the user during the training phase of the system. The final selected items were (listed in the same sequence of the final survey):

1. What is your stress level? (min: 1; max: 10)
2. This activity is a challenge (min: 1; max: 7)

3. This is something I'm good at (min: 1; max: 7)
4. I would rather be doing something else (min: 1; max: 7)
5. It takes me effort (min: 1; max: 7)

The average time for the completion of a full questionnaire is about 10-15 seconds. During a typical training week, 4-5 surveys per day are collected.

4.2 Analysis-decision module (ADM)

The ADM developed is composed of two main modules: the feature extraction and the classification module.

4.2.1 Data Exchange

The wearable sensors monitor patients and transfer data to web-servers, using a smartphone to collect and pre-elaborate data. In addition, the application allows users to track their own stress levels through a graphical representation (see Section 5.3). The chest band and its electronic act as masters that initiate Bluetooth communication with the Android phone. The Bluetooth protocol has a range of approximately 20 m and provides secure data transmissions. The communication between the phone and central DB is through Wi-Fi or 3G networks. The DSS makes use of information transmitted by the smartphone and stored within the central DB. For each subject is created an user profile able to maintains all history data. In particular the remote DB storage physiological data (HR, RR intervals and activity index) and the corresponding stress value extracted with physiological surveys. Given the *UserID*, the *TimeStamp*, and the *Session*, the ADM retrieve from DB the physiological data together with the questionnaires filled by the user about the stress level perceived. From these data features are extracted and used to train the classification module (training phase). After training, the ADM acts as an expert system and provide to the corresponding user the StressLevel automatically extracted after training of the classification module (testing phase).

The ADM acts asynchronously in respect to the sensor data collection process. At fixed time intervals, the new sensor data belonging to each subject are collected and a feature extraction process takes place in order to create a structured dataset. Classification module is trained and validated with most relevant features for automatic stress assessment.

4.2.2 Feature extraction

Once the data are sent to the remote server, concerning the RR intervals collected over time, the parts of the signals with artifacts are discarded and HRV features are extracted according to the traditional approach proposed by the International Guidelines of HRV [42], [43] to estimate cardiac vagal and sympathetic activities as markers of the autonomic interaction using a Data Exchange module (Figure 2). In time domain were

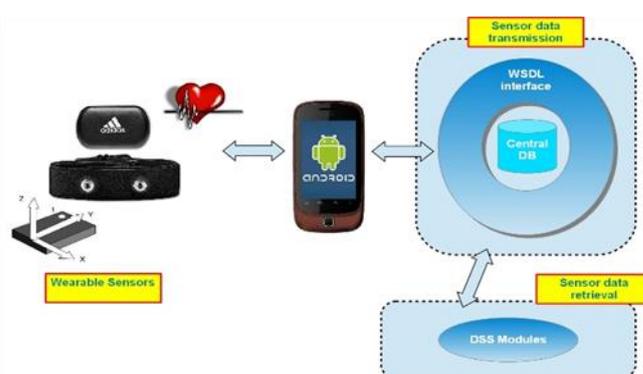


Figure 2. Data Exchange module.

extracted statistical indices from RR time series such as mean (mRR), standard deviation (σRR), root mean square of successive differences of intervals (RMSSD), difference between the longest and shortest RR intervals, and the number of successive differences of intervals which differ by more than 50 ms ($pNN50\%$ expressed as a percentage of the total number of heartbeats analysed), while in frequency domain are extracted parameters for each frequency band, low frequency (LF: 0.03-0.15 Hz) and high frequency (HF: 0.15-0.40 Hz), included absolute powers, peak frequencies (Max LF and Max HF) and the LF/HF power ratio that measure the global sympathetic-parasympathetic equilibrium. In particular, above the threshold, the curve reveals sympathetic dominance, below the threshold, the parasympathetic influence is dominant.

These features are extracted using an estimation of the Power Spectral density (PSD) analysis according to the Burg spectral estimation [51], where the optimal order p was estimated according to the Akaike information criterion [52]. The power of each band is normalized in respect to the total power of the spectrum. Also a nonlinear parameter was extracted, i.e. the Poincaré Plot a useful tool to investigate and combine the differences of the cardiac rhythms during the performed tasks. It is a graphical representation created by plotting all $RR(n)$ on the x-axis versus $RR(n+1)$ on the y-axis. Then, the data are fitted using an ellipse projected according to the line of identity and extracting the two standard deviations (SD) respectively [53] as shown in Figure 3.

4.2.3 Classification module

The classification module is based on machine learning (ML) models, such as artificial neural networks (ANN), based on inductive inference [54]. We decided to use ML model to cope with the non-linear and complex relations between the monitored parameters and the stress level prediction (Table 4). Artificial Neural Networks (ANNs) are particularly suited for solving such problems. They are biologically inspired computational models, which consist of a network composed of artificial neurons. For the implementation of ML for stress level detection a number of steps are needed:

- Initialization of parameters of implemented ML model.
- Training: the model is trained with the features extracted and selected to adapt itself to classify the given inputs. The loaded features, along with self-reported stress levels, generate the training set. The self-reported stress levels are collected

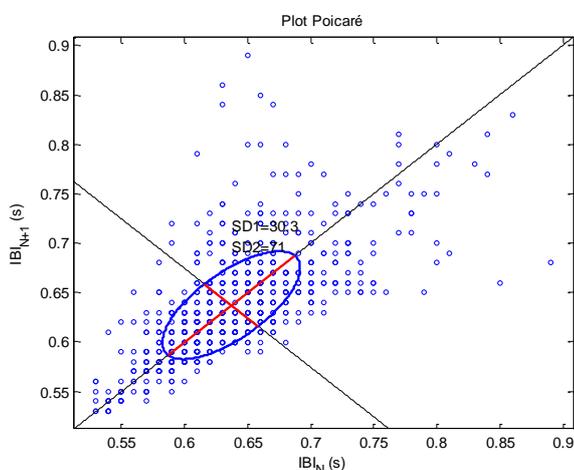


Figure 3. SD1 and SD2 of Poincaré plot observed for a portion of RR interval analysed.

Table 4. Features extracted and analysed from the signal.

No.	Features Extracted	Measure of	Signal
1	mRR	Mean RR interval	RR
2	σRR	Standard deviation RR interval	RR
3	RMSSD	Root mean square of successive differences of intervals	RR
4	$pNN50\%$	Number of successive differences of intervals which differ by more than 50 ms	RR
5	LF	Spectral estimation of low frequency power (0.03-0.15 Hz)	RR
6	Max LF	Max value of low frequency value	RR
7	HF	Spectral estimation of high frequency power (HF: 0.15-0.40 Hz)	RR
8	Min HF	Max value of high frequency value	RR
9	LF/HF	Spectral estimation of power ratio	RR
10	SD1, SD2	Standard deviations of Poincaré plot	RR
11	SMA	Signal magnitude area	Acc. X,Y,Z

during the training phase, in which the participant is prompted at random times during the day with a survey including five items (see Section 5.1.2) that allows the user to self-evaluate, on a Likert scale, perceived levels of stress, following a protocol described in Section 4.1.3. By matching this psychological “ground truth” with sensor data, synaptic weights of networks internal connections are modified in order to force the output to minimize the error with the presented example (in this case the stress level obtained from survey). In this step the architecture of the model and its hyper-parameters are optimized. The examples labelled with the stress-level are used to create a personalized stress prediction model.

- Validation: the model adequately trained, is able to classify the given input in order to present a consequent output value: the value obtained, is the inferred stress level. It is validated in order to guarantee good predictive properties.

During the analysis decision module fine-tuning design, we decided to develop a self-organizing maps (SOM) integrated with fuzzy rules. The SOM is a network structure which provides a topological mapping [55]-[57]. The main difference with the artificial neural network is that it is based on unsupervised learning. It is composed of two-dimensional layer in which all the inputs are connected to each node in the network (Figure 4).

A topographic map is autonomously organized by a cyclic process of comparing input patterns to vectors at each node. The node vector to which inputs match is selectively optimized to present an average of the training data. Then all the training data are represented by the node vectors of the map. Starting with a randomly organized set of nodes, and proceeding to the creation of a feature map representing the prototypes of the input patterns, the training procedure is as follows:

1. Initialization of the weights $w_{ij}(1 \leq i \leq n^F, 1 \leq j \leq m)$ to small random values, where n^F is the total number of selected features (input) and m is the total number of nodes in the map. Set the initial radius of the neighbourhood around node j as $N_j(t)$.
2. Present the inputs $x_1(t), x_2(t), \dots, x_{n^F}(t)$, where $x_i(t)$ is the i th input to node j at time t .
3. Calculate the distance d_j between the inputs and node j by the Euclidean distance to determine j^* which minimizes d_j :

$$d_j = \left\| W_j(t) - X(t) \right\| \quad (1)$$

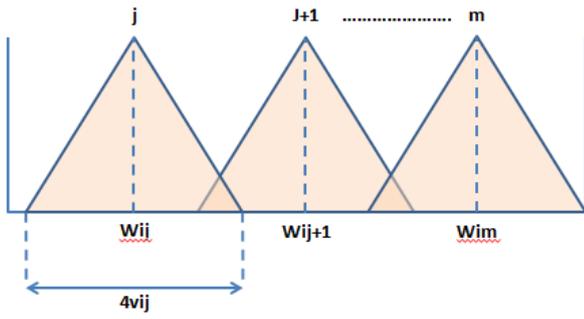


Figure 4. Generation of the fuzzy membership function for the i th input. The number of triangular functions is the equal to the SOM nodes.

Every node is examined to calculate which one's weights are most like the input vector. The winning node is commonly known as the Best Matching Unit (BMU). The radius of the neighborhood of the BMU is then calculated. This is a value that starts large, but diminishes each time-step. Any nodes found within this radius are deemed to be inside the BMU's neighborhood.

- Update the weights w_{ij} of the winning neuron j^* and of its neighborhood neurons $N_{j^*}(t)$ at the time t , for the input vector X , are modified according to the following equation (2) to make them more like the input vector:

$$w_{ij}(t) = w_{ij}(t - 1) + \alpha(t)[X(t) - w_{ij}(t - 1)], \quad (2)$$

where $\alpha(t)$ is the learning rate. Both $\alpha(t)$ and $N_{j^*}(t)$ are controlled so as to decrease in t .

If the process reaches the maximum number of iterations, stop; otherwise, go to (2). At the end of the training process, for each input variable x_i we generated the fuzzy membership function using triangular functions with the centre in the corresponding weight w_{ij} of the map and the corresponding variance v_{ij} , where i is the i th input and j represents the j th node of the map. The centers of the triangular membership functions in the i th input are $(w_{i1} \ w_{i2} \ \dots \ w_{im})$. The corresponding regions were set to $[w_{i1}-2v_{i1}, \ w_{i1}+2v_{i1}]$, $[w_{i2}-2v_{i2}, \ w_{i2}+2v_{i2}]$, ..., $[w_{im}-2v_{im}, \ w_{im}+2v_{im}]$, where m is the last node of the map.

We developed membership functions and fuzzy rules for each HRV parameter, including heart rate and motion activity. In order to reduce the number of fuzzy rules and to improve the system reliability, narrowly separated regions were combined to become a single region. Let the positions of the four corners of region j be l_j , lh_j , rh_j and rl_j (for a triangular membership function, $lh_j = rh_j$). Two neighboring regions $j-1$ and j were merged if they satisfied the following equation (3):

$$\frac{lh_j + rh_j}{2} - \frac{lh_{j-1} + rh_{j-1}}{2} \leq thr, \quad (3)$$

where thr is pre-specified threshold (set to 0.1 in our experiments). This process continued until all regions were well separated in terms of the threshold. Accordingly, some fuzzy regions had trapezoidal shapes instead of triangular ones as is shown in Figure 5.

After that, we generated fuzzy rules as a set of associations of the form "if antecedent conditions hold, then consequent conditions hold". Each feature was normalized to the range of $[0.0,1.0]$ and each region of fuzzy membership function was labeled as R1, R2, ..., RN. An input was assigned to the label of a region where the maximum membership value was obtained. In

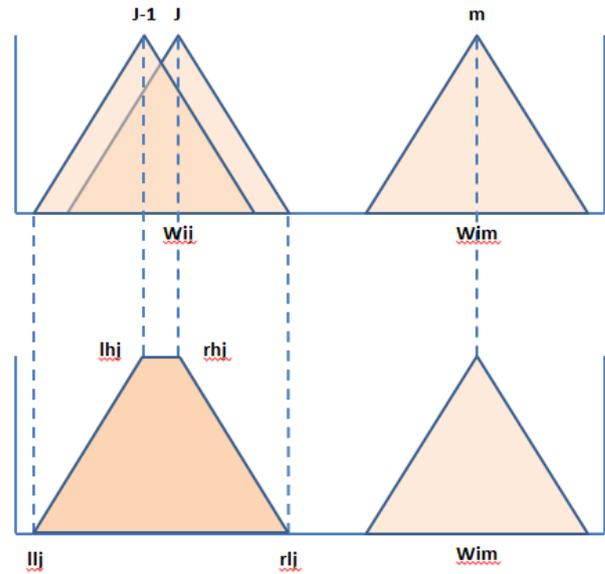


Figure 5. Trapezoidal function obtained for neighboring regions.

particular we adopted the method proposed by Wang et al. [56] where each training sample produced a fuzzy rule. An example of rule generated is: IF feature1 is R1 AND feature2 is RN AND feature3 is R2 AND feature4 is R3 AND feature5 is R6, AND feature6 is R8 AND feature M is R3 THEN it is Medium Stress level.

Finally, the number of all the fuzzy rules was the same order of the training samples. The problem was that a large number of training patterns may lead to repeated or conflicting rules. To deal with this problem, we recorded the number of rules repeated during the learning process. Those rules supported by a large number of examples were saved. A centroid defuzzification formula was used to determine the output for each input pattern:

$$Z = \frac{\sum_{i=1}^k D_p^i O^i}{\sum_{i=1}^k D_p^i}, \quad (4)$$

where Z is the output, k is the number of rules, O^i is the class generated by rule i and D_p^i measures how the input vector fit the i th rule. D_p^i is given by the product of degrees of the pattern in the regions which the i th rule occupies. The output is within $[0,5]$ for numeral recognition of stress level (0=unknown, 1=low stress, 2=mild stress, 3=elevated stress, 4=high stress, 5=severe stress). The output Z was adapted taking the nearest smaller integer value. Fuzzy rules do not necessarily occupy all fuzzy regions in input space. There could be some regions where no related rule exists. This is the case when the denominator in equation (4) is zero. We label the corresponding input stress level as unknown. After training the model was designed to be able to discriminate until 6 different classes during test phase with the trained fuzzy classifier.

For each user, the DSS is trained on the basis of the evaluation of the stress level supplied by the surveys, being such data collected by the mobile application and transmitted to the database. As the DSS training is completed, the DSS acts as an expert system, supplying the stress level information for each patient as new sensor data becomes available within the database (Figure 6).

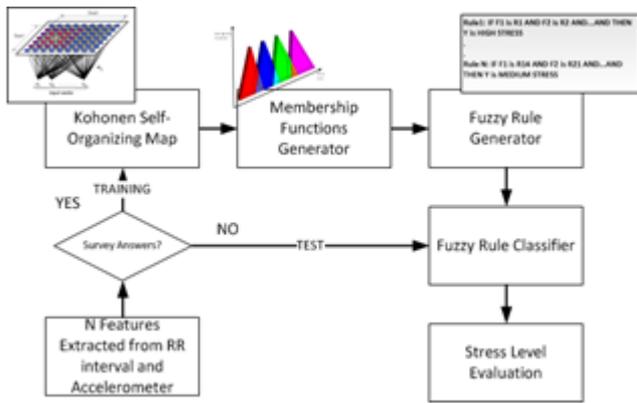


Figure 6. Architecture of the automatic Stress classification module.

4.3 Mobile visualization module

After data collection, the mobile application can poll the stress level report from the remote database provided by the analysis-decision module (see section 5.2) at definite time intervals, or at user's request. The *StressTracker* visualization component of the mobile application provide the user with a graphic visualization of measured stress events. This information is visualized by the *StressTracker* (Figure 7, right) which shows the number of detected stressful events over the course of last day, week, or month respectively.

5. CROSS VALIDATION

In order to check the stress detection capability of the personalized model, the performance of the classification task was evaluated using the leave one out cross validation process (LOOCV), where each fold consists of one session of data acquisition left-out. This method is an iterative process in which one session is recruited into the dataset each time for validation. The SOM classifier combined with fuzzy rules was trained using the remaining data and validated on the single, left-out validation point. This ensures that the validation is unbiased, because the classifier does not see the validation input sample during its training. One by one, each available session for each subject was recruited for validation. The performances of the system were assessed by using the confusion matrix, in which the generic element i,j indicates how many times (in mean percentage \pm SD)

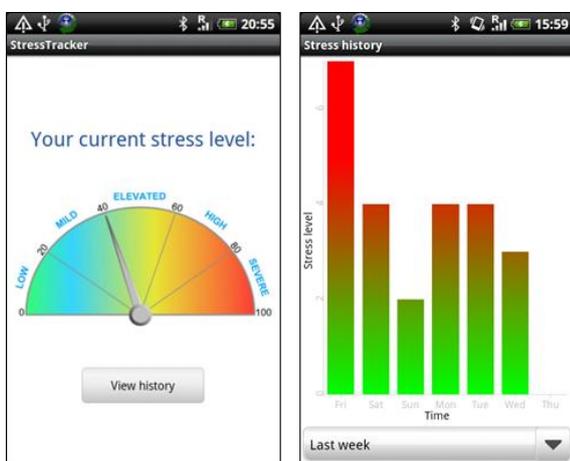


Figure 7. On the left side is reported the current stress level of the user, while on the right side are reported the stressful events detected over the last week.

Table 5. Confusion Matrixes.

A. The confusion matrix obtained with the first subject:

	Stress	No stress
Stress	89 \pm 6.2	11 \pm 2.1
No stress	13.3 \pm 3.6	86.6 \pm 5

B. The confusion matrix obtained with the second subject:

	Stress	No stress
Stress	66.6 \pm 11	33.3 \pm 6.7
No stress	30 \pm 3.1	70 \pm 5.9

C. The confusion matrix obtained with the third subject:

	Stress	No stress
Stress	95 \pm 3.3	5 \pm 4.6
No stress	26 \pm 3.1	73 \pm 8.2

a pattern belonging to the class i was classified as belonging to the class j .

5.1 Cross validation results

The system is able to recognize the presence of stress in the selected population. In particular, the results obtained with three subjects are reported in Table 5. The mean \pm SD percentages of the confusion matrices obtained with the first subject are reported in Table 5A. The model correctly identifies the presence of stress with percentage of correct classifications of 89.0 % and the absence of stress with percentage of correct classifications of 86.6 %. The mean \pm SD percentages of the confusion matrices obtained with the second subject are reported in the Table 5B. The model correctly identifies the presence of stress with percentage of correct classifications of 66.6 % and the absence of stress with percentage of correct classifications of 70.0 %.

The mean \pm SD percentages of the confusion matrices obtained with the third subject are reported in Table 5C. The model correctly identifies the presence of stress with percentage of correct classifications of 95.7 % and the absence of stress with percentage of correct classifications of 73.0 %. These results demonstrate the high discriminatory power of the system.

6. CONCLUSION

We described the design, key functional features and preliminary validation of a personalized system for monitoring stress in naturalistic environments. The original contribution of this work concerns the development of a new methodology, which allows to use subjective evaluation of potentially stressful situations in the calibration and training of the classification system. In particular, incorporating self-reported psychological data in the system knowledge base allows for a more comprehensive and personalized definition of the stress response as measured by HRV indices. An objective of future research is to validate the accuracy of the personalized stress detection model against other physiological markers, i.e. salivary cortisol collected during daily life activities.

ACKNOWLEDGEMENT

This work was supported by the European funded project "Interstress-Interreality in the management and treatment of stress-related disorders", Grant Number: FP7-247685.

The co-authors would like to thank Giovanni Pioggia and Andrea Gaggioli who have equally contributed to the manuscript.

REFERENCES

- [1] Y. Yang, Daily stressor, daily resilience, and daily somatization: The role of trait aggression, *Personality and Individual Differences*, vol. 165, 2020, 110141. DOI: [10.1016/j.paid.2020.110141](https://doi.org/10.1016/j.paid.2020.110141)
- [2] S. Scholten, K. Lavallee, J. Velten, X. C. Zhang, J. Margraf, The brief daily stressors screening tool: An introduction and evaluation, *Stress and Health*, vol. 36, no. 5, 2020, pp. 686-692. DOI: [10.1002/smi.2965](https://doi.org/10.1002/smi.2965)
- [3] L. Jandorf, E. Deblinger, J. M. Neale, A. A. Stone, Daily versus major life events as predictors of symptom frequency: A replication study, *Journal of General Psychology*, vol. 113, no. 3, 1986, pp. 205-218. DOI: [10.1080/00221309.1986.9711031](https://doi.org/10.1080/00221309.1986.9711031)
- [4] P. J. Brantley, G. N. Jones, Daily stress and stress-related disorders, *Annals of Behavioral Medicine*, vol. 15, no. 1, 1993, pp. 17-25. DOI: [10.1093/abm/15.1.17](https://doi.org/10.1093/abm/15.1.17)
- [5] T. D. Tran, J. Kim, N.-H. Ho, H.-J. Yang, S. Pant, S.-H. Kim, G.-S. Lee, Stress analysis with dimensions of valence and arousal in the wild, *Applied Sciences*, vol. 11, no. 11, 2021, 5194. DOI: [10.3390/app11115194](https://doi.org/10.3390/app11115194)
- [6] M. Kusserow, O. Amft, G. Tröster, Monitoring Stress-Arousal in the Wild - Measuring the Intangible, *IEEE Pervasive Computing*, vol. 12, no. 2, 2013, pp. 28-37. DOI: [10.1109/MPRV.2012.56](https://doi.org/10.1109/MPRV.2012.56)
- [7] J. A. Healey, R. W. Picard, Detecting stress during real-world driving tasks using physiological sensors, *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, 2005, pp. 156-166. DOI: [10.1109/ITITS.2005.848368](https://doi.org/10.1109/ITITS.2005.848368)
- [8] G. Tartarisco, G. Baldus, D. Corda, R. Raso, A. Arnao, M. Ferro, A. Gaggioli, G. Pioggi, Personal Health System architecture for stress monitoring and support to clinical decisions, *Computer Communications*, vol. 35, no. 11, 2012, pp. 1296-1305. DOI: [10.1016/j.comcom.2011.11.015](https://doi.org/10.1016/j.comcom.2011.11.015)
- [9] N. Carbonaro, P. Cipresso, A. Tognetti, G. Anania, D. De Rossi, A. Gaggioli, G. Riva, A mobile biosensor to detect cardiorespiratory activity for stress tracking, 3rd International Workshop on Pervasive Computing Paradigms for Mental Health, 2013. DOI: [10.4108/icst.pervasivehealth.2013.252357](https://doi.org/10.4108/icst.pervasivehealth.2013.252357)
- [10] P. Cipresso, A. Gaggioli, S. Serino, S. Raspelli, C. Vigna, F. Pallavicini, G. Riva, Inter-reality in the evaluation and treatment of psychological stress disorders: The INTERSTRESS project, *Annual Review of CyberTherapy and Telemedicine*, vol. 181, 2012, pp. 8-11. DOI: [10.3233/978-1-61499-121-2-8](https://doi.org/10.3233/978-1-61499-121-2-8)
- [11] N. Carbonaro, P. Cipresso, A. Tognetti, G. Anania, D. De Rossi, F. Pallavicini, A. Gaggioli, G. Riva, Psychometric assessment of cardio-respiratory activity using a mobile platform, in *Wearable Technologies: Concepts, Methodologies, Tools, and Applications*, 2018. DOI: [10.4018/978-1-5225-5484-4.ch037](https://doi.org/10.4018/978-1-5225-5484-4.ch037)
- [12] A. Gaggioli, F. Pallavicini, L. Morganti, S. Serino, C. Scaratti, M. Briguglio, G. Crifaci, N. Vetrano, A. Giulintano, G. Bernava, G. Tartarisco, G. Pioggia, S. Raspelli, P. Cipresso, C. Vigna, A. Grassi, M. Baruffi, B. Wiederhold, G. Riva, Experiential virtual scenarios with real-time monitoring (interreality) for the management of psychological stress: a block randomized controlled trial, *Journal of medical Internet research*, vol. 16, no. 7, 2014, e167. DOI: [10.2196/jmir.3235](https://doi.org/10.2196/jmir.3235)
- [13] A. Gaggioli, P. Cipresso, S. Serino, D. M. Campanaro, F. Pallavicini, B. K. Wiederhold, G. Riva, Positive technology: a free mobile platform for the self-management of psychological stress, *Studies in health technology and informatics*, vol. 199, 2014, pp. 25-29. DOI: [10.3233/978-1-61499-401-5-25](https://doi.org/10.3233/978-1-61499-401-5-25)
- [14] D. Giakoumis, A. Drosou, P. Cipresso, D. Tzovaras, G. Hassapis, A. Gaggioli, G. Riva, Real-time monitoring of behavioural parameters related to psychological stress, *Annual Review of Cyber Therapy and Telemedicine*, vol. 10, 2012, pp. 287-291. DOI: [10.3233/978-1-61499-121-2-287](https://doi.org/10.3233/978-1-61499-121-2-287)
- [15] L. Salahuddin, J. Cho, M. G. Jeong, D. Kim, Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings, 2007, pp. 4656-4659. DOI: [10.1109/IEMBS.2007.4353378](https://doi.org/10.1109/IEMBS.2007.4353378)
- [16] P. Melillo, M. Bracale, L. Pecchia, Nonlinear Heart Rate Variability features for real-life stress detection. Case study: Students under stress due to university examination, *BioMedical Engineering Online*, vol. 10, 2011, 96. DOI: [10.1186/1475-925X-10-96](https://doi.org/10.1186/1475-925X-10-96)
- [17] D. Kimhy, P. Delespaul, H. Ahn, S. Cai, M. Shikhman, J. A. Lieberman, D. Malaspina, R. P. Sloan, Concurrent measurement of 'real-world' stress and arousal in individuals with psychosis: Assessing the feasibility and validity of a novel methodology, *Schizophrenia Bulletin*, vol. 36, no. 6, 2010, pp. 1131-1139. DOI: [10.1093/schbul/sbp028](https://doi.org/10.1093/schbul/sbp028)
- [18] R. P. Sloan, P. A. Shapiro, E. Bagiella, S. M. Boni, M. Paik, J. T. Bigger Jr., R. C. Steinman, J. M. Gorman, Effect of mental stress throughout the day on cardiac autonomic control, *Biological Psychology*, vol. 37, no. 2, 1994, pp. 89-99. DOI: [10.1016/0301-0511\(94\)90024-8](https://doi.org/10.1016/0301-0511(94)90024-8)
- [19] G. G. Berntson, J. T. Cacioppo, *Heart Rate Variability: Stress and Psychiatric Conditions, in Dynamic Electrocardiography*, 2007. DOI: [10.1002/9780470987483.ch7](https://doi.org/10.1002/9780470987483.ch7)
- [20] P. Cipresso, D. Colombo, G. Riva, Computational psychometrics using psychophysiological measures for the assessment of acute mental stress, *Sensors (Switzerland)*, vol. 19, no. 4, 2019. DOI: [10.3390/s19040781](https://doi.org/10.3390/s19040781)
- [21] S. Cohen, R. C. Kessler, L. U. Gordon, Strategies for measuring stress in studies of psychiatric and physical disorders, *Measuring stress: A guide for health and social scientists*, Oxford University Press, 1995, pp. 3-26, ISBN: 978-0195121209.
- [22] M. Csikszentmihalyi, R. Larson, Validity and reliability of the experience-sampling method, in *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi*, 2014, pp 35-54. DOI: [10.1007/978-94-017-9088-8_3](https://doi.org/10.1007/978-94-017-9088-8_3)
- [23] M. Memar, A. Mocaribolhassan, Stress level classification using statistical analysis of skin conductance signal while driving, *SN Applied Sciences*, vol. 3, no. 1, 2021, 64. DOI: [10.1007/s42452-020-04134-7](https://doi.org/10.1007/s42452-020-04134-7)
- [24] K. Wang, P. Guo, An Ensemble Classification Model with Unsupervised Representation Learning for Driving Stress Recognition Using Physiological Signals, *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, 2021, pp. 3303-3315. DOI: [10.1109/ITITS.2020.2980555](https://doi.org/10.1109/ITITS.2020.2980555)
- [25] N. van Berkel, D. Ferreira, V. Kostakos, The experience sampling method on mobile devices, *ACM Computing Surveys*, vol. 50, no. 6, 2017, pp 1-40. DOI: [10.1145/3123988](https://doi.org/10.1145/3123988)
- [26] D. Kahneman, A. B. Krueger, D. A. Schkade, N. Schwarz, A. A. Stone, A survey method for characterizing daily life experience: The day reconstruction method, *Science*, vol. 306, no. 5702, 2004, pp. 1776-1780. DOI: [10.1126/science.1103572](https://doi.org/10.1126/science.1103572)
- [27] A. A. Stone, J. E. Schwartz, D. Schkade, N. Schwarz, A. Krueger, D. Kahneman, A population approach to the study of emotion: Diurnal rhythms of a working day examined with the day reconstruction method, *Emotion*, vol. 6, no. 1, 2006, pp. 139-149. DOI: [10.1037/1528-3542.6.1.139](https://doi.org/10.1037/1528-3542.6.1.139)
- [28] T. S. Conner, L. F. Barrett, Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine, *Psychosomatic Medicine*, vol. 74, no. 4, 2012, pp. 327-337. DOI: [10.1097/PSY.0b013e3182546f18](https://doi.org/10.1097/PSY.0b013e3182546f18)

- [29] C. Kirschbaum, D. H. Hellhammer, Salivary cortisol in psychobiological research: An overview, *Neuropsychobiology*, vol. 22, no. 3, 1989, pp. 150-169.
DOI: [10.1159/000118611](https://doi.org/10.1159/000118611)
- [30] G. E. Miller, E. Chen, E. S. Zhou, If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans, *Psychological Bulletin*, vol. 133, no. 1. 2007, pp. 25-45.
DOI: [10.1037/0033-2909.133.1.25](https://doi.org/10.1037/0033-2909.133.1.25)
- [31] D. J. Hruschka, B. A. Kohrt, C. M. Worthman, Estimating between- and within-individual variation in cortisol levels using multilevel models, *Psychoneuroendocrinology*, vol. 30, no. 7, 2005, pp. 698-714.
DOI: [10.1016/j.psyneuen.2005.03.002](https://doi.org/10.1016/j.psyneuen.2005.03.002)
- [32] G. H. Ice, A. Katz-Stein, J. Himes, R. L. Kane, Diurnal cycles of salivary cortisol in older adults, *Psychoneuroendocrinology*, vol. 29, no. 3, 2004, pp. 355-370.
DOI: [10.1016/S0306-4530\(03\)00034-9](https://doi.org/10.1016/S0306-4530(03)00034-9)
- [33] H. Mansikka, P. Simola, K. Virtanen, D. Harris, L. Oksama, Perceived mental stress and reactions in heart rate variability- a pilot study among employees of an electronics company, *International Journal of Occupational Safety and Ergonomics*, vol. 14, no. 3, 2008, pp. 1344-1352.
DOI: [10.1080/10803548.2008.11076767](https://doi.org/10.1080/10803548.2008.11076767)
- [34] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers, T. D. Wager, A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health, *Neuroscience and Biobehavioral Reviews*, vol. 36, no. 2. 2012, pp. 747-756.
DOI: [10.1016/j.neubiorev.2011.11.009](https://doi.org/10.1016/j.neubiorev.2011.11.009)
- [35] D. Kim, Y. Seo, J. Cho, C. H. Cho, Detection of subjects with higher self-reporting stress scores using heart rate variability patterns during the day, 2008, pp. 682-685.
DOI: [10.1109/ieembs.2008.4649244](https://doi.org/10.1109/ieembs.2008.4649244)
- [36] A. Muaremi, B. Arnrich, G. Tröster, Towards Measuring Stress with Smartphones and Wearable Devices During Workday and Sleep, *BioNanoScience*, vol. 3, no. 2, 2013, pp. 172-183.
DOI: [10.1007/s12668-013-0089-2](https://doi.org/10.1007/s12668-013-0089-2)
- [37] M. E. Morris, Q. Kathawala, T. K. Leen, E. E. Gorenstein, F. Guilak, W. DeLeeuw, M. Labhard, Mobile therapy: Case study evaluations of a cell phone application for emotional self-awareness, *Journal of Medical Internet Research*, vol. 12, no. 2, 2010, e10.
DOI: [10.2196/jmir.1371](https://doi.org/10.2196/jmir.1371)
- [38] B. Cinaz, B. Arnrich, R. la Marca, and G. Tröster, Monitoring of mental workload levels during an everyday life office-work scenario, *Personal and Ubiquitous Computing*, vol. 17, no. 2, 2013, pp. 229-239.
DOI: [10.1007/s00779-011-0466-1](https://doi.org/10.1007/s00779-011-0466-1)
- [39] A. Gaggioli, G. Pioggia, G. Tartarisco, G. Baldus, D. Corda, P. Cipresso, G. Riva, A mobile data collection platform for mental health research, *Personal and Ubiquitous Computing*, vol. 17, no. 2, 2013, pp. 241-251.
DOI: [10.1007/s00779-011-0465-2](https://doi.org/10.1007/s00779-011-0465-2)
- [40] N. Jacobs, I. Myin-Germeys, C. Derom, P. Delespaul, J. van Os, N. A. Nicolson, A momentary assessment study of the relationship between affective and adrenocortical stress responses in daily life, *Biological Psychology*, vol. 74, no. 1, 2007, pp. 60-66.
DOI: [10.1016/j.biopsycho.2006.07.002](https://doi.org/10.1016/j.biopsycho.2006.07.002)
- [41] R. Larson, P. A. E. G. Delespaul, Analyzing Experience Sampling data: a guide book for the perplexed, in *The Experience of Psychopathology*, 2010, pp. 58-78.
DOI: [10.1017/cbo9780511663246.007](https://doi.org/10.1017/cbo9780511663246.007)
- [42] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, P. J. Schwartz, Heart rate variability. Standards of measurement, physiological interpretation, and clinical use, *European Heart Journal*, vol. 17, no. 3. 1996, pp. 354-381.
DOI: [10.1093/oxfordjournals.eurheartj.a014868](https://doi.org/10.1093/oxfordjournals.eurheartj.a014868)
- [43] M. Malik, Heart rate variability: Standards of measurement, physiological interpretation, and clinical use, *Circulation*, vol. 93, no. 5, 1996, pp. 1043-1065.
DOI: [10.1161/01.cir.93.5.1043](https://doi.org/10.1161/01.cir.93.5.1043)
- [44] V. Magagnin, M. Mauri, P. Cipresso, L. Mainardi, E. Brown, S. Cerutti, M. Villamira, R. Barbieri, Heart rate variability and respiratory sinus arrhythmia assessment of affective states by bivariate autoregressive spectral analysis, in *Computing in Cardiology*, 2010, vol. 37, pp. 145-148.
- [45] M. Mauri, V. Magagnin, P. Cipresso, L. Mainardi, E. N. Brown, S. Cerutti, M. Villamira, R. Barbieri, Psychophysiological signals associated with affective states, 2010, pp. 3563-3566.
DOI: [10.1109/IEMBS.2010.5627465](https://doi.org/10.1109/IEMBS.2010.5627465)
- [46] R. Jane, P. Laguna, N. v. Thakor, P. Caminal, Adaptive baseline wander removal in the ECG: Comparative analysis with cubic spline technique, 1992, pp. 143-146.
DOI: [10.1109/CIC.1992.269426](https://doi.org/10.1109/CIC.1992.269426)
- [47] J. Pan, W. J. Tompkins, A Real-Time QRS Detection Algorithm, *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, 1985, pp. 230-236.
DOI: [10.1109/TBME.1985.325532](https://doi.org/10.1109/TBME.1985.325532)
- [48] J. H. Houtveen, P. F. C. Groot, E. J. C. de Geus, Effects of variation in posture and respiration on RSA and pre-ejection period, *Psychophysiology*, vol. 42, no. 6, 2005, pp. 713-719.
DOI: [10.1111/j.1469-8986.2005.00363.x](https://doi.org/10.1111/j.1469-8986.2005.00363.x)
- [49] F. H. Wilhelm, P. Grossman, M. I. Müller, Bridging the Gap between the Laboratory and the Real World- Integrative Ambulatory Psychophysiology, *Handbook of research methods for studying daily life*, 2012, The Guilford Press.
- [50] C. V. C. Bouten, K. T. M. Koekkoek, M. Verduin, R. Kodde, J. D. Janssen, A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity, *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 3, 1997, pp. 136-147.
DOI: [10.1109/10.554760](https://doi.org/10.1109/10.554760)
- [51] E. Pardo-Igúzquiza and F. J. Rodríguez-Tovar, *Maximum Entropy Spectral Analysis*, 2021, Springer.
DOI: [10.1007/978-3-030-26050-7_197-1](https://doi.org/10.1007/978-3-030-26050-7_197-1)
- [52] H. Akaike, Fitting autoregressive models for prediction, *Annals of the Institute of Statistical Mathematics*, vol. 21, no. 1, 1969, pp. 243-247.
DOI: [10.1007/BF02532251](https://doi.org/10.1007/BF02532251)
- [53] M. Brennan, M. Palaniswami, P. Kamen, Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability?, *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, 2001, pp. 1342-1347.
DOI: [10.1109/10.959330](https://doi.org/10.1109/10.959330)
- [54] S. R. Sain, V. N. Vapnik, *The Nature of Statistical Learning Theory*, *Technometrics*, vol. 38, no. 4, 1996, 409.
DOI: [10.2307/1271324](https://doi.org/10.2307/1271324)
- [55] T. Kohonen, The self-organizing map, *Neurocomputing*, vol. 21, no. 1-3, 1998, pp. 1-6.
DOI: [10.1016/S0925-2312\(98\)00030-7](https://doi.org/10.1016/S0925-2312(98)00030-7)
- [56] L. X. Wang, J. M. Mendel, Generating Fuzzy Rules by Learning from Examples, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 6, 1992, pp. 1414-1427.
DOI: [10.1109/21.199466](https://doi.org/10.1109/21.199466)
- [57] G. Tartarisco, N. Carbonaro, A. Tonacci, G. M. Bernava, A. Arnao, G. Crifaci, P. Cipresso, G. Riva, A. Gaggioli, D. De Rossi, A. Tognetti, G. Pioggia, Neuro-fuzzy physiological computing to assess stress levels in virtual reality therapy, *Interacting with Computers*, vol. 27, no. 5, 2015, pp. 521-533.
DOI: [10.1093/ivc/iwv010](https://doi.org/10.1093/ivc/iwv010)