Library Progress International
Vol.44 No. 3, Jul-Dec 2024: P.11167-11175

Print version ISSN 0970 1052
Online version ISSN 2320 317X

Original Article *Available online at www.bpasjournals.com*

Developing Standards for a Model Reference Dataset for Multimodal Human Stress Identification

[1]Reshu Gupta, [2]Aniket Kumar

[1]Research Scholar, Shobhit University, Meerut
[2]School of Electronics, Electrical & Mechanical Engineering (SEE & ME)
Shobhit Institute of Engineering & Technology (Deemed-to-be University)
NH-58, Modipuram, Meerut,Uttar Pradesh, India.

How to cite this article: Reshu Gupta, Aniket Kumar (2024) Developing Standards for a Model Reference Dataset for Multimodal Human Stress Identification**.** *Library Progress International*, 44(3), 11167-11175.

Abstract
Stress is essential for improving performance and enabling day-to-day efficient functioning. But when stress becomes more than a person can handle, it can negatively impact their productivity and health. An automated system for measuring stress can assist users in keeping their stress levels within healthy bounds by tracking stress levels via physiological and physical factors. It is crucial to create and evaluate different algorithms using a reference dataset that includes multimodal stress responses in order to establish such a system. Specific standards based on findings from both clinical and empirical research must be met by this dataset. In order to facilitate the establishment of a reference dataset for multimodal human stress detection, this work presents a thorough set of criteria. The suggested specifications cover both personal and technological elements, such as the selection of sample populations, the selection of stressors, the integration of various stress response modalities, the techniques for data annotation, and the choice of data collection tools. None of the stress datasets currently in the public domain entirely satisfied these predetermined standards, according to an analysis of them. Therefore, it is essential to give top priority to creating a reference dataset that complies with these particular specifications. The comparability and reliability of study findings in the field of stress detection will be improved by such an endeavour.

Keywords: *Multimodal Analysis, Stress Modulation, analysis of requirements, Health Monitoring Systems, Algorithm Development*

I. INTRODUCTION

One of the top 10 socioeconomic factors influencing health inequalities is stress. The detrimental consequences of stress on health and the enormous costs to society it entails have been brought to light by eminent organizations such as the World Health Organization (WHO), the American Psychological Association, and the Occupational Safety and Health Administration (OSHA). A conservative estimate from 2002 estimated that stress at work cost the European Union's economies alone €20 billion. Hans Selye coined the term "stress" with his General Adaptation Syndrome theory, which holds that each person's capacity for adaptation to stressful circumstances is unique and has a limit. Stress was once thought to be only a reaction to a perceived threat, but throughout time, it came to be recognized as a legitimate medical disease. In an effort to restore homeostasis, the body perceives a threat or challenge cognitively, which sets off the stress response. Although stress is a normal and necessary component of life, repeated, severe, and protracted exposure can impair immunity and cause delays in healing. Thus, keeping an eye on stress levels is essential to lowering exposure and encouraging a quicker recovery.

Stress responses are biologically triggered by the hypothalamic-pituitary-adrenal (HPA) axis and the sympathetic nervous system (SNS). The HPA axis is responsible for sending chemical and neurological signals that prime the body to respond to perceived threats. Chemical signals, mostly hormones delivered through the bloodstream, have a longer-lasting impact than neural signals, which are swift and fleeting. Feedback is given by cortisol, which is commonly recognized as a critical stress biomarker, to inhibit HPA activity. Even so, the blood-circulating hormone cortisol also sets off chemical reactions that may result in permanent cellular alterations. Increased concentrations of these substances have been linked to cardiovascular illnesses, which cause a considerable number of fatalities each year.

Numerous investigators have utilized a range of data collection techniques to identify stress early on and have created a variety of stress detection models. These methods have caught a variety of stress response modalities, used a range of stress stimuli, and used diverse sensors. Standardized stress detection algorithms have been hindered by the inconsistent collected datasets caused by the differences in data collection techniques. Given that stress modeling necessitates intricate, multivariate data processing, it is imperative that datasets contain important variables that impact stress response manifestation and assessment. These variables fall into two general categories: internal (such as age, gender, and health) and external (such as stressors, surroundings, and sensors).

To summarize, the lack of consistency in current datasets and the requirement to consider the various elements affecting stress reactions emphasize the need for the creation of a reference dataset for multimodal human stress detection. Reviewing clinical and technical literature is the goal of this project, which also aims to identify a set of specifications for this kind of reference dataset. A set of evaluation criteria is constructed based on the practices and findings from empirical and clinical research, and openly accessible datasets are evaluated in accordance with these criteria.

The several stress assessment techniques and the publicly accessible stress datasets are covered in Section II. The primary contribution and result of this work is outlined in Section III, which also describes the requirements for a reference dataset for multimodal stress recognition. The stress datasets are assessed in Section IV using standards that are derived from the given specifications. Ultimately, the study's result is presented in Section V.

II. RELATED WORK

A. Techniques for Stress Assessment

For more than forty years, computer scientists, psychologists, and doctors have been interested in stress studies. The goal of research is to comprehend, model, and categorize patterns of stress response. Though stress can be assessed in a variety of ways due to its manifestation across numerous physiological pathways, there is currently no consensus on a single, accurate method. For more than forty years, computer scientists, psychologists, and doctors have been interested in stress studies. The goal of research is to comprehend, model, and categorize patterns of stress response. Though stress can be assessed in a variety of ways due to its manifestation across numerous physiological pathways, there is currently no consensus on a single, accurate method.

Physiological, psychological, and behavioral domains are where stress response mechanisms can be divided into categories. Chemical and physical alterations in the organism are examples of physiological modalities. Elevated blood plasma cortisol and catecholamine levels are frequently regarded by clinicians as stress indicators . These hormones are not practical for daily stress monitoring, though, as their collection necessitates medical knowledge and bothersome sampling techniques.

Heart rate variability (HRV), which may be determined with an electrocardiogram (ECG), has been linked to fluctuations in cortisol levels, according to clinical research. Another sign of stress is skin conductance (SC), which can be determined by electrodermal activity (EDA) or galvanic skin responses (GSR). Stress-inducing stimuli have also been linked to pupil dilation, a drop in skin temperature, and an increase in perinasal sweating. Psychological research employs instruments like the Perceived Stress Scale and the State-Trait Anxiety Inventory to measure stress based on self-reported sensations. These scales are frequently used to indicate stress levels in data that has been gathered. Despite being seen as less accurate and having subjective biases, self-reports can nonetheless shed light on a person's perceived psychological health. Instantaneous self-reports can provide additional understanding of a person's locus of control and stress perception when paired with personality-related data.

It is crucial to collect this data via questionnaires since various factors, including food intake, medication and pre-existing conditions, and physical activity, might affect how people respond to stress. Documentation of task-dependent cognitive, sensory, and motor abilities is also necessary to take into consideration expected differences in responses to identical stimuli between individuals.

B. Public Human Stress Datasets

Drivedb: Five datasets that are accessible to the public have been chosen for assessment. This section includes a description of each dataset. These datasets will be compared in Section IV in order to determine how well they satisfy the criteria for a reference dataset.
One of the first studies in the field of automatic stress detection, the Drivedb dataset was assembled by Healey and Picard and is still one of the most often used public datasets . It was gathered with the intention of identifying drivers' general stress levels. Participants drove specified routes with different cognitive burdens, and stress responses were monitored during the drive. In order to confirm cognitive load and manually evaluate stress levels based on head movements, video footage of the drivers was recorded.

The dataset contains a variety of physiological modalities that were recorded in an ambulatory setting, including heart rate (HR), breathing, electromyogram (EMG), electrocardiogram (ECG), and galvanic skin reaction (GSR). The dataset has some drawbacks, even though it encouraged more study into stress detection. These include the inability to assess the driver's reaction time to stress stimuli because the video and sensor clocks were not synchronized, and the lack of data regarding the driver's cognitive state because self-reports were not gathered.

SWELL Knowledge Work (SWELL-KW) Dataset: The SWELL-KW dataset was developed by Koldijk et al. [26] to investigate stress-related behaviors in knowledge workers through the use of a ubiquitous, context-aware technology. In an office setting, distractions and time constraints were used as stressors. The collection contains physiological information including skin conductance (SC) and electrocardiogram (ECG), as well as computer interactions, face expressions, and body postures. Questionnaires were used to collect subjective self-reports of stress in order to establish ground truth. The dataset's preliminary findings showed that it was possible to distinguish between stressful and typical working environments. However, the authors had difficulty separating the impact of cognitive elements on the subjects' real stress levels because of their underreported cognitive ability.

Speech Under Stress (SUS) Datasets: Hansen et al. collected three datasets with the main goal of creating reliable speech processing algorithms to investigate how stress and emotion affect speech: Speech Under Stress Conditions (SUSC), Speech Under Simulated and Actual Stress (SUSAS), and the DERA License Plate (DLP) datasets. Single-word utterances were captured in contexts that were different from everyday work, such as aviation communication. Nevertheless, the investigation of interindividual variations in stress reactions is constrained by the fact that these datasets exclusively concentrate on speech, leaving out other stress-related modalities.

Schmidt et al. assembled the Wearable Stress and Affect Detection (WESAD): dataset in order to provide high-quality multimodal data for identifying states of amusement and stress (affect). The Trier Social Stress Test was used as the stress-inducing stimuli to gather data from 15 participants for the purpose of detecting stress. In order to offer context, the dataset incorporates triaxial acceleration along with a number of physiological modalities, including blood volume pulse (BVP), electrocardiogram (ECG), electrodermal activity (EDA), electromyogram (EMG), respiration, and body temperature. Two devices were used for data collection: the Empatica E4, worn on the wrist, and the RespiBAN Professional, worn on the breast. Four self-reports were given, together with any supplementary notes that could be found.

III. REQUIREMENTS ANALYSIS

A reliable stress response pattern can only be developed through a well-structured data collection process, whether conducted in ambulatory or laboratory-based settings. To ensure the validity and credibility of the study's conclusions based on the collected data, we established a five-part framework. Each component of this framework included key questions that guided the requirements analysis process. Our study was grounded in evidence from clinical and empirical research. The following methodology was employed to support the requirements analysis for a reference dataset for multimodal human stress recognition:

I. **Population Selection**: The study's findings should be generalizable to a broad population.
– Who should be selected as participants?
– How many participants are required?
II. **Stress Stimuli Selection**: The study's outcomes should be applicable to real-life situations.
– Which stress-inducing stimuli should be employed in a controlled laboratory environment?
III. **Stress Modalities Selection**: The study's outcomes should account for variations in individual stress responses.
– Which physiological and behavioral modalities should be recorded?
IV. **Sensing Device Selection and Configuration**: The study should be designed to facilitate practical, real-world implementation and reproducibility.
– Which sensors should be used for data collection?
– What specific details about the sensors should be documented?
– How should the sensors be configured?
V. **Self-Reported Information**: The study should capture the influence of both internal and external factors on stress.
– Which self-report scales should be used to annotate stress data?
– What prior information about the participant’s behavior could affect their stress response and recovery?
– What details about the participant’s activities prior to the experiment should be collected?

We carried out a comprehensive analysis of the data and conclusions from clinical and empirical research, based on the questions presented in this framework. The knowledge gained from this literature review, combined with our own judgments, has been organized into five groupings, each of which is described in more detail in the subsections that follow. The particular requirements generated from these categories are provided at the conclusion of each subsection. "REQ-" is used to identify these criteria, followed by the appropriate number.

A. Demographics and Size of Sample Population of the Dataset

Research indicates that factors such as age [7], [33] and gender [10], [34] significantly influence stress responses. Consequently, the sample population should encompass equal proportions of various demographic groups, particularly concerning age and gender. Collaboration with a statistician [35] is recommended to estimate an appropriate sample size that is representative of the population. Prior information related to the stress-related use case and relevant statistical data should be gathered; for example, assisted daily living may require a ratio of individuals suffering from depression or anxiety disorders compared to healthy individuals within a larger community. Once the statistical parameters are established, the sample size for the experiment can be determined using Cochran’s sample size formula [36] or other available sample size calculation tools.

REQ-1: The sample population size should be representative of the target population and should encompass various demographic factors.

B. Characteristics of the Stress Stimuli

Stress responses are known to be stimuli-specific [37]. Thus, the selected stress stimuli should be pertinent to everyday activities or tailored to the specific use case for which the reference dataset is being collected. An effective stress stimulus must be capable of eliciting HPA (Hypothalamus-Pituitary-Adrenal) responses. Four characteristics of stress stimuli that are essential for effectively eliciting stress responses in individuals include novelty, uncontrollability, unpredictability, and socio-evaluative threat [10]. A notable example of effective stress stimuli is the Trier Social Stress Test [38], which combines cognitive testing with public speaking.

REQ-2: The stress stimuli employed should be relevant and effective.

C. Multiple Reliable Modalities

As described in Section II-A, there is no consensus regarding a definitive ground truth for stress data annotation. However, due to the varying expressions of stress among individuals, it is crucial to capture multiple modalities. As mentioned in Section II-A, heart rate variability (HRV) [14], [15] and skin conductance (SC) [16] are correlated with cortisol levels. Therefore, the datasets should include a minimum of two reliable modalities, specifically the electrical responses of the heart (electrocardiogram, ECG) and the skin (electrodermal activity, EDA).

REQ-3: Multiple modalities containing complementary information should be recorded, specifically including the electrical responses of the heart (ECG) and the skin (EDA).

D. Self-Reported Information

Self-reports are employed by psychologists to capture instantaneous feelings experienced following a stress stimulus. Although self-reports are highly subjective, unreliable, and difficult to verify, they serve as useful tools for gaining insights into psychological responses to stimuli. It is advisable to utilize validated instantaneous self-report scales, such as the Perceived Stress Scale [20]. Prior to the experiment, information regarding the cognitive, sensory, and motor skills of the individual should be documented to relate perceived stress levels to the complexity of the stimulus.

Interpersonal variability in stress responses is influenced by various internal factors, including age [33], gender [10], [34], personality type [22], and medical conditions [24], as well as external factors such as exercise, dietary salt intake [25], and biological rhythms affecting sleep and shift work [39]. Medical conditions, such as chronic depression, can suppress physiological stress responses [40]. According to the American Psychological Association [41], significant natural causes of stress include an individual's socio-economic status, work conditions, and social relationships. Information regarding these factors that could potentially affect stress responses should be collected from participants before the experiment. This is crucial for analyzing and interpreting the interpersonal differences evident in the collected stress response data. Self-reports can also be utilized for this purpose.

REQ-4: Multiple self-reporting methods should be employed to record internal and external factors that influence stress response.

E. Data Acquisition Devices and Software

The quality of sensors significantly impacts the accuracy of the stress detection process. While achieving gold-standard performance from devices suitable for daily use may be challenging, it is advisable to utilize clinically validated sensing devices to mitigate their adverse effects on the performance of the stress detection algorithm. Device configuration should account for the characteristics of the physiological signals being recorded. For instance, signals such as skin conductance exhibit latency on the order of seconds between the stress stimulus and the response [42]. Therefore, careful attention should be given to device calibration and time synchronization to enable the integration of data from multiple sensors.

Given the inherent noise in any sensor, it is essential to specify the noise characteristics of the devices utilized. This information is critical for selecting appropriate noise filtering methods, improving the interpretation of stress responses, and enhancing the reproducibility of experiments. Stress detection systems that consider these noise characteristics will be better equipped to generalize across data obtained from sensors not included in the training dataset. Additionally, stress response measurements often need to be taken over multiple days to investigate intrapersonal variations. To minimize inconsistencies in measurements conducted on different days, it is necessary to establish a device setup protocol that outlines calibration parameters. This will help reduce errors, save time, and facilitate the extension of the reference dataset in the future by consistently calibrating the sensors to align with the existing dataset characteristics.

REQ-5: Clinically validated data acquisition devices should be utilized, and their noise characteristics should be documented; device calibration information should be included in a device setup protocol.

IV. EVALUATION RESULTS

It was observed that the datasets did not always provide all the information that was needed to understand how they were created. For example, participants' instantaneous self-reports are not included in the Drivedb dataset. As a result, only dataset information that was available to the general public was considered during the evaluation. Table II provides a summary of the evaluation criteria and the associated evaluation outcomes. The following are the evaluation's main findings:

- **Population**: The Distracted Driving dataset features an equal proportion of male and female subjects and includes two distinct age cohorts. However, the other datasets fail to meet these criteria.
- **Stimuli**: The stress stimuli in the Drivedb, SWELL-KW, Distracted Driving, and WESAD datasets are relevant to everyday living. Conversely, the stimuli in the SUS dataset do not reflect daily life scenarios.
- **Modalities**: The Drivedb, SWELL-KW, Distracted Driving, and WESAD datasets incorporate multiple modalities. However, the Distracted Driving and SUS datasets lack one or both of the reliable modalities, specifically ECG and EDA.
- **Self-reports**: The SWELL-KW, Distracted Driving, and WESAD datasets provide subjective reports regarding personality and perceived stress. WESAD also contains information on smoking, caffeine intake, and physical activity prior to the experiment. However, participants in the SWELL-KW dataset were instructed to refrain from smoking or consuming coffee three hours before the experiment. The Drivedb and SUS datasets do not provide relevant self-report data.
- **Data Acquisition Devices**: The Drivedb [43], SWELL-KW [44], WESAD, and Distracted Driving [31] datasets specify the devices employed for data acquisition. The SUS dataset indicates that microphones and telephone receivers were utilized, but does not furnish further details. None of the datasets provide information regarding the noise characteristics or calibration of the devices used.

In conclusion, it is evident that none of the datasets fully satisfy the evaluation criteria, and as such, they do not meet the requirements to be classified as a reference dataset.

V. CONCLUSION

Stress is a highly complex and subjective phenomenon. In order to build reliable, automatic stress detection systems, it is necessary to use a reference dataset that captures all the important aspects of the stress phenomenon. This paper specifies the requirements that should be fulfilled by a reference dataset for multimodal human stress detection. The requirements are derived by reviewing clinical and technical literature and are based on clinical practices and empirical evidence. A set of evaluation criteria is defined based on the requirements, and the existing publicly available human stress datasets are evaluated against it. It was found that none of these datasets fulfil all the requirements to qualify as a reference dataset. Therefore, future efforts should aim at establishing such a reference dataset, by considering the proposed requirements. The requirements listed in this paper are not exhaustive. Future work could also focus on extending the requirements by examining, for example, use case specific as well as chronic stress-related aspects.

REFERENCES

1. American Psychological Association, **Stress and Health Disparities: Contexts, Mechanisms, and Interventions Among Racial/Ethnic Minority and Low-Socioeconomic Status Populations**, Technical Report, 2018. Available: http://www.apa.org/pi/health-disparities/resources/stress-report.pdf
2. World Health Organization, **Comprehensive Mental Health Action Plan 2013-2020**, 2016. Available: https://www.who.int/mental_health/publications/action_plan/en/
3. ——, **Guidelines for the Management of Conditions Specifically Related to Stress: mhGAP Intervention Guide Module**, 2013. Available: http://www.ncbi.nlm.nih.gov/pubmed/24649518
4. EU-OSHA, **Calculating the Cost of Work-Related Stress and Psychosocial Risks**, Technical Report, 2014. Available: https://www.healthy-workplaces.eu/es/tools-and-resources/publications

5. H. Selye, "Stress and the General Adaptation Syndrome," **British Medical Journal**, vol. 1, no. 4667, pp. 1383–1392, 1950. Available: http://www.ncbi.nlm.nih.gov/pubmed/6649565
6. B. M. Kudielka, H. Hellhammer, and C. Kirschbaum, "Ten Years of Research with the Trier Social Stress Test—Revisited," **Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior**, pp. 56–83, 2007.
7. P. Foley and C. Kirschbaum, "Human Hypothalamus-Pituitary-Adrenal Axis Responses to Acute Psychosocial Stress in Laboratory Settings," **Neuroscience and Biobehavioral Reviews**, vol. 35, no. 1, pp. 91–96, 2010. Available: http://dx.doi.org/10.1016/j.neubiorev.2010.01.010
8. Z. Djuric, C. E. Bird, A. Furumoto-Dawson, G. H. Rauscher, M. T. Ruffin, R. P. Stowe, K. L. Tucker, and C. M. Masi, "Biomarkers of Psychological Stress in Health Disparities Research," **The Open Biomarkers Journal**, pp. 7–19, 2010.
9. S. H. Golden, G. S. Wand, S. Malhotra, I. Kamel, and K. Horton, "Reliability of Hypothalamic-Pituitary-Adrenal Axis Assessment Methods for Use in Population-Based Studies," **European Journal of Epidemiology**, vol. 26, no. 7, pp. 511–525, 2011. Available: https://www.ncbi.nlm.nih.gov/pubmed/21533585
10. S. J. Lupien, F. Maheu, M. Tu, A. Fiocco, and T. E. Schramek, "The Effects of Stress and Stress Hormones on Human Cognition: Implications for the Field of Brain and Cognition," **Brain and Cognition**, vol. 65, no. 3, pp. 209–237, 2007. Available: http://www.sciencedirect.com/science/article/pii/S0278262607000322
11. A. Steptoe and M. Kivimäki, "Stress and Cardiovascular Disease: An Update on Current Knowledge," **Annual Review of Public Health**, vol. 34, pp. 337–354, 2013. Available: http://www.annualreviews.org/doi/10.1146/annurev-publhealth-031912-114452
12. N. Sharma and T. Gedeon, "Objective Measures, Sensors, and Computational Techniques for Stress Recognition and Classification: A Survey," **Computer Methods and Programs in Biomedicine**, vol. 108, no. 3, pp. 1287–1301, 2012. Available: http://www.sciencedirect.com/science/article/pii/S0169260712001770
13. A. Alberdi, A. Aztiria, and A. Basarab, "Towards an Automatic Early Stress Recognition System for Office Environments Based on Multimodal Measurements: A Review," **Journal of Biomedical Informatics**, vol. 59, pp. 49–75, 2016. Available: http://www.sciencedirect.com/science/article/pii/S1532046415002750
14. B. von Dawans, C. Kirschbaum, and M. Heinrichs, "The Trier Social Stress Test for Groups (TSST-G): A New Research Tool for Controlled Simultaneous Social Stress Exposure in a Group Format," **Psychoneuroendocrinology**, vol. 36, no. 4, pp. 514–522, 2011. Available: http://dx.doi.org/10.1016/j.psyneuen.2010.08.004
15. S. S. Dickerson and M. E. Kemeny, "Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research," **Psychological Bulletin**, vol. 130, no. 3, pp. 355–391, 2004.
16. S. D. Vanderark and D. Ely, "Cortisol, Biochemical, and Galvanic Skin Responses to Music Stimuli of Different Preference Values by College Students in Biology and Music," **Perceptual and Motor Skills**, vol. 77, no. 1, pp. 227–234, 1993. Available: https://doi.org/10.2466/pms.1993.77.1.227
17. E. Granholm and S. R. Steinhauer, "Pupillometric Measures of Cognitive and Emotional Processes," **International Journal of Psychophysiology**, vol. 52, no. 1, pp. 1–6, 2004.
18. K. A. Herborn, J. L. Graves, P. Jerem, N. P. Evans, R. Nager, D. J. McCafferty, and D. E. McKeegan, "Skin Temperature Reveals the Intensity of Acute Stress," **Physiology and Behavior**, vol. 152, pp. 225–230, 2015. Available: http://dx.doi.org/10.1016/j.physbeh.2015.09.032

19. D. Shastri, M. Papadakis, P. Tsiamyrtzis, B. Bass, and I. Pavlidis, "Perinasal Imaging of Physiological Stress and Its Affective Potential," **IEEE Transactions on Affective Computing**, vol. 3, no. 3, pp. 366–378, 2012.
20. S. Cohen, T. Kamarck, and R. Mermelstein, "A Global Measure of Perceived Stress," pp. 385–396, 1983.
21. L. Laux, **Das State-Trait-Angstinventar (STAI): Theoretische Grundlagen und Handanweisung,** 1981.
22. E. Childs, T. L. White, and H. De Wit, "Personality traits modulate emotional and physiological responses to stress," **Behavioural Pharmacology,** vol. 25, no. 5-6, pp. 493–502, 2014.
23. U. Rimmele, R. Seiler, B. Marti, P. H. Wirtz, U. Ehlert, and M. Heinrichs, "The level of physical activity affects adrenal and cardiovascular reactivity to psychosocial stress," **Psychoneuroendocrinology,** vol. 34, no. 2, pp. 190–198, 2009.
24. D. H. Hellhammer, S. Wust, and B. M. Kudielka, "Salivary cortisol as a biomarker in stress research," **Psychoneuroendocrinology,** vol. 34, no. 2, pp. 163–171, 2009.
25. D. S. Goldstein and I. J. Kopin, "Evolution of concepts of stress," **Stress,** vol. 10, no. 2, pp. 109–120, 2007.
26. S. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, and W. Kraaij, "The SWELL Knowledge Work Dataset for Stress and User Modeling Research," in **Proceedings of the 16th International Conference on Multimodal Interaction,** ser. ICMI '14. New York, NY, USA: ACM, 2014, pp. 291–298. [Online]. Available: http://doi.acm.org/10.1145/2663204.2663257
27. J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," **IEEE Transactions on Intelligent Transportation Systems,** vol. 6, no. 2, pp. 156–166, 2005.
28. L.-l. Chen, Y. Zhao, P.-f. Ye, J. Zhang, and J.-z. Zou, "Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers," **Expert Systems with Applications,** pp. –, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417417300477
29. N. El Haouij, J.-M. Poggi, E. Ghozi Raja, S. Sevestre-Ghalila, and M. Jaïdane, "Random forest-based approach for physiological functional variable selection: Towards driver's stress level classification," 2017. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01426752
30. H. J. M. Steeneken and J. H. L. Hansen, "Speech under stress conditions: Overview of the effect on speech production and on system performance," in **1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258),** vol. 4, 1999, pp. 2079–2082 vol.4.
31. S. Taamneh, P. Tsiamyrtzis, M. Dcosta, P. Buddharaju, A. Khatri, M. Manser, T. Ferris, R. Wunderlich, and I. Pavlidis, "A multimodal dataset for various forms of distracted driving," **Scientific Data,** vol. 4, p. 170110, 2017. [Online]. Available: http://www.nature.com/articles/sdata2017110
32. P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, and K. Van Laerhoven, "Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection," in **Proceedings of the 20th ACM International Conference on Multimodal Interaction,** ser. ICMI '18. New York, NY, USA: ACM, 2018, pp. 400–408. [Online]. Available: http://doi.acm.org/10.1145/3242969.3242985
33. B. M. Kudielka, A. Buske-Kirschbaum, D. H. Hellhammer, and C. Kirschbaum, "HPA axis responses to laboratory psychosocial stress in healthy elderly adults, younger adults, and children: Impact of age and gender," **Psychoneuroendocrinology,** vol. 29, no. 1, pp. 83–98, 2004.
34. B. M. Kudielka and C. Kirschbaum, "Sex differences in HPA axis responses to stress: A review," **Biological Psychology,** vol. 69, no. 1 SPEC. ISS., pp. 113–132, 2005.

35. B. Rohrig, J.-B. du Prel, D. Wachtlin, R. Kwiecien, and M. Blettner, "Sample Size Calculation in Clinical Trials," **Deutsches Aerzteblatt Online,** vol. 107, pp. 552–556, 2010. [Online]. Available: https://www.aerzteblatt.de/10.3238/arztebl.2010.0552
36. W. G. Cochran, **Sampling techniques,** 1963.
37. K. Pacak and M. Palkovits, "Stressor specificity of central neuroendocrine responses: Implications for stress-related disorders," **Endocrine Reviews,** vol. 22, no. 4, pp. 502–548, 2001.
38. C. Kirschbaum, K.-M. Pirke, and D. Hellhammer, "The 'Trier Social Stress Test'—a tool for investigating psychobiological stress responses in a laboratory setting," **Neuropsychobiology,** vol. 28, pp. 76–81, 1993.
39. S. J. Lupien, B. S. McEwen, M. R. Gunnar, and C. Heim, "Effects of stress throughout the lifespan on the brain, behaviour and cognition," **Nature Reviews Neuroscience,** vol. 10, 2009. [Online]. Available: http://dx.doi.org/10.1038/nrn2639
40. H. M. Burke, M. C. Davis, C. Otte, and D. C. Mohr, "Depression and cortisol responses to psychological stress: A meta-analysis," **Psychoneuroendocrinology,** vol. 30, no. 9, pp. 846–856, 2005.
41. American Psychological Association, "Stress in America 2011," https://www.apa.org/news/press/releases/stress/2011/, 2012.
42. C. Setz, B. Arnrich, J. Schumm, R. L. Marca, G. Troster, and U. Ehlert, "Discriminating Stress from Cognitive Load Using a Wearable EDA Device," **IEEE Transactions on Information Technology in Biomedicine,** vol. 14, no. 2, pp. 410–417, 2010.
43. J. A. Healey, "Wearable and automotive systems for affect recognition from physiology," **Ph.D. dissertation,** 2000.
44. S. J. Koldijk, "Context-Aware Support for Stress Self-Management: From Theory to Practice," **Radboud Repository,** p. 144, 2016. [Online]. Available: http://repository.ubn.ru.nl/bitstream/handle/2066/93611/93611.pdf?sequence=1.