

Article

Not peer-reviewed version

Wearable Technologies for Electrodermal and Cardiac Activity measurements: a Comparison between Fitbit Sense, Empatica E4 and Shimmer GSR3+

[Vincenzo Ronca](#)*, Ana Martinez-Levy, [Alessia Vozzi](#), [Andrea Giorgi](#), [Pietro Aricò](#), [Rossella Capotorto](#), [Gianluca Borghini](#), [Gianluca Di Flumeri](#), [Fabio Babiloni](#)

Posted Date: 6 June 2023

doi: 10.20944/preprints202306.0378.v1

Keywords: Electrodermal activity; Photoplethysmography; Consumer wearables



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Wearable Technologies for Electrodermal and Cardiac Activity Measurements: A Comparison between Fitbit Sense, Empatica E4 and Shimmer GSR3+

Vincenzo Ronca ^{1,2,*}, Ana.C Martinez-Levy ^{2,3}, Alessia Vozzi ^{2,4}, Andrea Giorgi ^{2,4}, Pietro Arico ^{1,2}, Rossella Capotorto ^{1,3}, Gianluca Borghini ^{2,3}, Fabio Babiloni ^{2,3,5} and Gianluca Di Flumeri ^{2,3}

¹ Department of Computer, Control, and Management Engineering, Sapienza University of Rome, Italy; [vincenzo.ronca; pietro.arico, capotorto.1843967@studenti.]@uniroma1.it

² BrainSigns srl, Rome, Italy; vincenzo.ronca; alessia.vozzi; ana.martinez; andrea.giorgi; pietro.arico; gianluca.borghini; fabio.babiloni; gianluca.diflumeri@brainsigns.com

³ Department of Molecular Medicine, Sapienza University of Rome, Italy; [ana.martinez; gianluca.borghini; fabio.babiloni; gianluca.diflumeri]@uniroma1.it

⁴ Department of Anatomical, Histological, Forensic and Orthopaedic Sciences, Sapienza University of Rome, Italy; [alessia.vozzi; andrea.giorgi]@uniroma1.it

⁵ College of Computer Science and Technology, Hangzhou Dianzi University, China; fabio.babiloni@uniroma1.it

* Correspondence: vincenzo.ronca@uniroma1.it

Abstract: The capability of measuring specific neurophysiological and autonomic parameters plays a crucial role in the objective evaluation of the human's mental and emotional states. These human aspects are commonly known in scientific literature to be involved in a wide range of processes, such as the stress and arousal. These aspects represent a relevant factor especially in real and operational environment. Neurophysiological autonomic parameters, such as the Electrodermal Activity (EDA) and Photoplethysmographic data (PPG), have been usually investigated through research-graded devices, therefore resulting in a high degree of invasiveness, which could negatively interfere with the monitored user's activity. For such a reason, in the last decade the recent consumer-grade wearable devices, usually designed for fitness tracking purposes, are receiving an increasing attention from the scientific community, being characterized by higher comfort, easiness of use and, therefore, by higher compatibility with daily life environments. The present preliminary study aimed at assessing the reliability of a consumer wearable device, i.e., the Fitbit Sense, with respect to a research-graded wearable, i.e., the Empatica E4 wristband, and a laboratory device, i.e., the Shimmer GSR3+. The EDA and PPG were collected among 12 participants while performing multiple resting conditions. The results demonstrated that the EDA and PPG-derived features computed through the wearable and research devices were positively and significantly correlated, while the reliability of the consumer device resulted to be significantly lower.

Keywords: electrodermal activity; photoplethysmography; consumer wearables

Introduction

Nowadays, wearable technologies are constantly growing in terms of popularity, and wearable devices are increasing their employment in monitoring fitness and health related parameters [1,2]. Wearable technology relies on tiny, accurate, and cost-efficient sensors, and it can be used to track people's movement, collect biometric signals, and analyze daily activities [2–5]. One of the most popular products in terms of consumer-grade wearables corresponds to the wristbands. Such a technology is fully compatible with real-world applications, supported by a very low grade of

invasiveness, which permits to do not negatively interfere with the users' activities, by a consistent battery life and wireless connectivity to fetch and stream data [6–8].

Therefore, also the scientific community largely started to investigate the reliability of such wearable technologies in evaluating neurophysiological and autonomic parameters. In this context, several scientific works demonstrated the reliability of wearable devices, such as the Empatica E4, in collecting the Electrodermal Activity (EDA) and the Photoplethysmographic (PPG) signal [9–12]. As demonstrated by different previous works [4,10,13–15], the EDA corresponds to a consistent biomarker of the stress level, while the PPG-derived parameters, such as the Heart Rate (HR) and the Heart Rate Variability (HRV), were significantly correlated to the mental workload and emotional states variations [16–21]. In particular, Ragot and colleagues [19] successfully adopted the Empatica E4 wristband to measure physiological response in an emotion recognition task. Based on these evidence, wearable devices were also used to assess different mental states, such as the study conducted by Setz and colleagues [14], in which the mental drowsiness was detected during a driving simulation task through the HRV estimation. In this context, other studies demonstrated the capability of wearable devices in detecting crucial mental states in operational environments, such as the prediction of the stress and the mental workload in police academy training [22] and during professional physical activity [23] through the PPG, respiration, and motion activities data analyses. More recently, Giorgi and colleagues [9] positively assessed the reliability of the Empatica E4 in detecting EDA and cardiac activity with the respect to regular laboratory equipment, considered as the gold-standard in terms of accuracy.

To summarize, wearable technologies for research, e.g., the Empatica E4, result to be reliable in evaluating neurophysiological and autonomic parameters. However, they are still characterized by a relatively non-negligible economic cost (i.e., between 1 and 2 k€/k\$), that make them less compatible with a massive use such as large-scale studies involving multiple participants at the same time. Therefore, a further step could be taken in the sense of the consumer wearable technologies. In this regard, the last technological development allowed to launch on consumer markets devices able to capture autonomic parameters, such as the PPG and EDA. The Fitbit Sense resulted to be one of the first consumer wearables capable of detecting such signals.

The reliability of such a device in detecting autonomic and health-related parameters with the respect to research wearable and laboratory devices plays a crucial role, since its relevantly high economic accessibility. Nevertheless, there is still a consistent lack of scientific works to investigate the reliability of a consumer wearable compared to the accuracy characterizing research wearables and laboratory equipment. In particular, at the best of our knowledge, there is not any study so far aiming at investigating the reliability of a consumer-grade wearable smartband, i.e., the Fitbit one, in measuring neurophysiological autonomic data, namely PPG and EDA, with respect to research-grade devices. In this regard, it has to be noted that several works related to the employment of the most recent Apple Watch generations (Apple Inc., Cupertino, California - USA) in medical graded experimental protocol with promising results [24–26]. In particular, such consumer-grade wearable was employed for investigating its capabilities in detecting cardiac related pathologies, such as the atrial fibrillation ([27]), and its capability in assessing the noise level within the intensive care unit as proposed by Scquizzato and colleagues ([28]). In this context, it has to be observed that the reliability of the above-mentioned consumer-grade wearable, i.e., the Apple Watch, was not investigated in terms of collecting neurophysiological and autonomic data, such as the PPG and EDA.

Therefore, the aim of the present study consisted in assessing the reliability of the Fitbit Sense, in estimating EDA and PPG-derived parameters with the respect to a research wearable and a laboratory device, i.e., the Empatica E4 and the Shimmer GSR3+ respectively. Because of the difficulty in recruiting participants, given the fact that the complete experimental protocol was designed and conducted within the COVID-19 pandemic restrictions, the sample size includes in the experimental protocol was low (12 participants). Therefore, the presented work it has to be intended as a preliminary study, however it has to be considered in any case relevant because of the novelty of the topic, being the first study investigating a consumer-grade device's abilities to collect reliable data for scientific purposes.

Methods

Twelve (12) participants, young students (18 – 26 years old, 6 males and 6 females) from the Sapienza University of Rome, with normal or corrected-to-normal vision, were recruited on a voluntary basis. Informed consent was obtained from each participant after explanation of the study. The experiment was conducted following the principles outlined in the Declaration of Helsinki of 1975, as revised in 2000 and was approved by the Sapienza University of Rome Ethical Committee in Charge for the Department of Molecular Medicine (protocol number: 2104/2021, approved on 5 April 2021). To respect the privacy of participants, only aggregate results were reported, and any results based on single identity analysis was presented.

The experimental procedure was designed as simplest as possible, in order to allow a consistent reliability assessment of the three kinds of devices involved, i.e., the consumer wearable Fitbit Sense (Google Inc., Mountain View – USA), the research wearable Empatica E4 (Empatica, Milan – Italy), and the laboratory equipment Shimmer GSR3+ (Shimmer Sensing, Dublin – Ireland) (Figure 1). The experimental tasks corresponded to three consecutive 90-seconds-long resting conditions, that was a minimum requirement for collecting valid data from the Fitbit Sense. On a 90-seconds-long recording session, the Fitbit provides its estimation of EDA responses and HR each 30 seconds. In this way, for each participant, 9 data points were available. All the participants involved in the experimental protocol were instructed to do not perform any activities during the neurophysiological and autonomic data collection along the entire experiment.



Figure 1. Overview of the three different devices type investigated in the present work.

PPG signal recording and analysis

The PPG signal was collected simultaneously through the Empatica E4 and the Shimmer GSR3+. The first one was placed on the right participants' wrist, the Empatica E4 on the left participants' wrist, while the Shimmer GSR3+ PPG sensor was placed on the participants' thumb (Figure 2). Firstly, the PPG signals were filtered using a 5th Butterworth band-pass filter (1-15 Hz) in order to reject the continuous component and the high-frequency interferences, such as that related to the main power source. Subsequently, a signal preprocessing chain was applied to identify and remove movement-based artifacts. In particular:

- Signal filtering between 1 and 15 Hz.
- All the negative signal segments were made positive by checking their skewness.
- Signal windowing according to a 2-seconds length. This was based on the hypothesis for which along 2 seconds at least one physiological R-peak must occur.
- The local maxima of each 2 seconds-long window were considered.

- Identification of artefactual portions through a threshold method. This step was performed through specific sub-steps:
 - Each 2 seconds-long window was labelled as artefact if the corresponding local maximum was exceeding the threshold of $2 \times \text{median amplitude}$ among all the local maxima. This corresponded to an empiric and reasonable assumption for which if a signal peak exceeds twice the median among the other local maxima it corresponds to a non-physiological peak.
 - Each 2 seconds-long window was labelled as artefact if the corresponding local maximum did not exceed a minimum threshold set around 0. This corresponded to an empiric and reasonable assumption for which if a signal peak does not exceed a minimum threshold around 0 it corresponds to a non-physiological peak or, more likely, it corresponds to a 2 seconds-long windows in which no signal was recorded.
- Identification of artefactual portions through a threshold method and the first signal derivative. This step was based on the assumption for which specific signal artefacts, e.g., the signal discontinuities, are characterized by an amplitude within a physiological range and, therefore, they are not identifiable via the previous threshold method. In this regard, the signal first derivative was considered to identify non-physiological discontinuities. This step was performed through specific sub-steps:
 - The signal first derivative was standardized and squared.
 - The convolution between the signal first derivative standardized and squared and a 0.15 seconds-long-time window.
 - Each 2 seconds-long window was labelled as artefact if the first signal derivative standardized and squared was exceeding the threshold of 9. This value ensured that the labelled signal portion corresponded to a physiological outlier for the 99.7%, corresponding to $3 \times \sigma$ according to the cumulative distribution. This corresponded to an empiric and reasonable assumption for which if the signal first derivative exceeds such a threshold it corresponds to a non-physiological signal discontinuity.



Figure 2. Overview of the experimental settings.

Subsequently, the Pan-Tompkins [9,29,30] algorithm was applied exclusively to the non-artefactual PPG signal portions for the Inter-Beat Interval (IBI) estimation. Finally, the HR parameter

was estimated from the IBI during the task length for each participant, and averaged on 30-seconds-long windows accordingly to the measures provided by Fitbit device.

On the other side, since it is not possible to access the Fitbit raw data, the HR estimation computed by the Fitbit software (and related application) has been considered during the same task.

EDA recording and analysis

Similarly to the PPG signal, the EDA was collected through all the three devices categories. Actually, also in this case the Fitbit device does not provide access to the raw data, so the Shimmer and Empatica data have been recorded and processed as follows, while the Fitbit output has been considered for the analysis as described in the next sub-paragraph. The sampling frequency of the Shimmer3 GSR+ unit laboratory device corresponded to 64 Hz while the sampling frequency of the Empatica E4 was 4 Hz. Shimmer sensors were placed on the participant's no-dominant hand on the second and third fingers. Concerning the Empatica E4, the two electrodes were placed on the bottom part of the wrist, while the Fitbit Sense required the physical contact between the participants' hand palm and the upper part of the device in order to collect EDA. Regarding the signal processing related to the EDA collected through the Shimmer GSR3+ and the Empatica E4, the signal was firstly low-pass filtered with a cut-off frequency of 1 Hz and then processed by using the Ledalab suite [31], a specific open-source toolbox implemented within the MATLAB (MathWorks, Natick, Massachusetts) environment for EDA processing. The continuous decomposition analysis [32] was applied in order to estimate the tonic (SCL) and the phasic (SCR) EDA components [33]. The SCL corresponds to the slow-changing component of the EDA signal, consistently related with the arousal and stress levels. On the contrary, the SCR is the fast-changing component of the EDA signal usually related to single stimuli reactions. Finally, only the SCL was analyzed accordingly with the previous scientific evidence for which this parameter is strictly correlated with specific and relevant mental states, such as the stress [34–36]. The SCL estimated through the research wearable and laboratory equipment was computed during the task length and averaged on 30-seconds-long windows accordingly to the measures provided by Fitbit device.

Concerning the Fitbit Sense, such a device does not allow the continuous EDA raw data collection and access. The consumer wearable exclusively allowed to collect the EDA responses, which consists in a parameter related to the Skin Conductance Level (SCL) peaks overtime. Therefore, the "SCL responses" measure provided each 30 seconds was considered.

Statistical analysis

Given the main objective of the presented study, different type of correlation analyses were performed in order to compare the reliability of the consumer wearable with the respect of the research wearable and laboratory equipment. In particular, the Pearson's correlation analysis was performed to assess the similarity between the measurements collected through the consumer wearables, research wearables, and the laboratory equipment. Subsequently, the repeated measure correlation analysis [37] was additionally performed to estimate the reliability of the parameters estimated by the consumer and research wearable devices with respect to the laboratory one both at single - participant level and on the entire group, i.e., by taking into account intra- and inter-individual variability at the same time.

Results

Regarding the functional comparison between the consumer and research wearable devices and the laboratory equipment, the Pearson's correlation analysis revealed a consistent difference between the measurements collected through the consumer wearable and the research wearable, e.g., the Fitbit Sense and the Empatica E4 respectively, with the respect to the laboratory equipment, e.g., the Shimmer GSR3+, considered as the gold standard. In fact, the Pearson's correlation analysis performed on the PPG-derived parameters evaluations revealed a positive and significant correlation between the HR computed by the signal collected through the Empatica E4 and the Shimmer GSR3+

for the large majority of the participants (9/12) (Figure 2). No significant correlation for any participants was observed between the HR estimated through the Fitbit Sense and the Shimmer GSR3+.

Similarly, Figure 3 shows how the SCL parameters derived by the EDA collected through the Empatica E4 and the Shimmer GSR3+ were positively (all $R > 0.4$) and significantly (all $p < 0.05$) correlated for nearly all the participants (10/12); while the SCL parameters derived by the EDA gathered through the Fitbit Sense and the Shimmer GSR3+ revealed a positive (all $R > 0.53$) and significant (all $p < 0.05$) correlation only for 2 participants among 12.

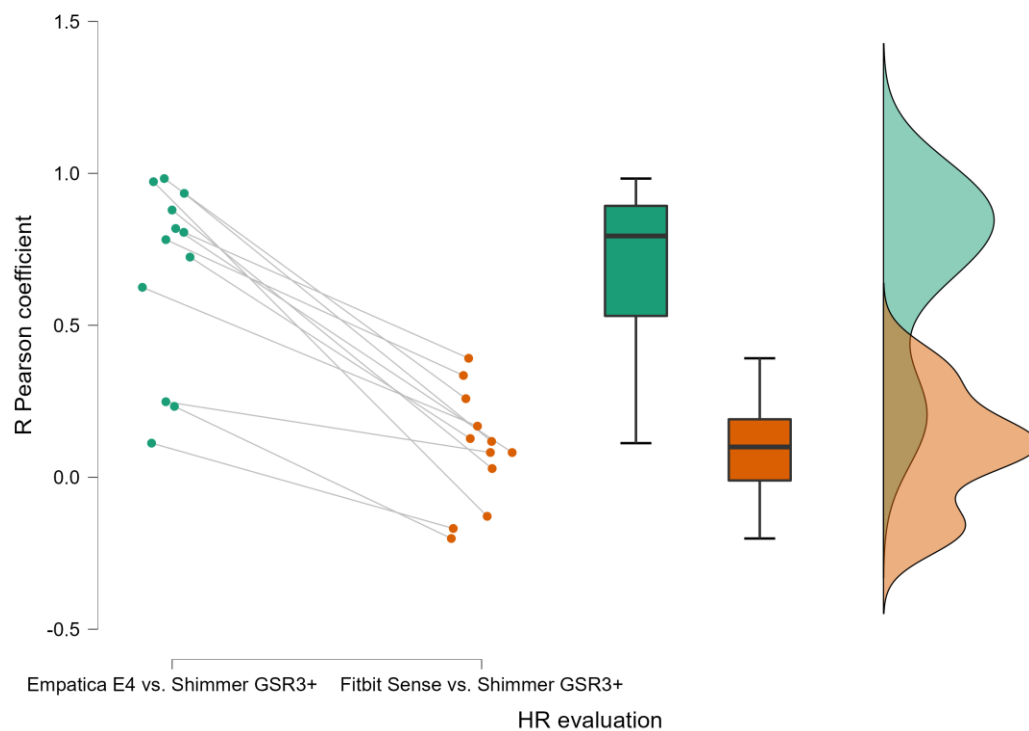


Figure 3. Person's correlations between the Empatica E4 and Shimmer (left bar) and the Fitbit Sense and the Shimmer (right bar) in terms of Heart Rate estimation.

Regarding the devices reliability assessment at single – participant level between the Empatica E4 and the Shimmer GSR3+, the repeated measure correlation analysis revealed a positive and significant correlation for both the SCL and HR parameters evaluations (SCL: $R = 0.681$, $p < 10^{-34}$; HR: $R = 0.596$, $p < 10^{-33}$) (Figures 4 and 5).

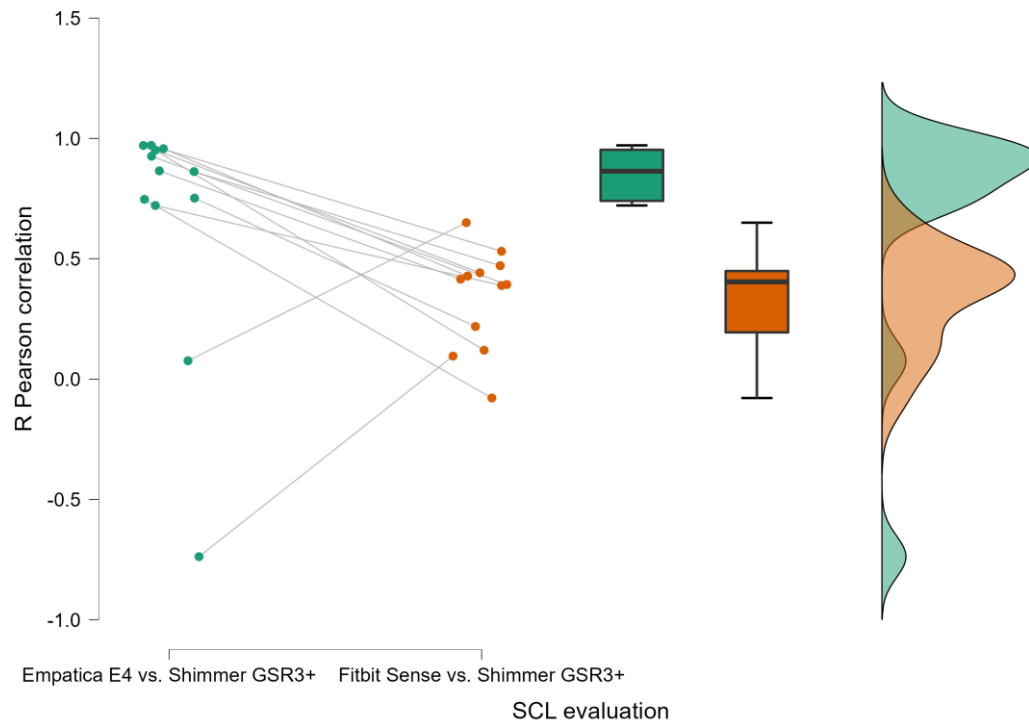


Figure 4. Person's correlations between the Empatica E4 and Shimmer GSR3+ (left bar) and the Fitbit Sense and the Shimmer GSR3+ (right bar) in terms of Skin Conductance Level estimation.

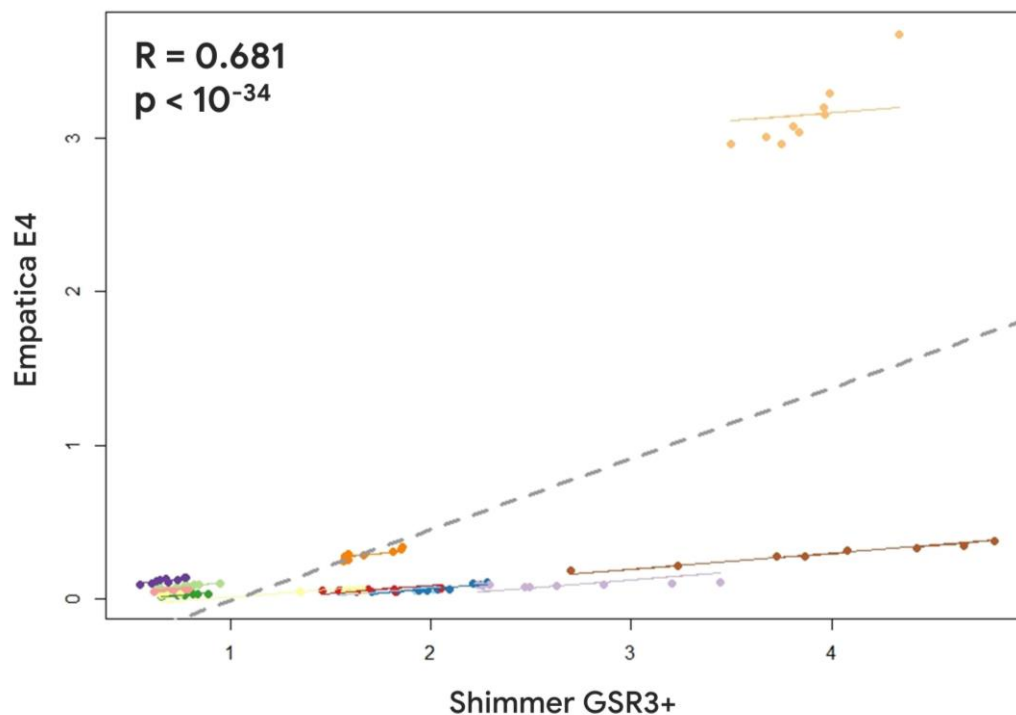


Figure 5. Repeated measure correlation between the Skin Conductance Level estimated through the Empatica E4 and the Shimmer GSR3+.

While the comparison between the Fitbit Sense and the Shimmer GSR3+ revealed a positive and significant correlation only for the SCL estimations ($R = 0.412$, $p < 10^{-25}$) (Figure 6). In fact, any positive and significant correlation were observed between the HR estimations provided by the Fitbit Sense, and one evaluated by the Shimmer GSR3+ ($R = -0.121$; $p = 0.1$) (Figure 7).

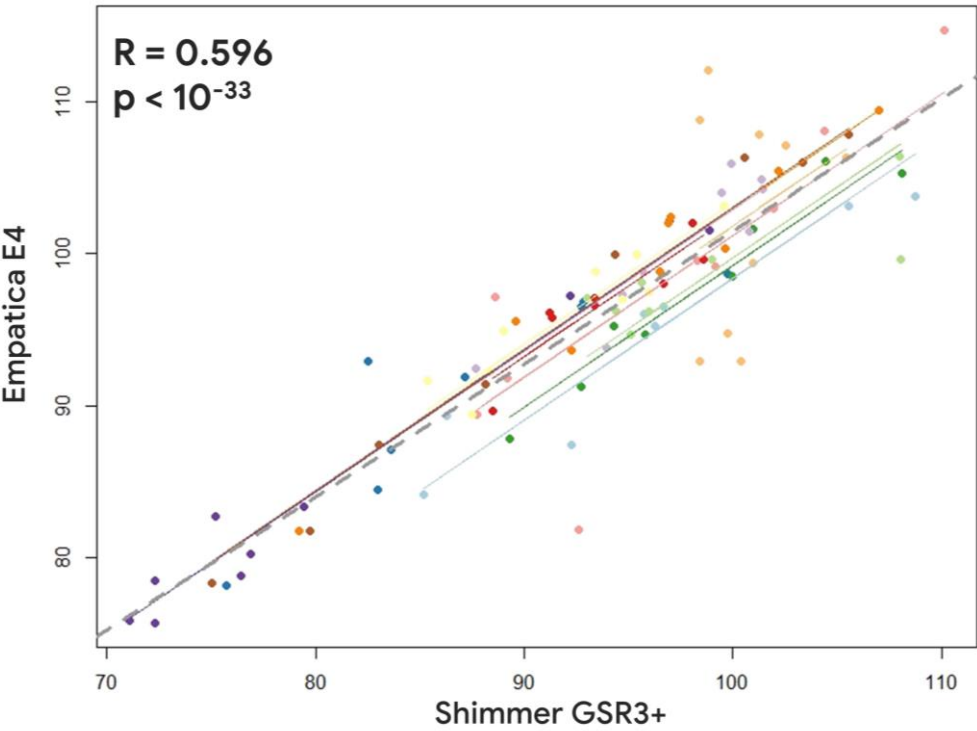


Figure 6. Repeated measure correlation between the Heart Rate estimated through the Empatica E4 and the Shimmer GSR3+.

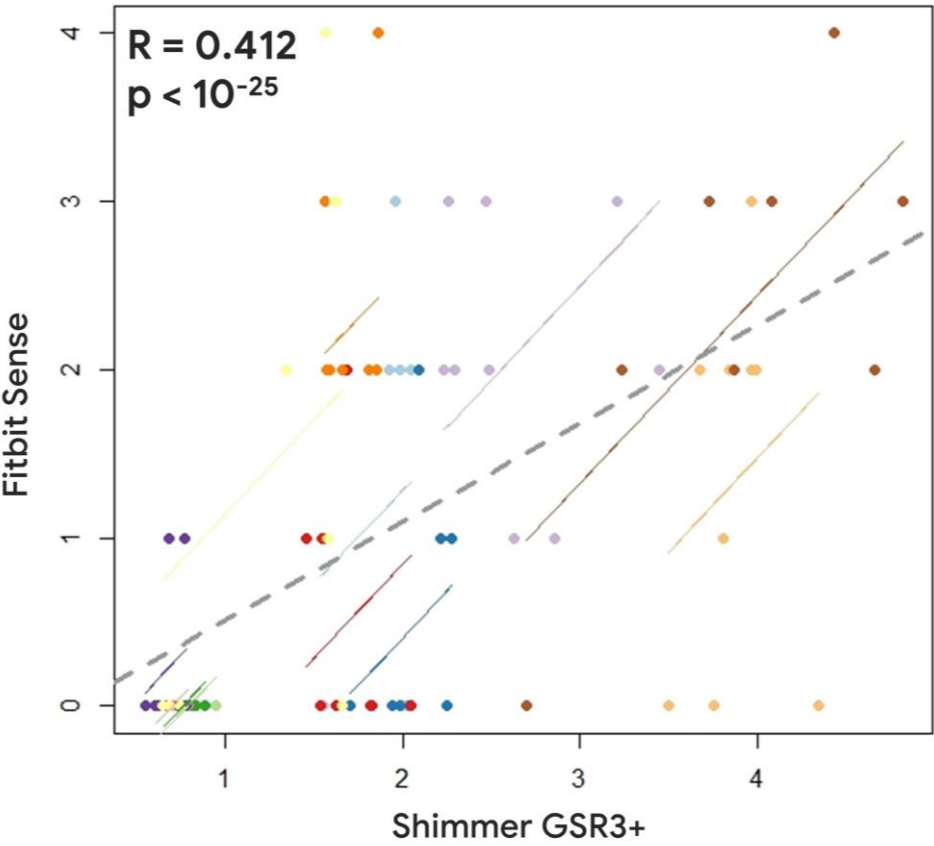


Figure 7. Repeated measure correlation between the Skin Conductance Level estimated through the Fitbit Sense and the Shimmer GSR3+.

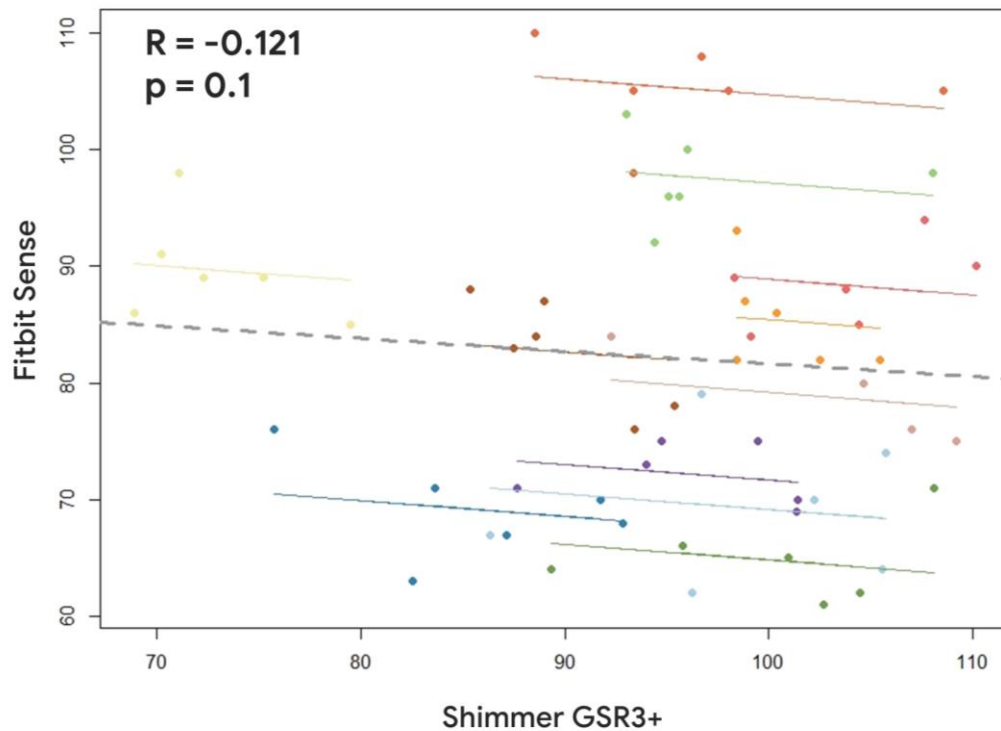


Figure 8. Repeated measure correlation between the Heart Rate estimated through the Fitbit Sense and the Shimmer GSR3+.

Discussion

The objective of the presented study consisted in the preliminary assessment of the consumer and research wearables' reliability with the respect to laboratory equipment, considered as gold standard, in collecting PPG signal and EDA in order to furtherly and reliably estimate autonomic parameters, such as the HR and SCL. This aspect is of relevant importance, since autonomic parameters result to be associated with the objective evaluation of specific mental states [10,38–40]. Being able to estimate these parameters under real conditions would open up the possibility of reliably assessing mental states in operational environments where particularly invasive instrumentation, such as laboratory equipment, cannot be used [41].

The presented results consistently demonstrated, accordingly to previous scientific evidences [2,8,9,14,42], the reliability of the research wearable, i.e., the Empatica E4, with the respect to the Shimmer GSR3+. In fact, both the Pearson's correlation (Figures 2 and 3) and the repeated measure correlation analyses (Figures 4 and 5) revealed that for the great majority of participants, at least for 9 participants among 12, the HR and SCL parameters estimated through the signals collected by the Empatica E4 were highly (all $R > 0.58$) and significantly (all $p < 0.05$) correlated with the ones estimated from the signals collected through the Shimmer GSR3+. On the contrary, the consumer wearable, i.e., the Fitbit Sense, did not show the same grade of reliability. In particular, the correlation analyses showed positive and significant correlations for a very low participants' number. The HR estimated through the Fitbit Sense resulted to not significantly correlated with the one estimated through the Shimmer GSR3+. A different discussion needs to be made with the regard to the SCL estimation through the Fitbit Sense. Even if the Pearson's correlation analysis highlighted a positive (all $R > 0.48$) and significant (all $p < 0.05$) correlations only for 2 participants among 12 (Figure 2), it has to be noted that the repeated measure correlation among all the participants resulted to be positive and significant (Figure 6). This result reflects a promising starting point to furtherly investigate the consumer wearables' reliability in scientific research field.

From a mere methodological and technical point of view, it is important to note that, actually, the low reliability of Fitbit measures can be caused by the sensors or by the integrated software for

metrics computation. In fact, the Fitbit does not provide access to raw signal recorded, that is the EDA and the PPG, but only to the parameters related to these two neurophysiological activities computed by their proprietary software, respectively EDA responses and Heart Rate. It means that despite these results, it is not said that with by improving the data processing and computation this consumer-grade device would increase its potential for this kind of applications. Anyhow, from a practical perspective the restricted access to raw data with third-party applications is a great limitation for scientific applications, where there is usually the need for data synchronization with other additional data, and for raw data access in order to investigate different data features and patterns as well as new procedures and algorithms for data processing. This precludes the possibility of applying techniques for identifying and correcting movement-based signal artifacts, which result to be very common in real-world applications and that might relevantly and negatively interfere with the correct estimation of the autonomic parameters of interest. In scientific literature, a plethora of works were already presented for validating data preprocessing techniques specifically developed for the PPG signal and EDA analyses in real-world applications [30,43,44].

As introduced, the present study is limited by the low sample size and probably by the novelty of the investigated consumer-grade device (the Fitbit Sense was just released on the market), therefore new studies on this path are strongly encouraged in order to verify, with larger sample sizes, whether technical improvements would have increased the reliability of these devices.

Conclusions

To conclude, the presented preliminary study demonstrated the consistent reliability associated to the usage of research wearables, i.e., the Empatica E4, in scientific research for real-world applications. Regarding the consumer wearables, it has to be observed how, at least for the EDA-derived parameter, the measurements were positively correlated at single – participant level with the ones provided by the gold standard equipment. This aspect leads to the conclusion that the consumer wearables' reliability must be furtherly investigated, especially by combining the raw data collection with specific preprocessing techniques. As highlighted in the Results section, the non-positive correlation observed between the HR evaluated through the consumer grade wearable device, i.e., the Fitbit Sense, and the ones estimated through the considered gold-standard, i.e., the Shimmer GSR3+, could relevantly be related to the lack of ad-hoc preprocessing techniques application for the PPG raw signal collected through the Fitbit Sense. To this regard, it has to be observed how consumer wearables represent the perfect trade-off between usability, since they can be used in real-world applications without negatively interfere with the user's activities, and cost, given their a very low economic cost with the respect to the research wearables and the laboratory equipment. By fulfilling this gap highlighted by the present work it will relevantly enlarge the possibility of employing consumer grade wearable devices, such as the Fitbit Sense, in scientific research for longitudinally assess relevant human's mental states.

Fundings: This work was co-funded by the European Commission by Horizon 2020 projects "FITDRIVE: Monitoring devices for overall FITness of Drivers" (GA n. 953432). The individual grants "AI-DRIVE: AI-based multimodal evaluation of car drivers' performance for on-board assistive systems" (Avvio alla ricerca 2021), provided by Sapienza University of Rome to Gianluca Di Flumeri, "The Smelling Brain: discovering the unconscious effect of the odours in industrial contexts" (Avvio alla ricerca 2022), provided by Sapienza University of Rome to Alessia Vozzi, "REMES – Remote tool for emotional states evaluation" provided to Vincenzo Ronca, and "HF AUX-Aviation: Advanced tool for Human Factors evaluation for the AUXiliary systems assessment in Aviation", provided by Sapienza University of Rome to Vincenzo Ronca are also acknowledged.

References

1. V. Ronca *et al.*, "Contactless Physiological Assessment of Mental Workload During Teleworking-like Task," in *Communications in Computer and Information Science*, Springer Science and Business Media Deutschland GmbH, Dec. 2020, pp. 76–86. doi: 10.1007/978-3-030-62302-9_5.
2. V. Ronca *et al.*, "A Video-Based Technique for Heart Rate and Eye Blinks Rate Estimation: A Potential Solution for Telemonitoring and Remote Healthcare," *Sensors*, vol. 21, no. 5, p. 1607, Feb. 2021, doi: 10.3390/s21051607.
3. M. Nasser *et al.*, "Signal quality and patient experience with wearable devices for epilepsy management," *Epilepsia*, vol. 61, no. S1, pp. S25–S35, Nov. 2020, doi: 10.1111/EPI.16527.
4. L. Zhu *et al.*, "Feasibility Study of Stress Detection with Machine Learning through EDA from Wearable Devices," *IEEE International Conference on Communications*, vol. 2022-May, pp. 4800–4805, 2022, doi: 10.1109/ICC45855.2022.9838970.
5. L. Zhu, P. Spachos, and S. Gregori, "Multimodal Physiological Signals and Machine Learning for Stress Detection by Wearable Devices," *2022 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2022 - Conference Proceedings*, 2022, doi: 10.1109/MEMEA54994.2022.9856558.
6. D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *Int J Biosens Bioelectron*, vol. 4, no. 4, p. 195, 2018, doi: 10.15406/IJBSBE.2018.04.00125.
7. J. C. Lo, E. Sehic, and S. A. Meijer, "Measuring Mental Workload With Low-Cost and Wearable Sensors: Insights Into the Accuracy, Obtrusiveness, and Research Usability of Three Instruments," *J Cogn Eng Decis Mak*, vol. 11, no. 4, pp. 323–336, Dec. 2017, doi: 10.1177/1555343417716040.
8. G. Borghini *et al.*, "Stress Assessment by Combining Neurophysiological Signals and Radio Communications of Air Traffic Controllers," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Institute of Electrical and Electronics Engineers Inc., Jul. 2020, pp. 851–854. doi: 10.1109/EMBC44109.2020.9175958.
9. A. Giorgi *et al.*, "Wearable Technologies for Mental Workload, Stress, and Emotional State Assessment during Working-Like Tasks: A Comparison with Laboratory Technologies," *Sensors*, vol. 21, no. 7, p. 2332, Mar. 2021, doi: 10.3390/s21072332.
10. L. Menghini, E. Gianfranchi, N. Cellini, E. Patron, M. Tagliabue, and M. Sarlo, "Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions," *Psychophysiology*, vol. 56, no. 11, Nov. 2019, doi: 10.1111/psyp.13441.
11. S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1. BioMed Central, p. 21, Apr. 2012. doi: 10.1186/1743-0003-9-21.
12. G. Borghini *et al.*, "Air Force Pilot Expertise Assessment during Unusual Attitude Recovery Flight," *Safety 2022, Vol. 8, Page 38*, vol. 8, no. 2, p. 38, May 2022, doi: 10.3390/SAFETY8020038.
13. G. Borghini *et al.*, "A multimodal and signals fusion approach for assessing the impact of stressful events on Air Traffic Controllers," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–18, May 2020, doi: 10.1038/s41598-020-65610-z.
14. C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, 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, Mar. 2010, doi: 10.1109/TITB.2009.2036164.
15. G. Borghini *et al.*, "Real-time Pilot Crew's Mental Workload and Arousal Assessment During Simulated Flights for Training Evaluation: a Case Study," *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3568–3571, Jul. 2022, doi: 10.1109/EMBC48229.2022.9871893.
16. N. Samadiani *et al.*, "A multiple feature fusion framework for video emotion recognition in the wild," *Concurr Comput*, p. e5764, Apr. 2020, doi: 10.1002/cpe.5764.
17. J. F. Brosschot and J. F. Thayer, "Heart rate response is longer after negative emotions than after positive emotions," *International Journal of Psychophysiology*, vol. 50, no. 3, pp. 181–187, Nov. 2003, doi: 10.1016/S0167-8760(03)00146-6.
18. D. Girardi, F. Lanubile, and N. Novielli, "Emotion detection using noninvasive low cost sensors," in *2017 7th International Conference on Affective Computing and Intelligent Interaction, ACII 2017*, Institute of Electrical and Electronics Engineers Inc., Jan. 2018, pp. 125–130. doi: 10.1109/ACII.2017.8273589.

19. M. Ragot, N. Martin, S. Em, N. Pallamin, and J. M. Diverrez, "Emotion recognition using physiological signals: Laboratory vs. wearable sensors," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2018, pp. 15–22. doi: 10.1007/978-3-319-60639-2_2.
20. V. Ronca et al., "Low-invasive Neurophysiological Evaluation of Human Emotional State on Teleworkers," in *ICETE International Conference on E-Business and Telecommunication Networks (International Joint Conference on Computational Intelligence)*, Science and Technology Publications, Lda, 2021, pp. 427–434. doi: 10.5220/0010726700003063.
21. A. Vozzi, V. Ronca, P. Cherubino, A. Trettel, and F. Babiloni, "Measuring the Emotional and Cognitive Consumers' Responses During Interaction with Marketing Stimuli," *Contributions to Management Science*, pp. 137–164, 2021, doi: 10.1007/978-3-030-67020-7_8/COVER.
22. A. Tiwari, R. Cassani, J. F. Gagnon, D. Lafond, S. Tremblay, and T. H. Falk, "Prediction of Stress and Mental Workload during Police Academy Training Using Ultra-Short-Term Heart Rate Variability and Breathing Analysis," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2020-July, pp. 4530–4533, Jul. 2020, doi: 10.1109/EMBC44109.2020.9175414.
23. S. Mach, P. Storzynski, J. Halama, and J. F. Krems, "Assessing mental workload with wearable devices – Reliability and applicability of heart rate and motion measurements," *Appl Ergon*, vol. 105, p. 103855, Nov. 2022, doi: 10.1016/J.APERGO.2022.103855.
24. C. L. Karmen, M. A. Reisfeld, M. K. McIntyre, R. Timmermans, and W. Frishman, "The Clinical Value of Heart Rate Monitoring Using an Apple Watch," *Cardiol Rev*, vol. 27, no. 2, pp. 60–62, Mar. 2019, doi: 10.1097/CRD.0000000000000243.
25. A. S. Jat and T. M. Grønli, "Smart Watch for Smart Health Monitoring: A Literature Review," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13346 LNBI, pp. 256–268, 2022, doi: 10.1007/978-3-031-07704-3_21/COVER.
26. J. M. Raja et al., "Apple Watch, Wearables, and Heart Rhythm: where do we stand?," *Ann Transl Med*, vol. 7, no. 17, pp. 417–417, Sep. 2019, doi: 10.21037/ATM.2019.06.79.
27. Y. Liao, C. Thompson, S. Peterson, J. Mandrola, and M. S. Beg, "The Future of Wearable Technologies and Remote Monitoring in Health Care.," *Am Soc Clin Oncol Educ Book*, vol. 39, no. 39, pp. 115–121, Jan. 2019, doi: 10.1200/EDBK_238919.
28. T. Scquizzato, A. Gazzato, G. Landoni, and A. Zangrillo, "Assessment of noise levels in the intensive care unit using Apple Watch," *Crit Care*, vol. 24, no. 1, pp. 1–3, Apr. 2020, doi: 10.1186/S13054-020-02852-3/FIGURES/2.
29. J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans Biomed Eng*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985, doi: 10.1109/TBME.1985.325532.
30. P. Ghosal, S. Himavathi, and E. Srinivasan, "PPG motion artifact reduction using neural network and spline interpolation," *2020 7th International Conference on Smart Structures and Systems, ICSSS 2020*, Jul. 2020, doi: 10.1109/ICSSS49621.2020.9202214.
31. D. R. Bach, "A head-to-head comparison of SCRalyze and Ledalab, two model-based methods for skin conductance analysis," *Biol Psychol*, vol. 103, no. 1, pp. 63–68, 2014, doi: 10.1016/j.biopsycho.2014.08.006.
32. M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *J Neurosci Methods*, vol. 190, no. 1, pp. 80–91, Jun. 2010, doi: 10.1016/j.jneumeth.2010.04.028.
33. J. J. Braithwaite, D. Derrick, G. Watson, R. Jones, and M. Rowe, "A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments.," 2015.
34. N. Tateyama, K. Ueda, and M. Nakao, "Development of an active sensing system for distress detection using skin conductance response," *2019 8th International Conference on Affective Computing and Intelligent Interaction, ACII 2019*, pp. 353–358, Sep. 2019, doi: 10.1109/ACII.2019.8925442.
35. D. S. Wickramasuriya and R. T. Faghih, "A Marked Point Process Filtering Approach for Tracking Sympathetic Arousal from Skin Conductance," *IEEE Access*, vol. 8, pp. 68499–68513, 2020, doi: 10.1109/ACCESS.2020.2984508.
36. S. C. Jacobs et al., "Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research," *Am Heart J*, vol. 128, no. 6, pp. 1170–1177, Dec. 1994, doi: 10.1016/0002-8703(94)90748-X.
37. J. Z. Bakdash and L. R. Marusich, "Repeated measures correlation," *Front Psychol*, vol. 8, no. MAR, p. 456, Apr. 2017, doi: 10.3389/FPSYG.2017.00456/BIBTEX.

38. G. Borghini, V. Ronca, A. Vozzi, P. Aricò, G. Di Flumeri, and F. Babiloni, "Monitoring performance of professional and occupational operators," in *Handbook of Clinical Neurology*, Elsevier B.V., 2020, pp. 199–205. doi: 10.1016/B978-0-444-63934-9.00015-9.
39. R. Laureanti *et al.*, "Emotion assessment using Machine Learning and low-cost wearable devices," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Institute of Electrical and Electronics Engineers Inc., Jul. 2020, pp. 576–579. doi: 10.1109/EMBC44109.2020.9175221.
40. M. Haghi, K. Thurow, and R. Stoll, "Wearable devices in medical internet of things: Scientific research and commercially available devices," *Healthc Inform Res*, vol. 23, no. 1, pp. 4–15, Jan. 2017, doi: 10.4258/hir.2017.23.1.4.
41. S. Gradl, M. Wirth, R. Richer, N. Rohleder, and B. M. Eskofier, "An overview of the feasibility of permanent, real-time, unobtrusive stress measurement with current wearables," in *ACM International Conference Proceeding Series*, New York, NY, USA: Association for Computing Machinery, May 2019, pp. 360–365. doi: 10.1145/3329189.3329233.
42. G. Borghini *et al.*, "Assessment of Athletes' Attitude: Physiological Evaluation via Wearable Sensors during Grappling Competitions," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2020-July, pp. 584–587, Jul. 2020, doi: 10.1109/EMBC44109.2020.9176401.
43. G. Georgieva-Tsaneva, "A Novel Photoplethysmographic Noise Removal Method via Wavelet Transform to Effective preprocessing," *ACM International Conference Proceeding Series*, pp. 113–118, Jun. 2022, doi: 10.1145/3546118.3546145.
44. S. S. Bashar, M. S. Miah, A. H. M. Z. Karim, M. A. Al Mahmud, and Z. Hasan, "A Machine Learning Approach for Heart Rate Estimation from PPG Signal using Random Forest Regression Algorithm," *2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019*, Apr. 2019, doi: 10.1109/ECACE.2019.8679356.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.