Validation Study of the Empatica EmbracePlus Wristband for Measuring Electrodermal Activity Utilizing a Virtual Reality Protocol

Yin Qu

Faculty of Behavioural, Management and Social Science (BMS), University of Twente

Health Psychology & Technology, Master’s Thesis (201900200)

Prof. Dr. Matthijs L. Noordzij, Magdalena Sikora

August 25, 2024

Abstract

Wearable technology for EDA monitoring has gained significant attention, yet validation studies remain inconclusive. This preliminary study assessed the EmbracePlus wristband's validity using a VR protocol and explored the effects of dynamic factors and temperature among 21 participants aged 18–32 with a BIOPAC system as the reference. Our findings challenged previous assumptions regarding the detection capabilities of wrist-worn EDA devices for significant stressors. While the EmbracePlus wristband demonstrated limitations in capturing EDA signals, parameters, and event-related effects compared to the reference device, it provided valuable insights into the sensitivity of wrist EDA to different stressors. Palm EDA exhibited greater responsiveness to both the VR stressor and dynamic factors, suggesting a synergistic effect. Wrist EDA, however, appeared less sensitive to both, highlighting the need for further research to understand the impact of stressor type on wrist EDA. Future research could focus on validating wearable devices across diverse stressors, incorporating advanced thermoregulatory indicators, and conducting long-term studies in real-world settings to optimize wearable technology for stress monitoring.

Keywords: wearable, validation, electrodermal activity (EDA), skin conductance level (SCL), skin conductance responses (SCRs), total amplitude, detrended cross-correlation analysis (DCCA), Bland–Altman plot, error bar plot, detrended partial cross-correlation analysis (DPCCA), Empatica EmbracePlus wristband, BIOPAC systems.

Stress management has grown in importance for individuals and organizations due to the proven negative impact of chronic stress on health (Morrison & Bennett, 2022). Studies have shown that prolonged exposure to stressors are associated with a wide range of mental and physical health outcomes, including unhealthy coping behaviour and cardiovascular problems (Morrison & Bennett). The estimated annual cost of work-related stress in the Netherlands was €4 billion (Blatter et al., 2005). A modern approach for ambulatory stress management involves the real-time monitoring of physiological processes and providing personal feedback. This approach is known as ambulatory biofeedback or biocueing, and is seen as a promising way for empowering individuals to gain control over their stress responses, potentially leading to improved stress management (Neupane et al., 2024; Sanches et al., 2010; Sehrt et al., 2024; Ter Harmsel et al., 2021).

Many commercial wearables have introduced biofeedback notifications indicating potential stress (Neupane et al., 2024). However, this technology is still in the infancy stage. Due to the novel nature, there is a growth in validation studies comparing the signals from wearable devices to that from more established reference tools. With this context in mind, the goal of this study is to compare physiological signals from two sources: the Empatica EmbracePlus wristband, newly designed for both commercial and research purposes (Empatica Inc., 2019, n.d.-a), and the extensively validated laboratory-grade system, the BioNomadix wireless system (BIOPAC Systems Inc., n.d.). This study focuses on the single physiological process of electrodermal activity (EDA), which shows potential to predict stress status (Zhu et al., 2022).

Additionally, this study adopted Van Lier et al. (2020)'s standardized validation protocol to bolster the comparability of validation outcomes. Moreover, recognizing that the intended applications of the wristband span real-life scenarios or semi-naturalistic laboratory settings, this study extends the traditional laboratory design by incorporating Virtual Reality (VR) technology to simulate ecological environments and evoke strong stress responses under controlled conditions. A VR high-altitude stressor, standing on a plank 80 floors high, was chosen based on its proven effectiveness in triggering significant stress responses (Aspiotis et al., 2022; Boccignone et al., 2021). Incorporating VR not only broadens the scope of investigation but also enhances the feasibility of Van Lier et al.'s protocol in conducting validations in an immersive environment.

Current Validations

The EmbracePlus wristband by Empatica represents a new advancement in smart sensor technology for EDA acquisition, having been applied in various scientific investigations across multiple domains (Böttcher et al., 2022; Firouz 2023; Grasser et al., 2023). This device is considered an upgrade over the previous Empatica E4 wristband (Empatica Inc., n.d.-c). However, existing published validations of EmbracePlus rarely focus on EDA (Gerboni et al., 2023; Sinichi et al., 2024), leaving a gap in understanding the validity of EDA data obtained from the EmbracePlus.

Previous assessments of the Empatica E4 model have yielded mixed results regarding EDA data. For instance, Ronca et al. (2023) found a significant correlation between E4 and a gold-standard SCL measurement in resting conditions. In contrast, several studies have reported negative findings. Borrego et al. (2019) observed none to moderate correlations between E4 and a laboratory-grade device during picture viewing in resting and temperature-controlled conditions. Costantini et al. (2023) concluded E4 failed to produce reliable data in video-simulated stressful driving. Milstein and Gordon (2020) did not find a correlation between the mean SCL obtained by the E4 and by a well-validated mobile device during a dyadic conversation, with 73% of E4 data deemed noise. Hu et al. (2024) reported high reliability but unacceptable validity for E4-derived parameters in both laboratory and ambulatory settings. Menghini et al. (2019) found no visual resemblance and only weak correlations between E4 and finger electrode signals under various laboratory and ecological conditions. Similarly, two other studies also observed no visual resemblance in EDA signals between the E4 and traditional reference devices. However, they provided new insights into wrist-specific advantages. Ollander et al. (2016) suggested that the wrist might have unique benefits for stress detection, while Van Lier et al. (2020) demonstrated the E4's feasibility in capturing physiological stress responses to a sustained, high-intensity social stressor.

The different patterns between wrist and palm signals were frequently discussed in these studies, in line with previous works described by Dawson et al. (2016, p. 217–243). Factors contributing to extent differences include sweat gland density (i.e., 108 glands/cm^2 on the forearm vs. 600 to 700 glands/cm2 on the palm), electrode materials (i.e., stainless steel vs. Ag/AgCl), and electrode size (i.e., around .5 vs. 1.1 cm in diameter). Higher gland density, the use of Ag/AgCl, and larger areas of contact can increase EDA parameters (Fowles et al., 1981; Freedman et al., 1994; Van Dooren & Janssen, 2012). Additionally, testing only one wrist may provide a partial understanding of arousal responses to stressors. Previous studies supporting the theory of multiple arousals illustrated that reactions from both wrists exhibited significant asymmetry to a classical mental arithmetic stressor (Payne et al., 2016). When participants perceived the stressor as threatening rather than positively arousing, testing only the nondominant wrist could underestimate emotional arousal (Picard et al., 2016).

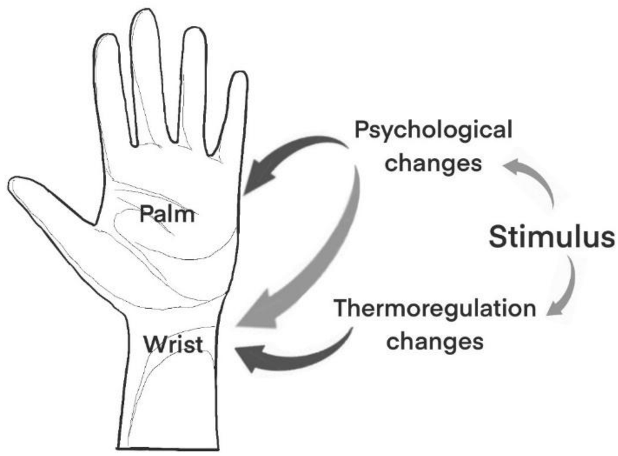
Prior validations have primarily attributed the distinct nature of EDA signals from different body sites to thermoregulatory influences (Menghini et al., 2019; Van Lier et al., 2020). Research indicates that palmar and plantar sweating are largely independent of thermoregulatory processes (Boucsein, 2012; Boucsein et al., 2012; Kerassidis, 1994; Turpin et al., 1983). For example, Kerassidis observed minimal sweating on the palms and soles in the absence of emotional arousal, even under high environmental temperatures. Other studies (e.g., Ollander et al., 2016; Van Lier et al., 2020) have highlighted the correlation between wrist-measured EDA and mental strain, emphasizing the wrist's sensitivity to both psychological and thermoregulatory changes.

The Role of Thermoregulation in EDA

The dual sensitivity of wrist EDA suggests that a stressor inducing significant thermoregulatory and psychological responses is likely to provoke sweating on both the wrist and palm. Conversely, a stressor that primarily elicits thermoregulatory responses may predominantly affect wrist sweating. A mind map (Figure 1) visually illustrates the proposed relationships among the stimulus, the psychological or thermoregulatory responses it provoke, and the subsequent EDA responses at the palm and wrist. Based on this model, exploring how thermoregulation impacts wrist EDA could provide further insight into potential discrepancies between palm- and wrist-worn devices.

Figure 1

A Possible EDA Response Model across Wrist and Palm



Thermoregulation, however, is a complex interactive system involving both internal heat transfer and external heat exchange. In laboratory settings, ambient temperature fluctuations are typically minimal due to the slow pace of natural temperature changes and controlled conditions. Whereas, psychological stimuli can trigger thermoregulatory responses as well. For instance, unpredictable stressors like psychological stress can elevate core body temperature while simultaneously decreasing skin temperature (Rimm-Kaufman & Kagan, 1996). Perceived negative emotions during rest can halt the rise in skin temperature (McFarland, 1985). If palm sweating is indeed unaffected by thermoregulation directly, either ambiently or psychologically induced thermoregulatory changes could exert a more pronounced influence on wrist sweating compared to palm sweating. However, the measurement of thermoregulatory changes remains challenging.

While several models estimate sweating rates using equations, their accuracy across different environments remains under investigation, and the measurement process can be complex. For example, the multi-node dynamic UTCI–Fiala model (Fiala et al., 2012) identifies sweating as one of four essential thermoregulatory responses managed by the central nervous system. Estimating sweating rates within this model requires measuring local skin temperature, weighted mean skin temperature, and hypothalamus temperature. Conducting such complex measurements for thermoregulation indicators was beyond the scope of this master's thesis project. Instead, we focused on exploring the influence of ambient and local peripheral temperatures, as they are direct and easily obtainable indicators.

Moreover, the stress detection capabilities of wrist EDA observed in two social stressor studies (Ollander et al., 2016; Van Lier et al., 2020) may be partly attributed to influences of social interactions and physical activity—factors inherently linked to thermoregulation (Davies, 1979; IJzerman et al., 2017). To address the limited understanding of these influences, the stressor design in this study deliberately separated the interaction and movement components from the stressor. This approach allowed us to observe the impact of a relatively "pure" psychological stimulus on wrist and palm EDA responses, while also investigating the differential effects of interaction and movement on these two sites. Below, we briefly summarize relevant information to provide more background and rationale for this validation study.

Challenges of Validations

De Geus and Gevonden (2024) summarized that the validation of wrist-placed dry electrode measurements against gold-standard wet electrode measures remains limited. While validation results have not always been promising, they emphasized the scarcity of large-scale validation studies for ambulatory EDA. They also highlighted the need for further research to identify optimal signals, study designs, and analytical approaches (De Geus & Gevonden, 2024).

To address the lack of standardization in validation results, Van Lier et al. (2020)

introduced a standardized validity assessment protocol to foster consensus in data analysis. This protocol systematically compares signals, parameters, and event effects across devices using established criteria, while prioritizing efficiency and replicability in the rapidly evolving wearable technology landscape. To ensure outcomes' comparability, this validation adopted Van Lier et al.'s protocol.

The protocol encompasses three levels of evaluation with corresponding standards. The first level, signal comparison, uses detrended cross-correlation analysis (DCCA) to compare the whole time series. Threshold of DCCA coefficients is set as .8. The second level, parameter comparison, involves analysing differences between two methods using the Bland–Altman plot. The main result of this step is to identify if potential systematic biases exist, as suggested by Giavarina (2015). For example, as shown in Figure 2, method A systematically gives lower values than method B, and the differences seem to be larger after the mean is higher than 100. The third level, event detection comparison, includes several steps. First, respectively plotting two signals means and standard errors within each task to observe the effect of event detected by two methods. When the error bars of baseline and task totally not overlap with each other, a significant event effect is identified. As can be seen in Figure 3, the error bars of preparing singing and singing are both not overlapping with the baseline's error bar. So that a significant event effect is identified by the RD. If both methods detected effect, then, plotting the differences between two methods within each task to explore the agreement. If the error bar for a task cross the zero axis, then, observing the error bars and boundaries (difference between the baseline and the task measured by the RD is set as the boundary size). If the difference mean and corresponding standard error lie between the priori defined boundaries, the devices show agreement for the task. This protocol comprehensively inspect the alignment between devices.

Figure 2

An Example of the Bland–Altman Plot (Giavarina, 2015)

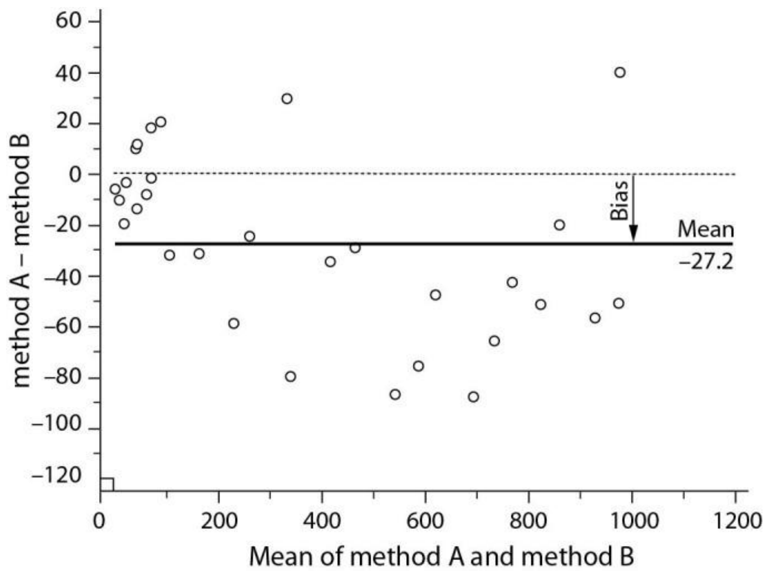
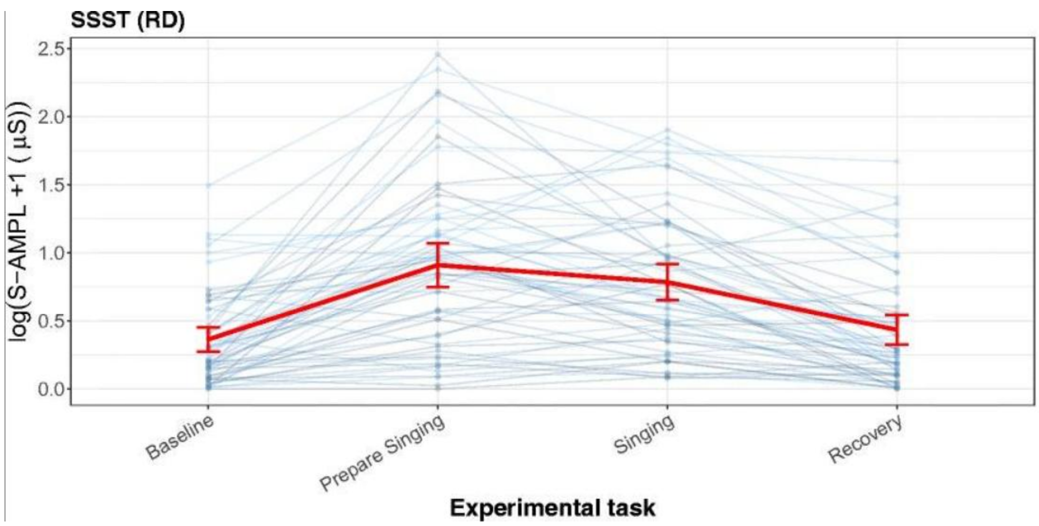


Figure 3

An Example of the Error Bar Plot (Van Lier et al., 2020)



Involved Parameters

The EDA comprises tonic and phasic components (Figure 4). Skin conductance level (SCL), the most common measure of the tonic component, reflects the slower baseline skin conductance. Changes in SCL are associated with gradual shifts in autonomic nervous system activity. The phasic component, represented by skin conductance responses (SCRs), encompasses rapid fluctuations in skin conductance. SCRs can be elicited by external stimuli or occur spontaneously. Individual SCRs typically last several seconds, while SCRs as a parameter often refers to their frequency per minute. SCR amplitude, calculated as the difference between response onset and peak (Figure 5), is another key measure.

Figure 4

EDA Data Visualization: Raw Data, Tonic, and Phasic Components (Posada-Quintero & Chon, 2020)

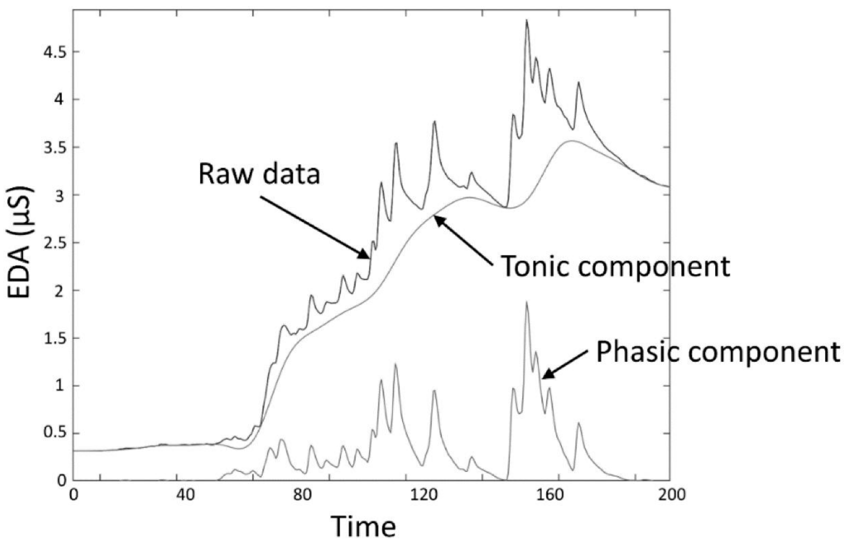
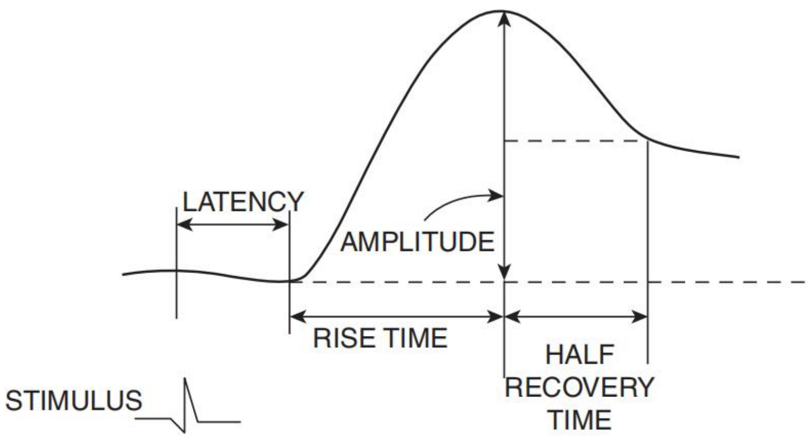


Figure 5

Graphical Representation of Principal EDA Components of a Single SCR (Dawson et al., 2016)



The typical values of electrodermal measures among healthy young adults are listed in the handbook of psychophysiology (Dawson et al., 2016). The SCL ranges from 2 to 20 µSiemens (µS). The frequency of SCRs in absence of identifiable eliciting stimulus usually ranges between 1 and 3 per minute (over 20 /min in high arousal situations (Boucsein, 2012)). The size of an elicited SCR amplitude typically ranges between .2 and 1 µS. However, it's common to consider a minimum amplitude that varies from .01 to .05µS, especially in the case of wrist-worn sensors. According to Van Lier et al.'s protocol, this study focuses on three parameters: SCL and SCRs and SCRs total amplitude (AMPL). The AMPL is the sum of the SCR amplitudes per minute, a combination of both the frequency and amplitude of the SCRs. Biologically plausible values for AMPL are between 0 and 3 µS and on average .3–1.3 µS (Braithwaite et al., 2013).

Application of VR in Stress Research

This study integrated VR technology into the laboratory validation process to introduce elements of ambulatory and naturalistic conditions. This approach was motivated by the demand for ambulatory validation of wearable devices and the consideration of the reference devices' limited portability in ambulatory settings (Hu et al., 2024). Previous research has established the significant impact of VR exposure on EDA in stress research (Van Dammen et al., 2022). However, the use of VR to simulate intense stressors for wearable device validation has been unexplored. This innovative approach offers ecologically valid conditions while maintaining experimental control, mitigating ethical concerns associated with real-world stressors. By integrating VR technology, this study expands the applicability of Van Lier et al.'s (2020) protocol to simulated naturalistic environments, providing insights into device and protocol performance in more ecologically relevant contexts.

Selection of Reference Device and Electrode Placement

The choice of BioNomadix system as the reference device is grounded in its extensive use in research and the consensus in prior studies that wet electrodes offer more accurate EDA measurements for short-term applications. Wet electrodes, which employ electrolyte paste to enhance skin contact, provide superior data accuracy for measurements lasting less than two hours compared to dry electrodes (electrolyte-free) used in wearables (Boucsein, 2012). The BioNomadix was placed on the palm area, a traditional gold-standard site for EDA acquisition. The laboratory validity of EDA responses to various stimuli has been largely established on the palm site (Dawson et al., 2016; Egger et al., 2019; Rohrbaugh, 2016), making it reliable for accurate EDA measurements.

Research Questions

The primary objective of this study is to evaluate the concordance between EDA data generated by EmbracePlus and an RD using Van Lier et al. (2020)'s protocol. The research questions (RQs) aligned with the protocol are:

RQ1: On the signal level, what is the level of agreement between raw EDA signals from EmbracePlus and the RD, as indicated by the optimal detrended cross-correlation coefficients, with a threshold of >.8?

RQ2: On the parameter level, are there systematic biases evident in SCL, SCRs, or AMPL values between EmbracePlus and the RD, as assessed through Bland-Altman plots?

RQ3: On the parameter level, do the 95% limits of agreement (LoA) for SCL, SCRs, and AMPL fall within the predefined boundaries of ±1.6 μS, ±2.5, and ±.6 μS, respectively?

RQ4: On the event level, do both EmbracePlus and the RD detect significant effects of the stressor, and if so, do these effects align? Significance will be determined by non-overlapping error bars, while agreement will be assessed by within-effect-size difference error bars.

The secondary objective is to investigate the role of thermoregulation-related factors in EDA measurements. This involves examining both environmental and physiological influences. Specifically, the effects of movement and social interaction—factors known to be linked to thermoregulation—are examined by comparing EDA data during dynamic and static periods. Additionally, the study explores the impact of ambient and local peripheral temperatures as direct and accessible indicators of thermoregulation. The following secondary research questions (SRQs) aim to provide additional insights:

SRQ1: Are there significant differences in EDA signals between dynamic periods (with natural interactions and movements) and static periods (with minimal interactions and movements)? It is hypothesized that dynamic periods will exert a more pronounced influence on wrist EDA compared to palm EDA.

SRQ2 (secondary goal): What is the correlation between EDA signals and ambient and local peripheral temperature? It is hypothesized that the correlation between these temperature measures and wrist EDA will be stronger than that with palm EDA.

Method

Participants

According to Van Lier et al. (2020)'s sample estimation calculated through a detailed power analysis, a final sample of 55 is sufficient for validating physiological wearables in a lab study using their standardised protocol. However, due to the scope of this master thesis study, an intended final sample size 20 was decided, where other master students might continue to sample more participants in follow-up thesis projects. Prior to the formal experiment, the researchers practiced data acquisition on themselves and conducted testing experiments with five cases (one researcher and four friends). Considering a 40% attrition rate observed in this pilot (Appendix 1.4) and a 20% nonresponse rate noted in Van Lier et al.'s research, the sample size for data acquisition was set at 50.

Participants were informed that the study was a validity assessment for the EmbracePlus wristband, without mentioning that tasks were related to stress. Exclusion criteria included use of medication affecting the SNS, obesity, and certain diseases (e.g., cardiovascular disease, high blood pressure, epilepsy). Due to time constraints and recruitment challenges, pre-screening was not conducted, and eligibility was assessed by a questionnaire during the experiment. Data collection occurred in the BMS lab on campus from May 17 to May 29, June 10, June 11, and June 21 to June 28.

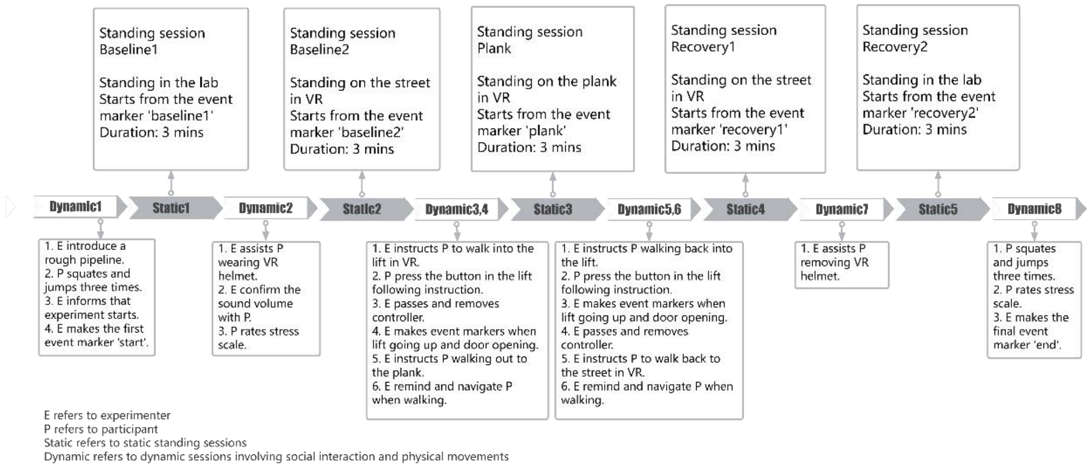
Participants were recruited from the University of Twente (UT) student and staff populations, as well as healthy adults from the researchers' contacts. Recruitment methods included self-selection through the campus SONA subject platform, convenience sampling via invitations sent to contacts, and selective sampling by inviting people on campus. Participants visited the lab, read the information sheet (Appendix 1.1), and signed informed consent forms (Appendix 1.2) before enrolling in the experiment formally. After the experiment, participants received a debriefing and could request an explanation of their personal signal graph, with the option to photograph it. Bachelor students participating through the SONA platform received course credits. No other incentives were provided. Informed consent was obtained from all 29 participants, and the study was approved by the University's Ethical Committee and Personal Data Protection Code (Ethics Code: 240530).

Design

This validation study employed a prospective, non-randomized, partially blind, within-subjects design. The experiment, as illustrated in Figure 6, comprised five main standing sessions: two baseline recordings ("baseline1", "baseline2"), one stress task ("plank"), and two recovery recordings ("recovery1", "recovery2"), each lasting three minutes. Eight dynamic sessions bridged these main sessions, involving physical movements and interactions between the researcher and participants, with durations ranging from under a minute to several minutes.

Figure 6

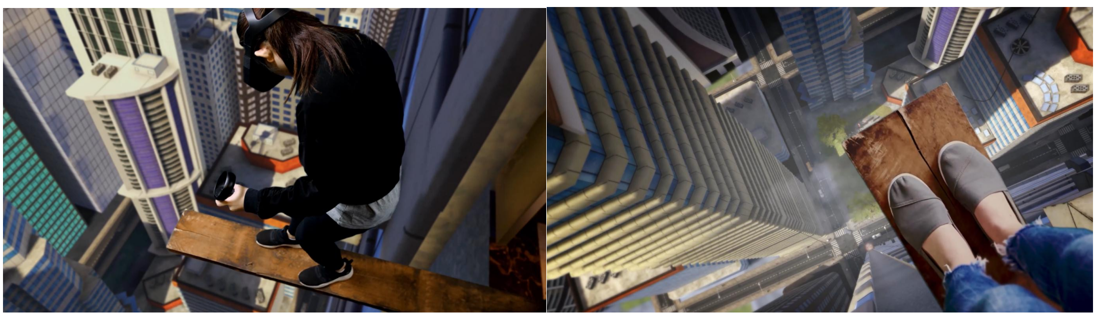
Experiment Design



The stress task involved a strong and sustained VR-simulated environmental stressor using Richie's Plank Experience (Toast Interactive, 2016), where participants stood on a virtual plank 80 stories high as shown in Figure 7. This VR game had previously been used to simulate intensive real-world stressors for assessing stress indicators, such as electroencephalography (EEG) (Aspiotis et al., 2022), heart rate, and EDA (Boccignone et al., 2021). It had also been used for exploring exposure therapy or performance training (Hu et al., 2018; Osborne et al., 2022; Ramdhani et al., 2019).

Figure 7

Game Effect Display (Toast Interactive, n.d.)



At the start of the experiment, participants were guided through the procedure and instructed to perform three squats and jumps. During the first static session, participants stood in the lab for baseline recording. They then wore the head-mounted display (HMD) device with the assistance of the researcher, and acknowledged the game's terms by pressing the “I Agree” button in VR and pulling the button under the controller. The second baseline recording was done while they stood on a VR street. Participants then entered a virtual elevator, pressed the “Plank” button, ascended to the plank floor, and walked to the end of the virtual plank. During the third static session, they stood on the plank above a cityscape. Afterward, they returned to the elevator, descended to the ground floor, walked back to the street, and completed the final two static sessions: standing on the VR street and then in the lab without the VR headset.

Given previous negative results regarding the E4’s EDA validity, this study aimed to assess the EmbracePlus’s EDA validity conservatively. Past studies have emphasized the influence of physical activity intensity when evaluating the accuracy of wearable devices (Hu et al., 2024; Menghini et al., 2019; Milstein & Gordon, 2020). Hence, this study used simple standing tasks similar to the active orthostatic test in Menghini et al.'s protocol, but without the stand up process. Each session lasted three minutes, aligning with Menghini et al.'s protocol and considering potential adverse effects of prolonged VR height exposure for participants with potential acrophobia or basophobia. Additionally, the three-minute duration was deemed acceptable based on the five-case pre-experiment (Appendix 1.4).

Before the experiment, participants completed a "Stairs" task designed to enhance EDA data collection on the wrist, as suggested by Picard et al. (2016). This task, which involved walking up and down one floor three times at a normal pace (different from Picard et al.'s original setting), extended the wearing time and introduced moderate physical activity before the main experiment. Previous studies have indicated that longer recording durations and physical activity can improve EDA signals for non-responders (Kleckner et al., 2021; Menghini et al., 2019). Additionally, participants performed a "Squats and Jumps" task at both the beginning and end of the experiment to assess the feasibility of using accelerometer data to synchronize EDA signals. This task involved three squats followed by three hops, with arms naturally hanging slightly away from the body. This physical activity at the beginning further enhanced the wrist signals.

Materials

Devices and Software

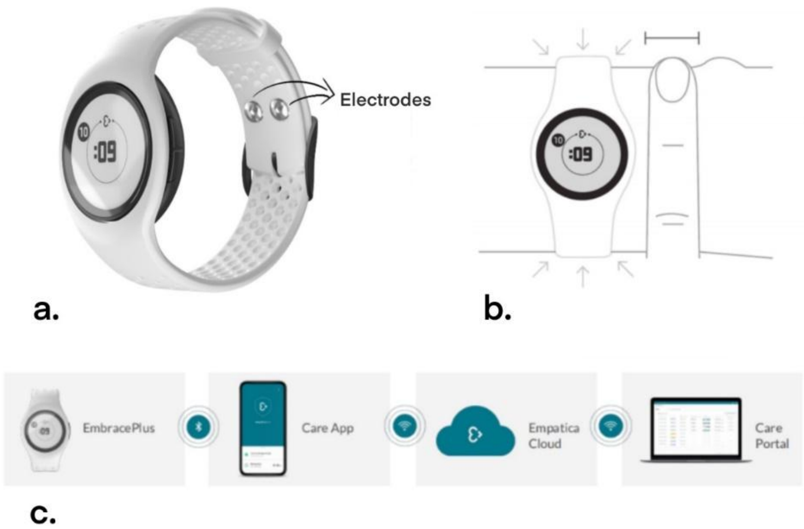
The real-world high-altitude scenario was simulated using a VR game, Richie's Plank Experience (Toast Interactive, 2016), running on different VR headsets: Oculus Quest (Facebook Inc., 2019) for participants 1–16, Meta Quest 2 (Meta Platforms, Inc., 2020) for participant 17, and Varjo VR3 (Varjo HQ, n.d.) for participants 18–29. These changes were necessitated by compatibility issues following software updates. In half the cases (details in Appendix 2.2), a wooden plank matching the virtual one was placed on the floor to enhance the immersive experience, as done in previous studies (Aspiotis et al., 2022; Boccignone et al., 2021). This setup aimed to explore the feasibility of using a real plank, though subgroup analyses are not reported in this study. EDA data was recorded using three different systems, which will be detailed later, and processed using Python. Questionnaire surveys were administered through the Qualtrics platform on an iPad.

The Empatica EmbracePlus is a research-grade wearable device designed for continuous physiological monitoring, including EDA, with a 3-axis accelerometer. It is classified as a medical device under the European Union Medical Device Directive (EU

MDD) 93/42/EEC (Empatica Inc., n.d.-b). The EDA data is collected by two stainless steel (SUS 316L) electrodes on the ventral side of the band (Figure 8a). The device should be worn a finger's width from the wrist bone on the non-dominant wrist (Figure 8b). Data synchronization occurs via Bluetooth between the sensor and a smartphone, and through Wi-Fi between the Empatica Care Lab App on the smartphone, the Empatica cloud service, and the Care portal on a computer (Figure 8c). The raw EDA data is sampled at 4 Hz, with a range of .01 to 100 µS and a resolution of 55 pSiemens.

Figure 8

EmbracePlus Wristband's Features



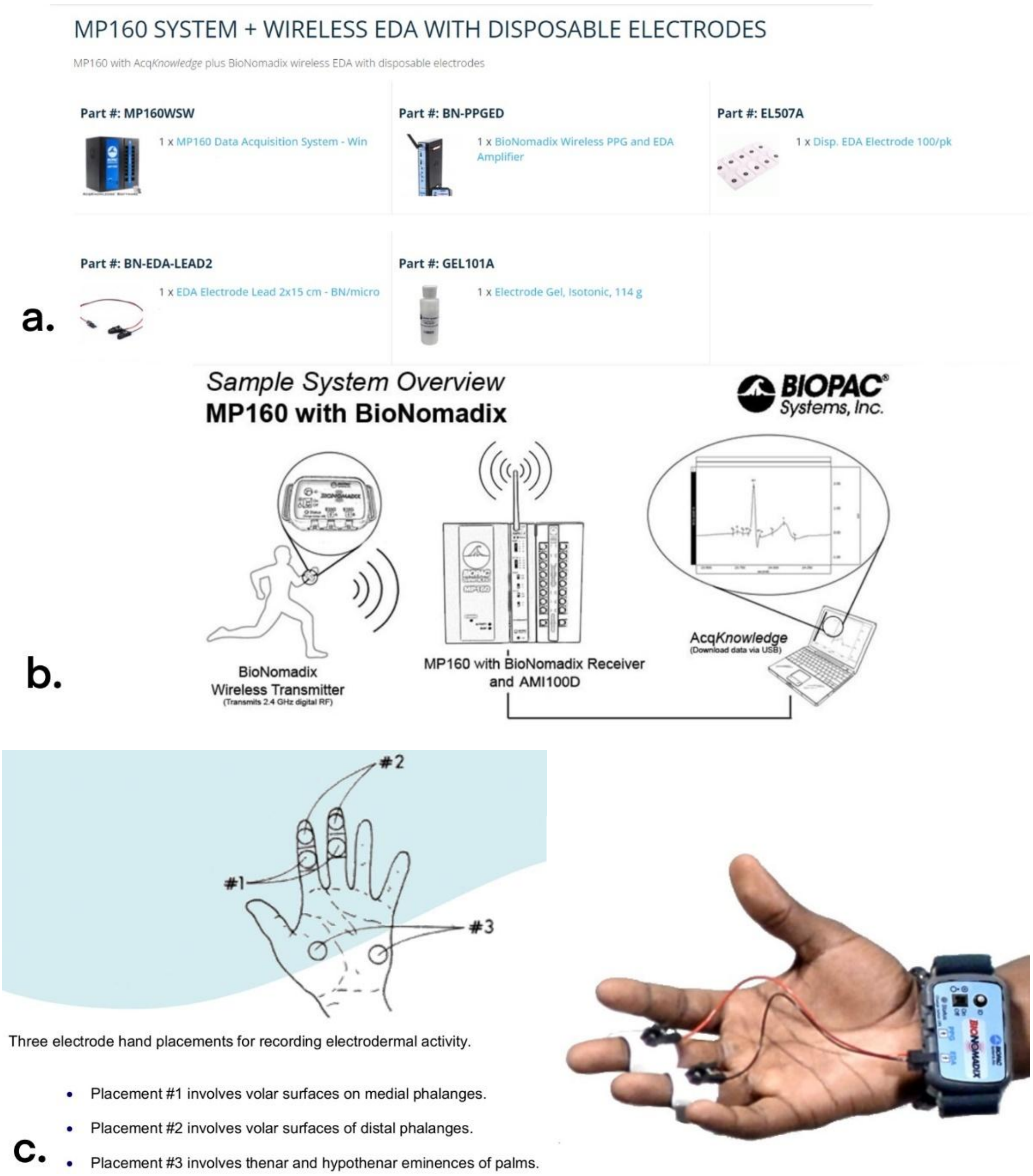
Note. a. EDA sensors; b. correct placement; c. data synchronization. Pictures are from Empatica Inc. (2024).

BioNomadix wireless system is a research-oriented EDA data acquisition solution and is in use in laboratories around the world (BIOPAC Systems, Inc., n.d.-a). It allows high-frequency EDA recordings. The components of this system are shown in Figure 9a. As shown in Figure 9b, the data collected from wet electrodes is transferred through leads to the

BioNomadix–PPGED transmitter, then wirelessly transferred to the BioNomadix–PPGED receiver and MP 160 module, and synchronized with the computer, then shown in Acqknowledge 5.0 software as real-time visualization. The clip-shaped electrode leads and the recommended placement can be seen in Figure 9c. The EDA data collected by BioNomadix is raw data without algorithms, with a sampling rate of 2000 Hz.

Figure 9

BioNomadix System's Features



Note. a. Components; b. Data synchronization; c. Recommended placement and electrode leads. Pictures are from BIOPAC Systems, Inc. (n.d.-b).

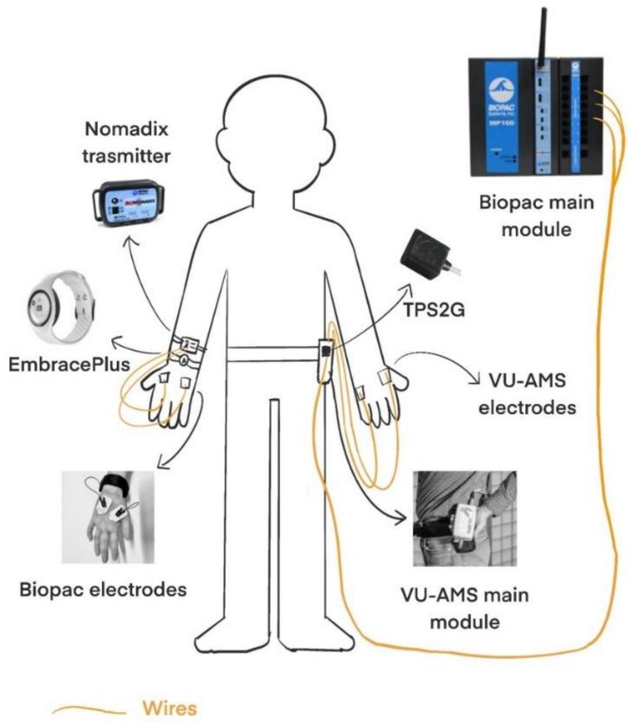
The Vrije Universiteit Ambulatory Monitoring System (VU-AMS) (version 5.0) developed at Vrije University, Amsterdam, is a research-oriented device for stress study. It is a gold standard device garnered peer approvals for EDA measuring (De Geus & Gevonden, 2024). This study evaluated the feasibility of incorporating this device, though its data was not analysed.

Data Collection

The three devices used in this study featured different electrodes and leads. The EmbracePlus wristband employed dry electrodes, while the BioNomadix system used EL507A disposable wet electrodes with BN-EDA25-LEAD2 leads. The VU-AMS system also utilized EL507A electrodes with its proprietary wiring. The EmbracePlus wristband and BioNomadix transmitter collected EDA data from the non-dominant hand, whereas the VU-AMS collected data from the dominant hand (details on the distribution of dominant-hand samples are provided in Appendix 2.2). Figure 10 illustrates how the devices were positioned. To ensure consistent contact between the skin and electrodes, a harmless isotonic gel (TD-246 Skin Conductance Electrode Paste) was applied for both the BioNomadix and VU-AMS systems (Figure 11). All data was stored on a secured cloud drive at UT.

Figure 10

The Wearing of devices



Note. This figure shows the wearing of devices for left-hand-dominant participants using eminence placement. The BioNomadix transmitter and EmbracePlus were worn on the right wrist. BIOPAC and VU-AMS electrodes were attached to both hands. The VU-AMS main module and the BIOPAC accelerometer TPS2G were fastened to the left hip using waist belts.

Figure 11

Electrolyte Application

While the distal phalanges of the index and middle fingers are often recommended for optimal EDA acquisition (Boucsein, 2012; Dawson et al.,2016), this study employed two electrode placements (Figure 12) due to experimental constraints. Participants used one hand to operate a controller, requiring finger movement. Eminence placement allowed for unrestricted finger use but shortened the distance between two pairs of electrodes to approximately 9 cm (calculated based on average hand length; Agnihotri et al., 2005), potentially introducing interference. Finger placement doubled this distance but heightened the risk of movement artifacts. To address these challenges, both placements were evaluated for feasibility (details in Appendix 2.2).

Figure 12

Finger and Eminence Electrode Placements

a.


b.


Note. a. BIOPAC electrode leads with eminence placement; b. VU-AMS electrode leads with finger placement.

Boucsein (2012) recommended controlling room temperature around 23°C and keeping humidity constant to avoid environmental influence on SCL. He also suggested controlling diurnal rhythms across experimental conditions due to their effects on EDA. Although this study did not rigorously control these environmental factors, ambient temperature, humidity, and time of day were recorded to supplement the interpretation of results (Appendix 2.4). The lab settings are shown below.

Figure 13

Lab Settings

a.  b. 

Note. a. lab before relocation; b. lab after relocation.

Surveys

A pre-experiment questionnaire (Q1–Q14) was used to confirm participant eligibility and collect demographic information. After the experiment, the Perceived Stress Scale (PSS), a widely used instrument (Cohen et al., 1994), measured the perception of stress over the past month (Q18–Q27). The PSS is a 10-item questionnaire assessing stress frequency. Three additional questions (Q15–Q17) asked participants to rate their stress on a scale of 0-100 after baseline1 and at the end of the experiment. Three post-experiment questions (Q28–Q30) investigated prior experience with similar games, fear of heights, and related diagnoses. The survey was answered on Qualtrics. The content is available in Appendix 1.3.

Procedures

Participants were not required to wash their hands beforehand, due to the noted fall in conductance following the use of soap and water (Venables & Christie, 1973) and the abrasion involved in drying hands. At the beginning of the experiment, participants were seated in front of an iPad and informed they would be attached to sensors and experience a naturalistic scenario in a VR game. Then, the researcher applied the wristband to the participant's non-dominant hand and guided them through a three-round stair-walking task with giving explanation and receiving participants' permission. Participants were then asked to read an information sheet and sign an informed consent form if they chose to participate. After enrolment, the two laboratory devices were attached. Participants then completed the first part of survey (Q1–Q14). Following this, the researcher checked the signals in real-time on the computer and conducted a breathing calibration for the BIOPAC and VU-AMS systems. This involved asking participants to take a deep breath, hold it for two seconds, and then exhale while observing the signal. They then were reminded to minimize movement and talking during the main experiment to avoid noise in the collected signals. In cases involving a real plank, participants were reminded of its presence for safety and told that the researcher would guide them to avoid it during the VR session.

Participants were informed that total five standing sessions were involved, each session would last three minutes, data would be processed anonymously, and that participation was voluntary, with the option to stop the experiment at any time. The researcher then marked the start of the experiment with the first event marker start (eleven event markers were used in total: “start”, “baseline1”, “baseline2”, “going up”, “door open”, “plank”, “going down”, “door open”, “recovery1”, “recovery2”, and “end”). During standing sessions, the researcher monitored the signals on the computer and VR images and made notes sometimes, without disturbing the participant unless necessary. Participants were instructed to stand upright naturally, keep their hands naturally at their sides, minimize movement, and avoid pressing the electrodes during standing sessions.

Participants performed three slight squats and jumps before the first session began. When they were ready, the researcher made a three-minute timer and marked baseline1 simultaneously in event palette of Acqknowledge. After the first baseline recording, participants rated their subjective stress level at that moment in survey (Q15) and donned a VR HMD device with the researcher's assistance. Then, the volume of sound effect was confirmed with participants.

When participants were ready, the researcher made another three-minute timer and event marker baseline2. Following the second baseline recording, participants were guided to avoid wires (and a plank) in reality and enter the virtual elevator. The researcher handed a controller to the participant for pressing the elevator button, and then took it back once the participant finished pressing. The researcher made the event marker going up when hearing the sound effect of door closing, and marker door open when hearing the sound effect of door opening. After arrival at plank floor, the researcher reiterated the option of reporting the feeling of uncomfortable or the decision of withdrawing at any time. Then, participants walked on to the plank and tried to achieve the end (not all participants stood at the end of the plank, some stood in the middle, some stood inside the elevator depending on acceptability). When participants were ready, the researcher made the third timer and event marker plank.

Following completion of the plank task, participants were guided to return the elevator, take the controller, and press the ground floor button. The researcher made the event marker going down when hearing the sound effect of door closing, and marker door open when hearing the sound effect of door opening. After arrival at ground floor, the researcher took away the controller and navigated participants walking back to the street. When participants were ready, the researcher made the fourth timer and event marker recovery1. Following this, participants removed the HMD device with the help of the researcher, and started the final standing session. The researcher made the fifth timer and event marker recovery2. After completion of the second recovery recording, participants rated their subjective stress level during the plank task and at that moment in survey (Q16&17). After participants completing the rest of the questionnaire (Q18–Q30), the researcher removed the electrodes, debriefed the participants, and thanked them for their participation. The detailed instructions and experiment protocol are attached in the appendix 2.1 and 2.3.

Time Synchronization

BioNomadix and VU-AMS systems record data based on the computer's time, while the EmbracePlus system uses the smartphone's time. Since the time servers and synchronizing timings for the computer and smartphone are different, this can cause time drifts between the devices. Before each experiment, the researcher ensured that the time zones were correctly set on both the computer and the smartphone. Additionally, the VU-AMS system's time was synchronized by a click on "Set Device Time to Computer Time" button in VU-DAMS software prior to every experiment.

All three systems included an accelerometer. The BIOPAC system used the TPS 2g accelerometer, while EmbracePlus and VU-AMS had built-in accelerometers. The EmbracePlus wristband featured a three-axis microelectromechanical accelerometer with a ±16g range, a sampling frequency of 64 Hz, and a resolution of 16 bits. The BIOPAC system's accelerometer, designed to detect subtle movements, operated on a smaller scale. This study explored the feasibility of using accelerometer data for time synchronization, though the findings were not reported in this thesis.

Thermoregulation

To investigate the SRQs, room and skin temperatures were collected. Ambient temperature was recorded by a SwitchBot, a commercial thermometer with a resolution of .2 °C and a sampling rate of 1 sample per minute. Peripheral skin temperature was measured using a Texas Instruments TMP117M digital temperature sensor embedded in the EmbracePlus wristband, which has a resolution of .01 °C and a sampling rate of 4 Hz. The SwitchBot thermometer was typically placed on a position of the computer desk near the VR game play zone.

Data analysis

Data Quality Assessment

Following the recommendations of Van Lier et al. (2020), two researchers conducted a visual inspection of the EDA data to identify and filter out failed measurements. Signals without any SCRs were considered as missing data or zero responses. If either researcher classified the data as non-responding, the entire dataset was excluded. The exclusion criteria were set as follows, based on previous studies (Braithwaite et al., 2013; Van Der Mee et al., 2021; Van Lier et al., 2020): SCL lower than .5 μS, SCR amplitude lower than .01 μS, and visual plot shown a flat line. Obvious artifacts appeared in the data visualization were removed using the AcqKnowledge "Connect Endpoints" function, which performed a linear interpolation between the left and right edges of selected area (BIOPAC Systems, Inc., 2020). Participants who signed the informed consent but did not complete the experiment were classified as withdrawers.

Signal Comparison—Detrended Cross-Correlation Analysis

Following Van Lier et al. (2020)'s standardized validation protocol, the first step involved signal comparison using the cross-correlation function. In this study, detrended cross-correlation analysis (DCCA) was applied to test the correlation between two signals over time (Podobnik & Stanley, 2008; Shen, 2015). Within the DCCA functions, the data was normalized using mean centering and detrended using linear polynomial fitting to ensure the data was stationary and comparable. Data from the BioNomadix and EmbracePlus systems were first resampled to 16 Hz. Boucsein (2012) noted that downsampling EDA data to frequencies higher than 10 Hz does not enhance data quality. The choice of 16 Hz was due to it being a multiple of 4 and a factor of 2000 more than 10, ensuring adequate resampling.

The two signals were then aligned and inputted into the Python “DCCA” library with a detrending scale of 2880 and a time lag setting of (-8, 8) to compare correlation at multiple lags, aiming to find the optimal cross-correlation with the corresponding lag (Python scripts are attached in appendix 3.5). The time lag range was decided based on the time estimation of nerve conduction through the distance from wrist to finger (Van Lier et al., 2020). The selection of the detrending scale at 2880 (the sample number of three minutes) was a considered decision based on multiple factors. The detrending scale was chosen based on visual fluctuation patterns, as a time scale set too short would insufficiently remove larger trends. For EDA signals in an interval-related experimental design, obvious fluctuations are typically related to task duration. Therefore, the most direct fluctuations can be observed from session to session. Each static session lasted three minutes, while the dynamic sessions (interspersing between each pair of static sessions) had non-fixed durations. The researcher experimented with detrending scales of 3, 3.5, and 4 minutes and observed minimal differences between them. Thus, a duration of one standing session seemed sufficient for capturing the most significant fluctuations. Ultimately, a detrending scale of 3 minutes was chosen for the analysis.

Parameter Comparison—The Bland–Altman Plot

The second step involves parameter comparison using the Bland–Altman plot, a widely used analysis method in validation studies (Giavarina, 2015). The Bland–Altman plot visually represents the agreement between two devices on a particular parameter over the averages and differences of measurements. Prior to plotting, the raw signals was resampled and aligned as in the signal comparison, then decomposed into three parameters.

This study used the Python "Neurokit2" library (Makowski et al., 2021) with its default smoothing, decomposing, and finding peaks settings to process the raw signals. Retrieved parameters included the mean SCL, the SCRs per minute, and the AMPL per minute. Normality were visually checked. Amount of missing data was verified. The Shapiro–Wilk test (Shapiro & Wilk, 1965) was conducted to calculate the statistical confidence to reject the normality assumption. If the data was not normal, a log transformation was applied. For the Bland–Altman plot, the mean of the two measurements was plotted on the x-axis, and the difference between the two values on the y-axis. The acceptable boundaries suggested by Van Lier et al. (2020) are 10 percent of the parameter's empirical range (±1.6 μS for SCL, ±2.5 for SCRs, ±.6 μS for AMPL).

Event Detection Comparison—Error Bar Plots

The event detection comparison involved an error bar plot, which visualizes the mean effect and its confidence interval. The study focused on a self-designed environmental stimulus lasting three minutes, where a significant difference from the baseline in AMPL was expected. Data was resampled and aligned similarly to the signal comparison. Tonic and phasic EDA parameters were extracted using the Neurokit2 library, as in the parameter comparison. AMPL was calculated by summing the amplitudes of SCRs every minute for static sessions and combined dynamic sessions, excluding the last group if it was shorter than a minute. For each dynamic session, the AMPL was estimated by multiplying the amplitude per sample by the sample number per minute (an arithmetic estimation). This approach was necessary because some dynamic periods were shorter than a minute. However, since AMPL is not evenly distributed over time, direct comparison between each dynamic session and static session is not valid. Therefore, the results within each calculation were examined separately to ensure meaningful analysis.

Normality was visually checked, and log transformations were applied if necessary, similar to the parameter comparison. Two line plots with means and errors were created, each representing a device, to visualize the effect captured by the two devices. If both plots showed non-overlapping error bars between the stress task and baseline, a plot with the means and errors of differences between the two devices would be created with predefined boundaries (difference between the stress task and baseline recording from the BioNomadix). If both devices captured the event effect significantly and all error bars in the difference event plot overlapped with the zero axis, a Bland–Altman-like analysis with predefined boundaries would be conducted to investigate the wearable's capability to distinguish between the baseline and the stressor.

Explorations and Post Hoc Tests

Since the method of signal comparison is applicable exclusively to whole time series, the differences between dynamic and static periods were investigated on the rest two levels concurrently when analyzing the main research questions. A post hoc test using the detrended partial cross-correlation analysis (DPCCA) was conducted to explore the influence of temperature. DPCCA, derived from DCCA, allows for the investigation of multiple non-stationary time series in complex systems and reveals intrinsic relations between the considered signals (Yuan et al., 2015). The codes for conducting DPCCA were sourced from the study by Ide et al. (2017).

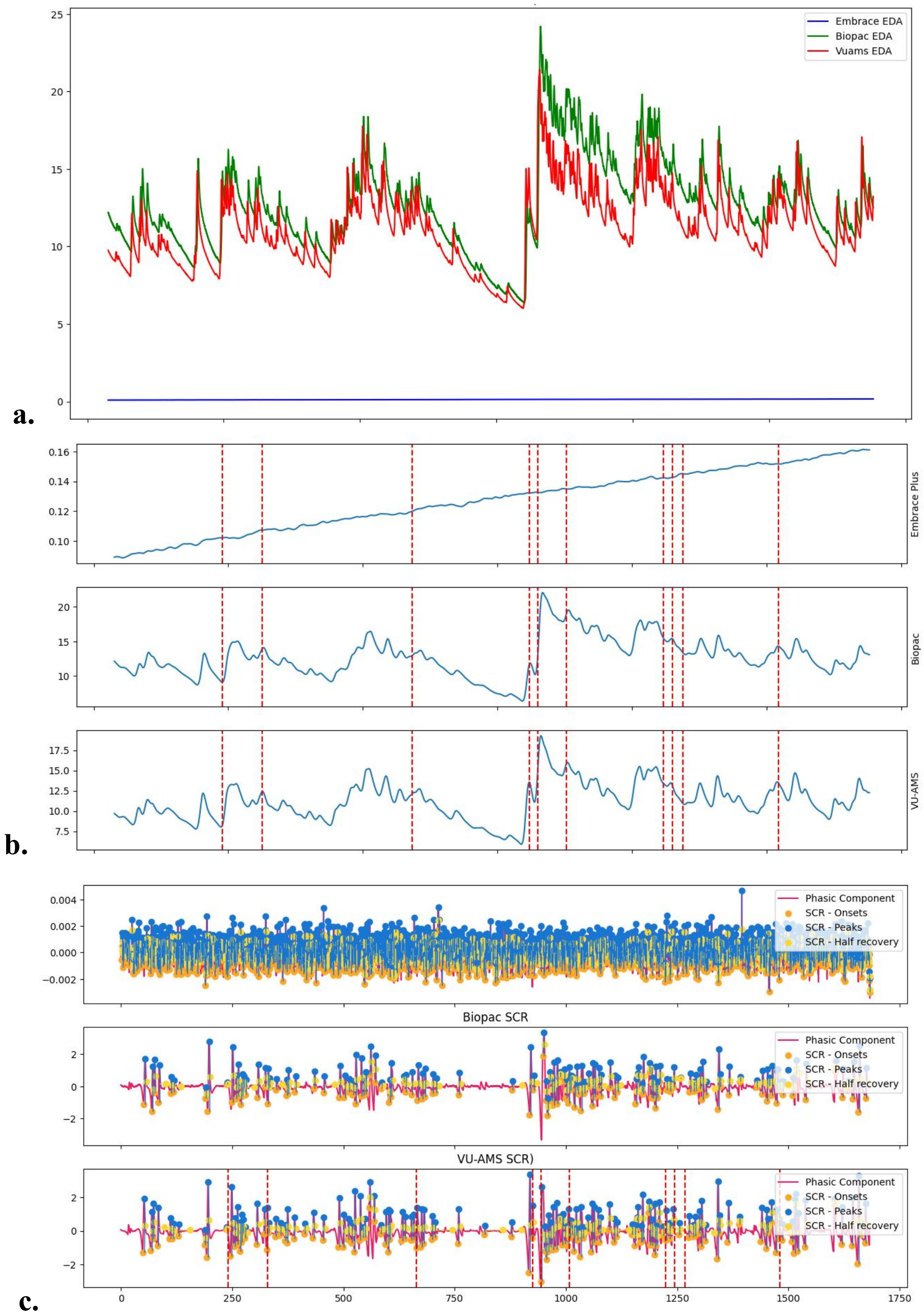
Results

Data Quality Assessment

Of the 31 participants, 2 (6.5%) withdrew from the study—one due to dizziness and the other due to technical issues. The data from the remaining 29 participants were visually inspected. During this process, 8 participants (27.6%) were identified as non-responders by two coders who reached a consensus. These non-responders were removed from the dataset, resulting in a final sample of 21 participants (achieved the planned final sample size) for the research. Figure 14 illustrates an example of excluded signals, showing visual presentations from three devices (VU-AMS signals were visually inspected but not further analysed). In this example, the EmbracePlus signal displayed a flat line very close to the x-axis and parameter values under the predefined thresholds, indicating a non-response.

Figure 14

An Example of a Non-responder's Signals.



Note. a. Raw signals on the same scale; b. Tonic signals with specific scales; c. Phasic signals with specific scales.

During the visual inspection, signals with missing values or parameters outside the empirical range were checked. No missing values were found in the raw signals. However, three participants had missing values in their wearable-retrieved amplitudes, with a missing rate averaging 1.7% (.7%–3.1%), while another two participants had missing amplitudes in RD signals, both around .6%. For the wearable, on average 17.3% (0%– 87.5%) of amplitudes are lower than .01 µS, all within the empirical range (0–3 µS). For the RD, on average 3.4% (0%–35.4%) of amplitudes are lower than .01 µS, 10.0% (0%–49.5%) higher than the empirical range (0–3 µS). No zero amplitudes were found. Missing values were ignored in further analysis.

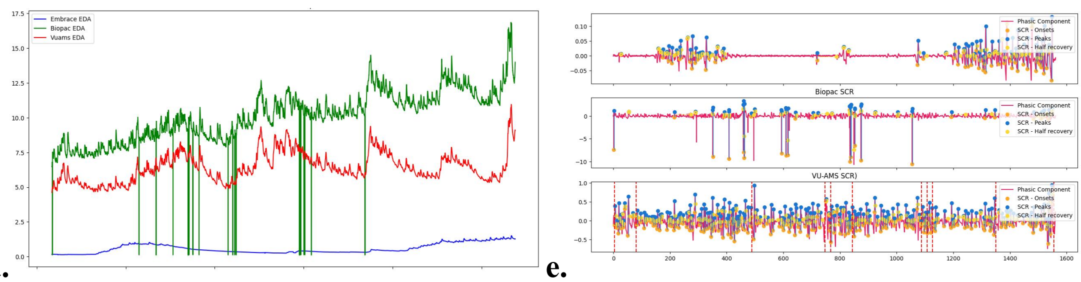
The SCRs were all in the empirical range of (0–25 per minute), 29.3% (0–81.8%) of wearable retrieved SCRs and 5.7% (0–22.2%) of RD retrieved SCRs were zero. There was no missing values for SCRs. All the SCL values of wearable device were in the empirical range of (0–16 µS). 59.1% (0–100%) of the reference retrieved SCL values were higher than the range. No zero or missing values were found for SCL.

After excluding non-responsive participants, four participants' signals were identified with significant artifacts (visible sudden drops in raw signals or abnormal troughs in phasic signals). These artifacts were removed. Figure 15 illustrates the four signal sets with artifacts.

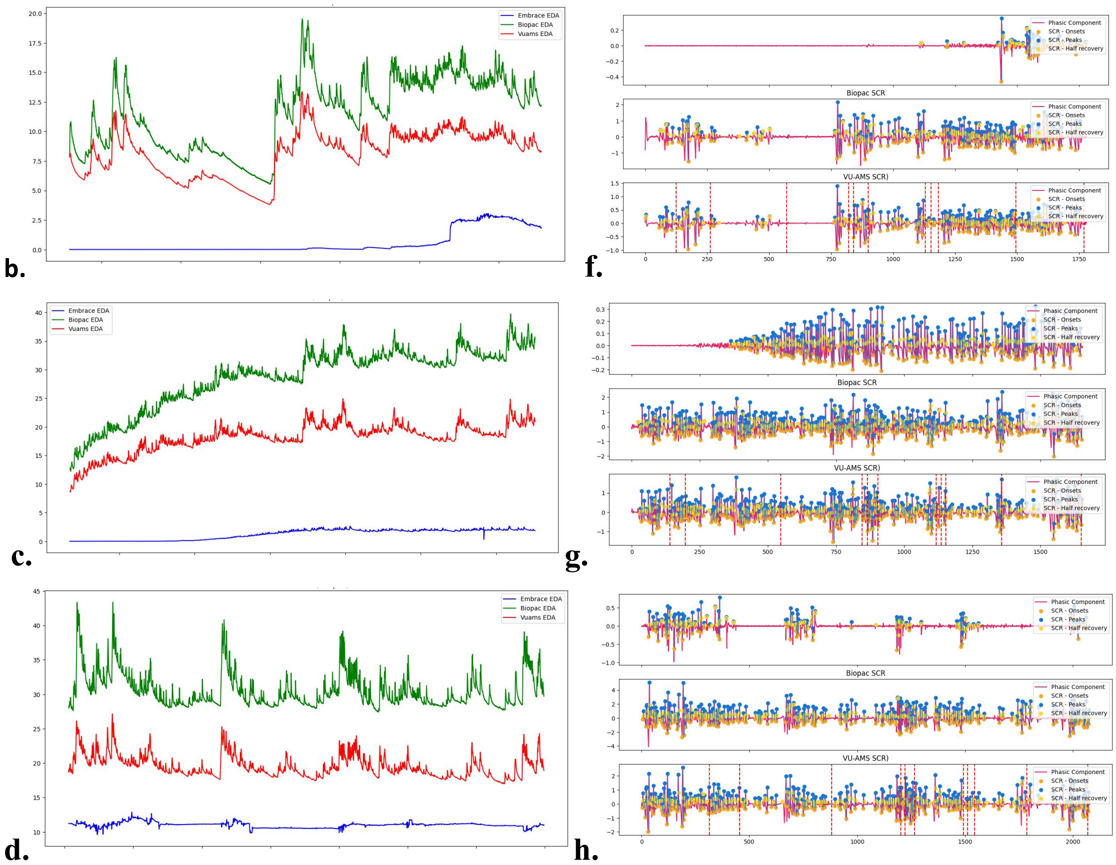
Figure 15

Four Sets of Signals With Visible Artifacts.

a.



e.



Note. Each row of pictures represents a signal set from a participant. Left column of pictures (Figure 15.a., b., c., d.): raw signals' comparison; Right column of pictures (Figure 15.e., f., g., h.): phasic signals' comparison.

[a]The first row (Figure 15.a., e.) showed artifacts in RD's raw and phasic signal. [b]The second row (Figure 15.b., f.) showed an artifact in EmbracePlus phasic signal. [c]The third row (Figure 15.c., g.) showed an artifact in EmbracePlus raw signal. [d]The fourth row (Figure 15.d., h.) showed artifacts in EmbracePlus raw and phasic signal.

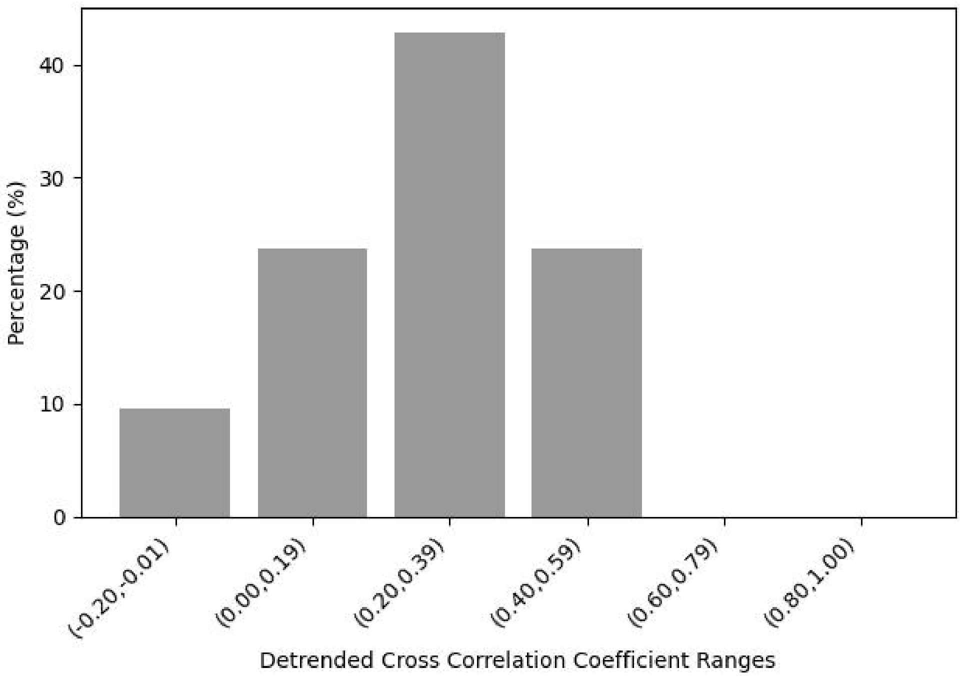
Signal Comparison: DCCAs

For the majority (76.2%) of participants, the correlations were considered low (i.e., below .4), with two participants' signals showing a very weak negative correlation (.05, .005). For 23.8% of participants, the correlation coefficients were moderate (between .4 and .6) (see Figure 16). On average, the cross-correlation was .27, ranging from -.05 to .54. Similar to

Van Lier et al. (2020)'s findings, the optimal correlations were at different lags across participants (Appendix 3.3). Visual inspection of the accelerometer signals showed that the time drifts for each participant did not follow identical trends as well.

Figure 16

Distribution of the Optimal DCCAs at the Time Scale of 2880



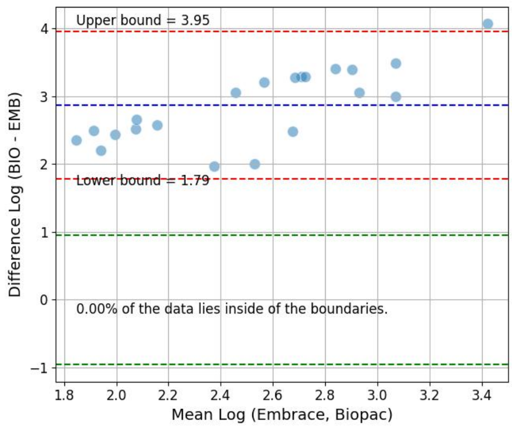
Note. The coefficients for each participant presented in the figure were determined on the basis of the optimal lag.

Parameter Comparison: Bland–Altman Plots

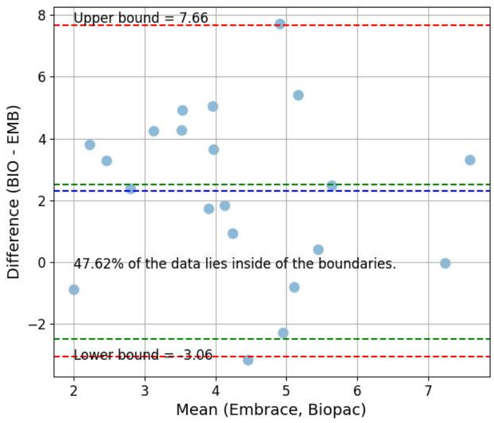
Bland–Altman plots for the mean SCL, number of SCRs, and AMPL of the entire EDA signal are shown in Figure 17. SCL and AMPL were log-transformed to achieve normality (with a small constant of 1e-8 added to handle zero values). The number of SCRs was plotted without normalization. The plots show that 47.62% of SCRs and their mean line lie within the pre-set boundaries. For SCL and AMPL, the percentages are 0% and 4.76%, respectively. Mean biases are found in SCL and AMPL, as all differences are above zero, even the lower LoA lie above the higher boundary. It seems the wearable systematically underestimated SCL and AMPL. Additionally, an ascending trend is presented in the plots for SCL and AMPL, indicating that the signals' differences tend to expand when their means are higher.

Figure 17

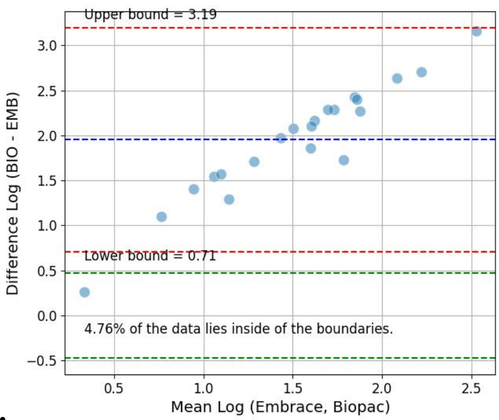
Bland–Altman Plots across Parameters



a.



b.



c.

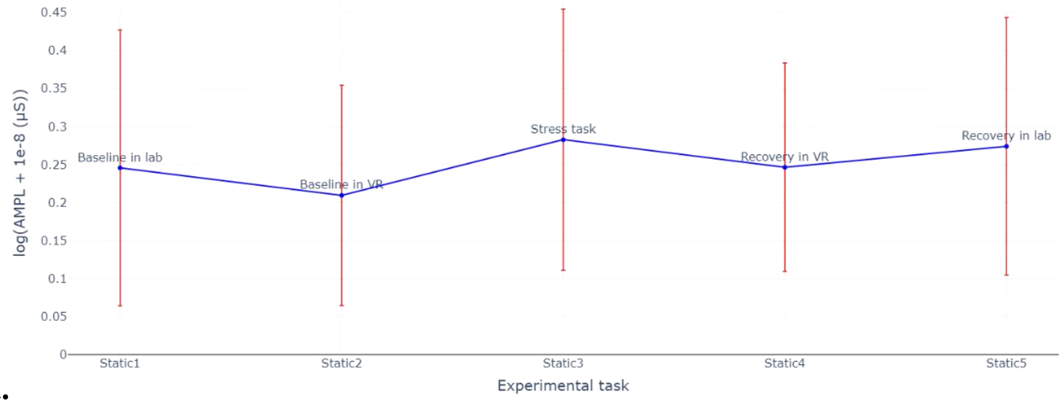
Note. a. SCL plot; b. SCR plot; c. AMPL plot. Each dot represents one participant. The difference between and the average of the two measures are represented on the y-axis and the x-axis, respectively. The blue line shows the mean bias. The green lines represent the priori chosen acceptable boundaries, whereas the red lines (also marked with upper bound and lower bound) represent the actually found 95% LoA. At the bottom of each figure, the percentage of values within the proposed boundaries is given. SCL, skin conductance level; SCRs, number of skin conductance responses per minute; AMPL, total amplitude of the skin conductance responses per minute.

Event Detection Comparison: Error Bar Plots

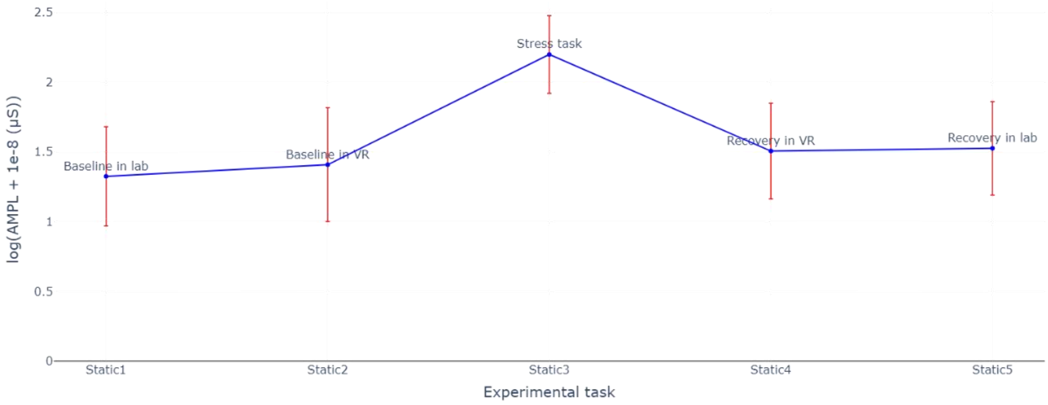
The main stress task used to determine stress reactions was the third static session or plank session, where participants stood on a virtual plank 80 stories high. Figure 18 graphically represents the five static sessions for EmbracePlus and RD, respectively. The RD showed the expected increase in AMPL during the stress task, with approximate increases of 70% compared to the baseline in reality, 60% to the baseline in VR, 50% to the recovery in VR, and 40% to the recovery in reality. In contrast, much lower changes were observed for the EmbracePlus wearable, that were 20%, 30%, 20%, 3%.

Figure 18

Error Bar Plots of Static Sessions



a.



b.

Note. Error bar plot for static sessions including the plank task, measured with the RD and EmbracePlus wearable, with AMPL as the parameter of interest. a. Plot of EmbracePlus. b. Plot of the RD. the means are plotted in blue, and its error bars are plotted in red. The AMPL was log transformed due to rejection of normality. The first two static sessions are neutral baselines, the third static session is the VR stress task, and the fourth and last static sessions are recovery recordings (considered also as baselines).

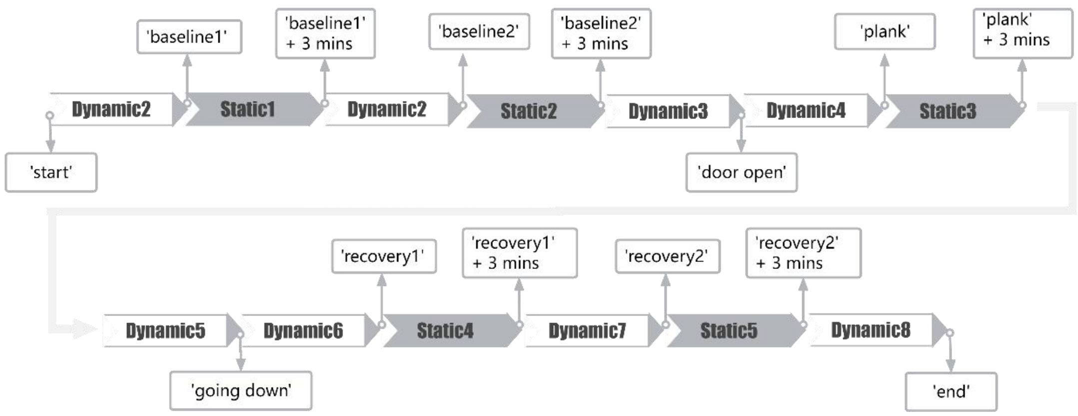
Figure 18 shows obvious overlap between the error bars for all standing sessions in the wearable's data, while non-overlapping error bars between the stress task and baseline were observed in the RD's data. This indicates that the wearable did not detect the stressor effectively, whereas the RD did. As a result, further analysis on this level was not conducted. Error bar plots of SCL were inspected due to observed bimodality in visual inspection for distributions (Appendix 3.7), whereas no significant effects were found (Appendix 3.2), reinforcing Van Lier et al. (2020)'s decision to prioritize AMPL as the event-level indicator.

Comparison Between Static and Dynamic Sessions

At the parameter level, no significant differences were observed between the whole signal, signals within static sessions, and signals within dynamic sessions, as shown in Appendix 3.1. Further exploration was conducted at the event detection level. Figure 19 illustrates the exact time points used to divide the static and dynamic sessions. Figure 20 provides a graphical representation of the five static sessions, the combined dynamic sessions, and the separate dynamic sessions for both devices.

Figure 19

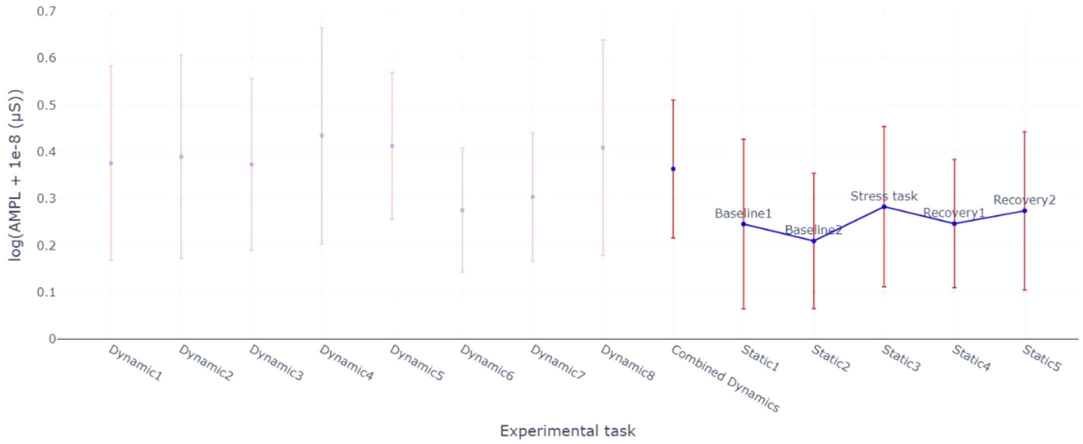
Time Points for Dividing the Static and Dynamic Sessions



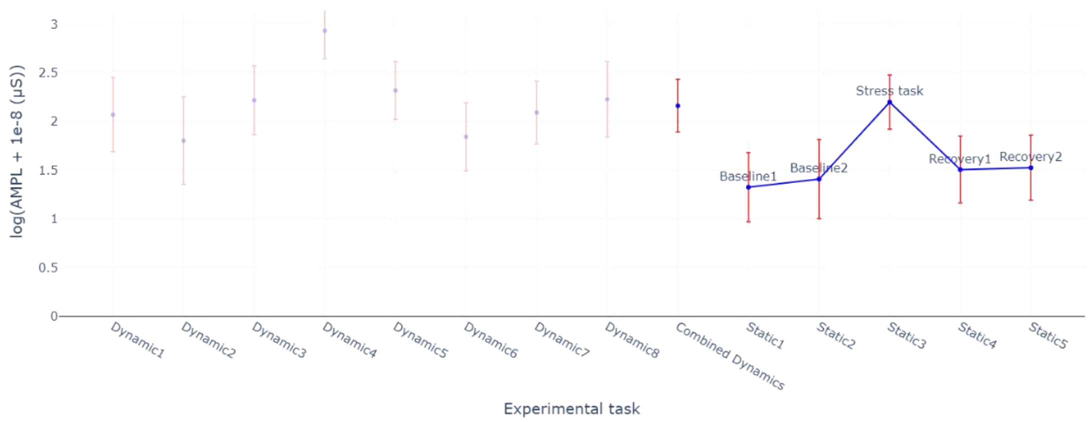
Note. Words in quotations represent event markers. Markers are made upon either the beginning or the end of a session.

Figure 20

Error Bar Plots for All Sessions



a.



b.

Note. Error bar plot for all sessions, measured with the RD and EmbracePlus wearable, with AMPL as the parameter of interest. a. Plot of EmbracePlus. b. Plot of the RD. The means are plotted in blue, and its error bars are plotted in red. The AMPL was log transformed due to rejection of normality. The shaded areas show error bars of AMPL's arithmetic estimations for each dynamic period. The error bar of combined dynamic session is added to the non-transparent side of the plot.

In the half-transparent area of Figure 20, the arithmetic estimation of AMPL per dynamic period shows similar patterns to the static sessions. For the RD's data, the error bar for Dynamic 4 (from the door opening to the start of plank standing, involving the height stressor but also conversations and movements) does not overlap with the error bars of the other dynamic periods, indicating a distinct response. In contrast, the error bars for each dynamic period overlap in the EmbracePlus data, indicating a lack of detectable effect of the stressor.

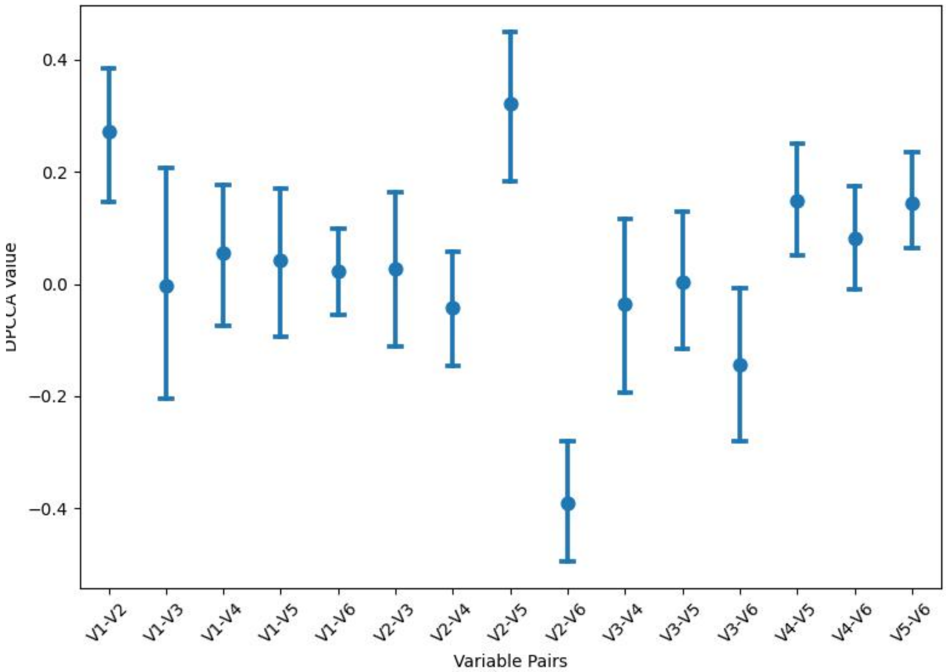
The combined dynamic session was compared with static sessions based on the same calculation method for AMPL. For the entire dynamic period, the error bar for RD's data does not overlap with baselines (Static 1, 2, 4, 5), suggesting a detectable difference between dynamic and static baselines. However, for EmbracePlus data, the error bar overlaps with all sessions, providing little information about the difference between the dynamic periods and static baselines.

Temperature's Influence

Due to technical issues of the thermometer, ambient temperature recording was completed for only 14 participants (details in Appendix 2.2). The mean DPCCA between two EDA signals (.27, with 64.3% below .4) was similar with the mean of DCCA coefficients (.27). The correlation between RD signal and the occurrence of stressor also showed a weak correlation (average .32, with 64.3% below .4). The correlation between RD signal and the occurrence of static sessions showed a negative weak correlation (-.39, with 50% below .4). The DPCCA coefficients between each pair of six variables were listed in Appendix 3.4. Figure 21 illustrates that the estimated correlation range for V1–2 (two EDA signals), V2–5 (RD signal and stressor) and V2–6 (RD signal and static session) overlap less with the rest.

Figure 21

The Error Bar Plot of DPCCAs across Six Pairs of Variables



Note. V1 = Raw signals from EmbracePlus. V2 = Raw signals from the RD. V3 = Wrist temperature. V4 = Ambient temperature. V5 = Dummy variable of stressor. V6 = Dummy variable of static periods.

Sample Composition

As shown in Table 1, data from 21 participants (13 male, 8 female; mean age: 23 years, SD = 4.1) were analysed. Twelve participants (57.14%) with removing overlap did not meet the ideal eligibility criteria due to factors like medication intake, coffee consumption, or inadequate sleep. Two participants had incomplete data for the second recovery recording (Static session 5), one due to dizziness and the other due to experimenter error. Considering the small sample size, these participants were included in the analysis. The current study solely described the sample composition without conducting further analysis due to the small scale of groups. Visual inspections (Appendix 2.5) and descriptive statistic comparisons (Table 1) revealed no significant differences between subgroups. Inferential statistical comparisons (Shapiro-Wilk, Levene's, independent t-tests, Mann-Whitney U tests) could be conducted for in-depth analysis with larger samples in future.

Table 1

Sample Composition According to Features and Conditions

Features	Statistics			
	n	%	M	SD
Gender				
Female	8	38.1	.27	.21
Male	13	61.9	.26	.16
Race				
Black or Africa	0	0		
Asian	4	19.1		
White	14	66.7		
Other (Persian, Croatian)	2	9.5		
None	1	4.8		
Follicular phase(% of females)	4	50		
Medication intake	3	14.3		
Drank coffee prior	4	19.1		
Slept less than 6 hours last night	4	19.1		
Without VR experience	3	14.3		
Rating similar stress levels for the plank, baseline, and recovery sessions	2	9.5		
Rating higher stress levels for the baseline than for the plank session	3	14.3		
Played the game or other similar games prior	5	23.8		
Diagnosed with acrophobia/basophobia)	0	0		
Features (unit, range; actual range)	Statistics			
	M	SD		
Age (year, -; 18–32)	23	4.1		
Fear of the height (score, 0–100; 5–85)	33.6	24.5		
Perceived Stress Scale (score, 0–40; 2–27)	16.4	6.7		
Conditions	Statistics			
	n	%	M	SD
Real plank				
Yes	10	47.6	.33	.16
No	11	52.4	.21	.18
Dominant hand				
Right	15	71.4	.27	.18
Left	6	28.6	.23	.18
Placement of reference devices				

Finger placement	11	52.4	.28	.20
Eminence placement	10	47.6	.25	.16

Note. N = 21. n = count; % = percent; M = mean of correlations; SD = standard deviation of correlations. When the subgroup size was smaller than 5, the M and SD were not reported.

Discussion

Wearable technology offers new opportunities for stress management in daily life. Palmar or plantar areas were preferred for laboratory EDA measurement due to their sensitivity to psychophysiological stimuli. However, these sites are impractical for wearable applications due to their interference with daily activities. Consequently, the wrist has emerged as a more viable option. This study aimed to assess the validity of the EmbracePlus wristband for measuring EDA responses to a VR stressor by comparing it to an RD using a standardized protocol. The goal was to establish the wristband's suitability for real-world stress assessment, where stationary equipment is infeasible. To provide insights into the main results, the experiment incorporated dynamic and static conditions, as well as temperature measurements to examine the impact of movement, interaction, and temperature on EDA.

Primary Goals: Findings Across Three Levels of the Protocol

RQ1: Signal Comparison

Findings from this preliminary study involving approximately 20 participants suggest that the EmbracePlus wristband does not capture EDA signals equivalent to those from an RD using a self-designed VR stressor. Raw signals exhibited consistently low correlations across all participants, falling below the accepted validity threshold of .8, aligning with Hu et al. (2024) and Van Lier et al. (2020) regarding the nonsignificant cross-correlation between E4 and RDs. This is consistent with other previous E4 validation studies that reported negative results, despite using different correlation types or testing parameter correlations (Borrego et al., 2019; Costantini et al., 2023; Menghini et al., 2019; Milstein & Gordon, 2020). While some participants' signals showed visual similarities (Appendix 3.6), these instances were rare, aligning with Menghini et al., Ollander et al. (2016), and Van Lier et al.'s findings of none or weak visual resemblance. The observed low correlation is likely due to several factors: distinct sweating patterns between wrist and palm, variations in gland density, asymmetry between the two wrists, and differences in electrode materials and sizes, as previously discussed.

RQ2: Parameter Comparison: Systematic Biases

Based on the protocol's criteria, the EmbracePlus wristband may be unsuitable for capturing EDA parameters in a VR lab task. Most participants exhibited significant and consistent lower wrist SCL and AMPL values compared to palm measurements, similar to Costantini et al. (2023) and Hu et al. (2024)'s findings of a systematic underestimation of SCL by E4. While no consistent difference in SCRs between devices was observed, the majority of wrist-based SCRs were smaller than those recorded by the reference device, similar to Hu et al. and Van Lier et al. (2020)'s reports. This suggests that SCRs are more stable across different stressors, possibly tracking SNS activity independent of thermoregulatory needs, replicating prior findings (Dementienko et al., 2000; Hu et al., 2024; Menghini et al., 2019; Van Der Mee et al., 2021).

The observed systematic bias, where differences in SCL and AMPL tended to increase with their mean values, aligns with Hu et al. (2024) 's findings for SCL but contrasts with Van Lier et al. (2020) 's findings of no systematic biases. This discrepancy may be attributed to the task type and exposure duration. Both this study and Hu et al.'s validation used tasks with minimal social pressure, while Van Lier et al. employed an intensive social stressor, sing-a-song stress test (SSST). Wrist EDA may exhibit a proportionally weaker response compared to palm EDA when transitioning from social stressors to tasks with minimal social challenges. Additionally, the longer task durations (>2 min in this study and Hu et al.'s study) compared to SSST's one-minute exposure may have enabled the detection of this systematic bias.

RQ3: Parameter Comparison: LoA and Boundaries

Ideally, LoA values should fall within predefined boundaries as per protocol guidelines. However, in this study, the LoA for SCL and AMPL were wide and did not overlap with the boundaries, similar to the findings of Van Lier et al. (2020) and Costantini et al. (2023). The stringent 10% difference boundary, adopted from heart rate equipment standards, might be overly restrictive for EDA data, as suggested by Van Lier et al. In contrast, the LoA for SCRs overlapped with the acceptable range, again reflecting the relative stability of SCRs compared to SCL and AMPL and aligning with previous studies (Hu et al., 2024; Menghini et al., 2019; Van Der Mee et al., 2021).

RQ4: Event Detection Comparison

The EmbracePlus wristband demonstrated limited ability to reliably detect stress responses to a sustained, intensive VR task minimizing movement and interaction, unlike the RD. This contradicts previous social stressor research suggesting wrist EDA's sensitivity to high-intensity stressors (Ollander et al., 2016; Van Lier et al., 2020). As discussed in RQ2, the varying levels of social pressure involved in these studies may explain this discrepancy.

Secondary Goals: the Influence of Thermoregulatory Factors

SRQ1: Dynamic Factors

The wrist device did not detect significant EDA differences between dynamic and static periods, suggesting reduced sensitivity to dynamic conditions. Contrary to the initial hypothesis, these dynamic conditions showed more pronounced impacts on palm EDA than wrist EDA. This indicates that natural interactions and movements can augment palm-retrieved EDA responses to stressors involving these factors. This finding aligns with and may provide insight into previous research highlighting lower wearable validity in dynamic tasks (Hu et al., 2024; Menghini et al., 2019; Milstein & Gordon, 2020). Palm devices appear to benefit from a synergistic effect between dynamic factors and other components of stressors on EDA, outperforming wrist devices.

SRQ2: Temperature

The anticipated significant difference in temperature's influence on the two signals was not observed in the post hoc test. While the weak impact on palm signals aligned with expectations, the lack of a significant effect on wrist EDA may be attributed to insufficient temperature fluctuations and the limited ability of the temperature indicators to accurately reflect thermoregulatory changes, as discussed in the "Current Validations" section. It's possible that the temperature changes induced by either the VR stressor or dynamic factors were insufficient to trigger noticeable wrist EDA changes.

In summary, the low validity of EmbracePlus found in this study largely aligns with previous E4 validation results. However, the cross-device differences cannot be solely attributed to device performance. Physiological characteristics of the sensor placement (wrist vs. palm) also play a significant role. If we consider the earlier proposed EDA response model (Figure 1), some interesting inferences can be drawn.

Synthesizing the outcomes of RQ2-4, SRQ1-2, and mentioned previous studies, while VR stressors and social stressors both caused detectable reactions in palm EDA, social stressors appeared more influential on wrist EDA. Social stressors are known to effectively elicit acute psychosocial stress. The VR-based stressor used in this study, while involving a high-altitude threat, minimized social challenges, interaction, movement, and facial exposure. If the proposed model is correct, palm EDA is primarily caused by psychological changes. Both stressors likely induced significant psychological responses. However, the weak wrist EDA response to the VR stressor suggests a limited stimulation of psychological changes on the wrist. Therefore, thermoregulatory changes may have been the primary driver of wrist EDA.

The effects of psychological and thermoregulatory changes are dynamic and can interact in complex ways, potentially neutralizing each other when they move in opposite directions and have similar magnitudes. If social stressors consistently produce noticeable changes in wrist EDA, this suggests that they may have induced significant thermoregulatory changes that overshadow the psychological effects on wrist EDA. These inferences imply that social pressure might lead to more substantial thermoregulatory changes, resulting in significant wrist EDA alterations. Conversely, artificial high-altitude stress may not generate sufficient thermoregulatory responses to elicit robust wrist EDA. This inference also extends to dynamic factors. However, these ideas cannot be supported as the current study did not find direct evidence for significant effects of thermoregulation on the wrist.

Data Quality and Experiment Feasibility

The wearable device exhibited a non-responder rate of 27.6% (8/29), between the 22% rate reported by Van Lier et al. (2020) and the 73% rate reported by Milstein and Gordon (2020) using the same criterion. Implementing preliminary physical activity tasks (stairs, squats, jumps) and a 15-minute wear-in period may have contributed to the relatively low non-responder rate. While these measures effectively lowered the rate, they did not entirely eliminate non-responders. It could potentially explained by Picard et al. (2016)'s finding on response latency (30–120 min) for a wearable device.

This study examined the sample composition and assessed the feasibility of different palm placements, dominant hand involvements, and real plank settings. The current sample size restricted the ability to conduct inferential analyses on subgroup differences. However, the collection of demographic data and the implementation of different experimental conditions were carried out smoothly, providing reference for future relevant studies.

The RD signals demonstrate the feasibility of the VR-simulated height-based stressor for EDA, further supporting VR's potential as a promising approach to recapitulate traditional standardized stressors (Van Dammen et al., 2022). Furthermore, accelerometer signals effectively captured time drifts across devices for most participants. Visual inspection indicated that the use of an additional reference device and accelerometer did not interfere with the main experiment.

However, several issues require attention. First, switching VR devices midway through the experiment introduced complications and extended the duration. This was primarily due to differences in controllers, helmets, and sound output. Second, real-time event markers required additional personnel and coordination, with synchronous wire management and computer marking proving challenging. Seeking participants’ cooperation to make event markers on portable devices could be a viable alternative. At last, despite instructions, some participants exhibited unintentional movements and electrode presses. Surprisingly, some unexpected activities did not create artifacts, as confirmed by visual inspections.

Strength and Limitations

This preliminary study introduced a novel height-based VR stressor, demonstrating its feasibility for validation studies. The stressor's design was tailored to specific research questions, allowing for a focused investigation of how both immersive static stressors and dynamic factors influence signals from two different devices. Notably, it suggested a synergistic effect between dynamic factors and other stressor components, potentially explaining previously observed lower validity in wearable devices during dynamic task (Hu et al., 2024; Menghini et al., 2019; Milstein & Gordon, 2020).

Employing a standardized protocol ensured comparable results across experiments. It also enhanced methodological rigor and statistical validity of this study, facilitating the identification of distinct EDA response patterns from different skin sites when exposed to different stressors. The replication of the majority of results from Van Lier et al. (2020)’s validation demonstrates the protocol's feasibility across contexts.

There are also a few limitations of our study that need to be taken into account while interpreting our results. First, some sampling issues might have restricted the generalizability of the findings. The participant pool, primarily consisting of university students and employees, may not adequately represent the broader population. Additionally, the small sample size (N = 21) fell short of the protocol's recommended sample size (55) and the average size of previous eight E4 validations (36). Furthermore, the inclusion of participants with suboptimal eligibility might have introduced unique features into the sample.

Second, The thermoregulation indicators used in this study, particularly the single local temperature, may be less relevant than some others recommended by current research (e.g., Fiala et al., 2012; Rubinstein & Sessler, 1990). This could potentially limit the reliability of the results. Ambient temperature varies slowly, and a single local temperature measurement may not accurately reflect body temperature changes under stress (Vinkers et al. 2013). Caution should be exercised when interpreting the findings based on these indicators.

Third, the statistical methods used for SRQs were subject to peer review. While focusing on the comparison of ranges rather than specific values helped mitigate the impact of the arithmetical estimation of AMPL of each dynamic session, the risk of misrepresentation still existed. Moreover, the accuracy of DPCCA for post hoc tests with dummy variables remains uncertain. Consequently, the error bar plot based on DPCCA should be interpreted cautiously.

Finaly, there were some limitations related to instruments. The reliance on an unvalidated stressor, despite providing some evidence of its validity, limits the strength of the study's conclusions. Additionally, the use of three simple rating questions instead of standardized instruments to assess perceived stress levels during the experiment may have resulted in a less precise and consistent understanding of subjective arousal. A similar concern applies to the simple inquiry about the dominant hand, which lacked the rigor of a standardized assessment. In addition, the accuracy of the commercial thermometer used to measure ambient temperature is uncertain, as the experimenter observed discrepancies when comparing it with another commercial thermometer in daily life.

Recommendations for Future Research

Moving forward, understanding wrist EDA's sensitivity to different stressors, including those encountered in daily life, is crucial for optimizing wearable technology. Future research could explore how validation results vary based on stressor type and severity, identify stress scenarios more detectable by wrist-worn devices, and test these wearables in everyday environments, where data quantity might compensate for data quality (Geus & Gevonden, 2024).

Additionally, developing gold standards, stressors, and protocols for wrist-site EDA measurement using wet electrodes in both laboratory and ambulatory settings could be a promising approach. This would enable more accurate comparisons of wrist-worn devices against a wrist reference, accounting for site-specific physiological differences. When comparing wrist devices to palm references, it's crucial to consider the influence of thermoregulation during validation. A deeper understanding of thermoregulation's role in wrist EDA can contribute to the development of more effective and reliable wearable devices for stress monitoring. To further investigate the potential correlation, future research could incorporate more precise indicators of thermoregulation, such as the sweating rate calculated from the UTCI–Fiala model equations (Fiala et al., 2012), or the skin-surface temperature gradients from forearm to fingertip (Rubinstein & Sessler, 1990). Infrared thermography offers another non-invasive option for continuous temperature monitoring.

Conclusion

While the EmbracePlus wristband offers a promising tool for real-time EDA monitoring, this study highlights its significant limitations in accurately capturing EDA signals and detecting stress responses during a height-based VR task that requires minimal movement and interaction. The observed synergistic effect of interaction and movement on palm EDA, rather than wrist EDA, provides an explanation for the previously noted lower reliability of wrist devices in dynamic tasks. Given the growing importance of effective stress management and the potential of ambulatory biofeedback, further research is needed to validate the EmbracePlus across various stressors, develop wrist-specific gold standards, and investigate the role of thermoregulation in EDA responses. Additionally, assessing the device's performance in naturalistic settings is crucial for its practical application in real-world stress management.

References

Agnihotri, A., Purwar, B., Jeebun, N., & Agnihotri, S. (2005). Determination of sex by hand dimensions. Internet Journal of Forensic Science, 1(2), 1–3.

Aspiotis, V., Miltiadous, A., Kalafatakis, K., Tzimourta, K. D., Giannakeas, N., Tsipouras, M. G., ... & Tzallas, A. T. (2022). Assessing electroencephalography as a stress indicator: A VR high-altitude scenario monitored through EEG and ECG. Sensors, 22(15), 5792. https://doi.org/10.3390/s22155792

BIOPAC Systems, Inc. (2020, November 19). Cutting Through the Noise—Simple Ways to Remove Artifact. https://blog.biopac.com/data-cleaning-how-to-remove-artifact/

BIOPAC Systems, Inc. (n.d.-a). BioNomadix wireless systems [Apparatus and software]. https://www.biopac.com/product-category/research/bionomadix-wireless-physiology/

BIOPAC Systems, Inc. (n.d.-b). MP160 System + wireless EDA with disposable electrodes. https://www.biopac.com/product/mp150-system-70/

Blatter, B., Houtman, I., van den Bossche, S., Kraan, K., van den Heuvel, S., Terwee, C., ... & Blonk, R. (2005). Health damage and costs as a result of RSI and psychosocial work pressure in the Netherlands. Hoofddorp: TNO.

Boccignone, G., Gadia, D., Maggiorini, D., Ripamonti, L. A., & Tosto, V. (2021, September). Predicting real fear of heights using virtual reality. In Proceedings of the Conference on Information Technology for Social Good. 103–108. https://doi.org/10.1145/3462203.3475882

Borrego, A., Latorre, J., Alcañiz, M., & Llorens, R. (2019, July). Reliability of the Empatica E4 wristband to measure electrodermal activity to emotional stimuli. In 2019

International Conference on Virtual Rehabilitation. https://doi.org/10.1109/ICVR46560.2019.8994546

Boucsein, W. (2012). Electrodermal activity. Springer Science & Business Media.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben–Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Society for psychophysiological research ad hoc committee on electrodermal measures. Publication recommendations for electrodermal measurements. Psychophysiology, 49(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Böttcher, S., Vieluf, S., Bruno, E., Joseph, B., Epitashvili, N., Biondi, A., ... & Loddenkemper, T. (2022). Data quality evaluation in wearable monitoring. Scientific reports, 12(1), 21412. https://doi.org/10.1038/s41598-022-25949-x

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments (Technical Report, version 2.0). Birmingham, UK: Selective Attention & Awareness Laboratory (SAAL) Behavioural Brain Sciences Centre, University of Birmingham. https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/psychology/ABriefGuideforAnalysingElectrodermalActivityv4.pdf

Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. Measuring stress: A guide for health and social scientists (pp. 235–283). John Wiley & Sons.

Costantini, S., Chiappini, M., Malerba, G., Dei, C., Falivene, A., Arlati, S., ... & Storm, F. A. (2023). Wrist-worn sensor validation for heart rate variability and electrodermal activity detection in a stressful driving environment. Sensors, 23(20), 8423. https://doi.org/10.3390/s23208423

Davies, C.T.M. (1979). Thermoregulation during exercise in relation to sex and age. Europ. J. Appl. Physiol. 42, 71–79. https://doi.org/10.1007/BF00421907

Dawson, M., Schell, A., & Filion, D. (2016). The Electrodermal System. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), Handbook of Psychophysiology (Cambridge Handbooks in Psychology, pp. 217–243). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781107415782.010

Dementienko, V. V., Dorokhov, V. B., Koreneva, L. G., Markov, A. G., Tarasov, A. V., & Shakhnarovich, V. M. (2000). The hypothesis of the nature of electrodermal reactions. Human Physiology, 26, 232-239. https://doi.org/10.1007/BF02760098

de Geus, Eco J. C., & Gevonden, M. J. (2024). Acquisition and Analysis of Ambulatory Autonomic Nervous System Data. In M. R. Mehl, M. Eid, C. Wrzus, G. Harari, & U. W. Ebner-Priemer (Eds.), Mobile Sensing in Psychology: Methods and Applications. Guilford Press.

Egger, M., Ley, M., & Hanke, S. (2019). Emotion recognition from physiological signal analysis: A review. Electronic Notes in Theoretical Computer Science, 343, 35–55. https://doi.org/10.1016/j.entcs.2019.04.009

Empatica Inc. (2019, November 19). Empatica Inc. Retrieved May 11, 2024, from https://www.empatica.com/blog/introducing-embraceplus-the-world-s-most-mobile-flexible-and-accurate-clinical-platform-in-neurology-research.html

Empatica Inc. (n.d.-a). EmbracePlus [Apparatus and software].

https://www.empatica.com/en-int/embraceplus/

Empatica Inc. (n.d.-b). Legal & Compliance. Retrieved May 11, 2024, from https://www.empatica.com/legal/

Empatica Inc. (n.d.-c). Meet EmbracePlus, the E4 Wristband’s Next-gen Successor. Retrieved January 31, 2024, from https://www.empatica.com/en-int/research/e4/

Facebook Inc. (2019). Oculus Quest 2. [Apparatus and software]. https://www.oculus.com/quest/refurbished/

Fiala, D., Havenith, G., Bröde, P., Kampmann, B., & Jendritzky, G. (2012). UTCI–Fiala multi-node model of human heat transfer and temperature regulation. International journal of biometeorology, 56, 429–441. https://doi.org/10.1007/s00484-011-0424-7

Firouz, N. (2023). VR Public Speaking Training: The Effectiveness of a Relaxation Micro-Intervention Training in Reducing Symptoms Related to Public Speaking Anxiety among Students in Higher Education (Master's thesis, University of Twente). https://purl.utwente.nl/essays/96739

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., & Venables, P. H. (1981). Publication recommendations for electrodermal measurements. Psychophysiology, 18(3), 232-239. https://doi.org/10.1111/j.1469-8986.1981.tb03024.x

Freedman, L. W., Scerbo, A. S., Dawson, M. E., Raine, A., McCLURE, W. O., & Venables, P. H. (1994). The relationship of sweat gland count to electrodermal activity. Psychophysiology, 31(2), 196–200.

https://doi.org/10.1111/j.1469-8986.1981.tb03024.x

Gerboni, G., Comunale, G., Chen, W., Lever Taylor, J., Migliorini, M., Picard, R., ... & Regalia, G. (2023). Prospective clinical validation of the Empatica EmbracePlus wristband as a reflective pulse oximeter. Frontiers in Digital Health, 5, 1258915. https://doi.org/10.3389/fdgth.2023.1258915

Giavarina D. (2015). Understanding Bland Altman analysis. Biochemia medica, 25(2), 141–151. https://doi.org/10.11613/BM.2015.015

Grasser, L. R., Erjo, T., Goodwin, M. S., Naim, R., German, R. E., White, J., ... & Brotman, M. A. (2023). Can peripheral psychophysiological markers predict response to exposure-based cognitive behavioral therapy in youth with severely impairing irritability? A study protocol. BMC psychiatry, 23(1), 926. https://doi.org/10.1186/s12888-023-05421-4

Hu, F., Wang, H., Chen, J., & Gong, J. (2018, August). Research on the characteristics of acrophobia in virtual altitude environment. In 2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR). https://doi.org/10.1109/IISR.2018.8535774

Hu, X., Sgherza, T. R., Nothrup, J. B., Fresco, D. M., Naragon-Gainey, K., & Bylsma, L. M. (2024). From lab to life: Evaluating the reliability and validity of psychophysiological data from wearable devices in laboratory and ambulatory settings. Behavior Research Methods, 1–20. https://doi.org/10.3758/s13428-024-02387-3

Ide, J., Cappabianco, F., Faria, F., & Li, C. S. R. (2017). Detrended partial cross correlation for brain connectivity analysis. Advances in Neural Information Processing Systems,

30.
http://papers.nips.cc/paper/6690-detrended-partial-cross-correlation-for-brain-connectivity-analysis.pdf

IJzerman, H., Heine, E. C., Nagel, S. K., & Pronk, T. M. (2017). Modernizing relationship therapy through Social Thermoregulation Theory: Evidence, hypotheses, and explorations. Frontiers in Psychology, 8, 635. https://doi.org/10.3389/fpsyg.2017.00635

Jansen, A. S., Van Nguyen, X., Karpitskiy, V., Mettenleiter, T. C., & Loewy, A. D. (1995). Central command neurons of the sympathetic nervous system: basis of the fight-or-flight response. Science, 270(5236), 644–646. https://www.science.org/doi/abs/10.1126/science.270.5236.644

Kerassidis, S. (1994). Is palmar and plantar sweating thermoregulatory?. Acta physiologica scandinavica, 152(3), 259–263. https://doi.org/10.1111/j.1748-1716.1994.tb09805.x

Kleckner, I. R., Feldman, M. J., Goodwin, M. S., & Quigley, K. S. (2021). Framework for selecting and benchmarking mobile devices in psychophysiological research. Behavior research methods, 53, 518–535. https://doi.org/10.3758/s13428-020-01438-9

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. Behavior Research Methods, 53(4), 1689–1696. https://doi.org/10.3758/s13428-020-01516-y

Menghini, L., Gianfranchi, E., Cellini, N., Patron, E., Tagliabue, M., & Sarlo, M. (2019).

Stressing the accuracy: Wrist-worn wearable sensor validation over different conditions. Psychophysiology, 56(11), e13441. https://doi.org/10.1111/psyp.13441

Meta Platforms, Inc. (2020). Meta Quest 2 [Apparatus and software]. https://www.meta.com/nl/en/quest/products/quest-2/

McFarland, R. A. (1985). Relationship of skin temperature changes to the emotions accompanying music. Biofeedback and Self-regulation, 10, 255–267. https://doi.org/10.1007/BF00999346

Milstein, N., & Gordon, I. (2020). Validating measures of electrodermal activity and heart rate variability derived from the empatica E4 utilized in research settings that involve interactive dyadic states. Frontiers in Behavioral Neuroscience, 14, 148. https://doi.org/10.3389/fnbeh.2020.00148

Morrison, V. & Bennett, P. (2022). Introduction to Health Psychology (5th ed.). Harlow, England: Pearson Education Limited.

Neupane, S., Saha, M., Ali, N., Hnat, T., Samiei, S. A., Nandugudi, A., ... & Kumar, S. (2024). Momentary Stressor Logging and Reflective Visualizations: Implications for Stress Management with Wearables. arXiv preprint arXiv:2401.16307. https://doi.org/10.1145/3613904.3642662

Ollander, S., Godin, C., Campagne, A., & Charbonnier, S. (2016, October). A comparison of wearable and stationary sensors for stress detection. In 2016 IEEE International Conference on systems, man, and Cybernetics (SMC). https://doi.org/10.1109/SMC.2016.7844917

Osborne, M., Glasser, S., & Loveridge, B. (2022). ‘It’s not so scary anymore. It’s actually exhilarating': A proof-of-concept study using virtual reality technology for music performance training under pressure. ASCILITE Publications, e22116–e22116. https://doi.org/10.14742/apubs.2022.116

Payne, A. F., Schell, A. M., & Dawson, M. E. (2016). Lapses in skin conductance responding across anatomical sites: Comparison of fingers, feet, forehead, and wrist. Psychophysiology, 53(7), 1084–1092. https://doi.org/10.1111/psyp.12643

Picard, R. W., Fedor, S., & Ayzenberg, Y. (2016). Multiple arousal theory and daily-life electrodermal activity asymmetry. Emotion review, 8(1), 62–75. https://doi.org/10.1177/1754073914565517

Podobnik, B., & Stanley, H. E. (2008). Detrended Cross-Correlation Analysis: A New Method for Analyzing Two Nonstationary Time Series. Physical review letters, 100(8), 084102. https://doi.org/10.1103/PhysRevLett.100.084102

Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. Sensors, 20(2), 479. https://doi.org/10.3390/s20020479

Empatica Inc. (2024). Professional Users Instructions for Use. Empatica Inc.

Ramdhani, N., Akpewila, F., Faizah, M., & Resibisma, B. (2019, July). It's so real: Psychophysiological reaction towards virtual reality exposure. In 2019 5th International Conference on Science and Technology (ICST), 1, 1–5. https://doi.org/10.1109/ICST47872.2019.9166212

Rimm-Kaufman, S. E., & Kagan, J. (1996). The psychological significance of changes in skin temperature. Motivation and Emotion, 20(1), 63–78.

https://doi.org/10.1007/BF02251007

Rohrbaugh, J. (2016). Ambulatory and Non-Contact Recording Methods. In J. Cacioppo, L. Tassinary, & G. Berntson (Eds.), Handbook of Psychophysiology. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781107415782.014

Ronca, V., Martinez-Levy, A. C., Vozzi, A., Giorgi, A., Aricò, P., Capotorto, R., ... & Di Flumeri, G. (2023). Wearable technologies for electrodermal and cardiac activity measurements: a comparison between fitbit sense, empatica E4 and shimmer GSR3+. Sensors, 23(13), 5847. https://doi.org/10.3390/s23135847

Rubinstein, E. H., & Sessler, D. I. (1990). Skin-surface temperature gradients correlate with fingertip blood flow in humans. Anesthesiology, 73(3), 541–545. https://doi.org/10.1097/00000542-199009000-00027

Sanches, P., Höök, K., Vaara, E., Weymann, C., Bylund, M., Ferreira, P., ... & Sjölinder, M. (2010, August). Mind the body! Designing a mobile stress management application encouraging personal reflection. In Proceedings of the 8th ACM conference on designing interactive systems, 47–56. https://doi.org/10.1145/1858171.1858182

Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., ... & Seneviratne, A. (2017). A survey of wearable devices and challenges. IEEE Communications Surveys & Tutorials, 19(4), 2573–2620. https://doi:10.1109/COMST.2017.2731979.

Sehrt, J., Yilmaz, U., Kosch, T., & Schwind, V. (2024, May). Closing the Loop: The Effects of Biofeedback Awareness on Physiological Stress Response Using Electrodermal Activity in Virtual Reality. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 1–7. https://doi.org/10.1145/3613905.3650830

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. https://doi.org/10.2307/2333709

Shen, C. (2015). Analysis of detrended time-lagged cross-correlation between two nonstationary time series. Physics Letters A, 379(7), 680–687. https://doi.org/10.1016/j.physleta.2014.12.036

Sinichi, M., Gevonden, M., & Krabbendam, L. (2024, July 1). Quality in Question: Assessing the accuracy of four heart rate wearables and the implications for psychophysiological research. https://doi.org/10.31219/osf.io/wkzsn

Ter Harmsel, J. F., Noordzij, M. L., Goudriaan, A. E., Dekker, J. J. M., Swinkels, L. T. A., van der Pol, T. M., & Popma, A. (2021). Biocueing and ambulatory biofeedback to enhance emotion regulation: A review of studies investigating non-psychiatric and psychiatric populations. International journal of psychophysiology : official journal of the International Organization of Psychophysiology, 159, 94–106. https://doi.org/10.1016/j.ijpsycho.2020.11.009

Toast Interactive. (2016). Richie’s Plank Experience (Version 3.43) [Computer software]. Toast Interactive. https://toastinteractive.com/richiesplankexperience/

Toast Interactive. (n.d.). Richie’s Plank. Retrieved July 1, 2024, from https://toastinteractive.com/richiesplankexperience/

Turpin, G., Shine, P., & Lader, M. (1983). Ambulatory electrodermal monitoring: effects of ambient temperature, general activity, electrolyte media, and length of recording. Psychophysiology, 20(2). https://doi.org/10.1111/j.1469-8986.1983.tb03291.x

Varjo HQ. (n.d.). VR3 [Apparatus and software]. https://varjo.com/products/varjo-vr-3/

van Dammen, L., Finseth, T. T., McCurdy, B. H., Barnett, N. P., Conrady, R. A., Leach, A. G., ... & Shirtcliff, E. A. (2022). Evoking stress reactivity in virtual reality: A systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews, 138, 104709. https://doi.org/10.1016/j.neubiorev.2022.104709

van Der Mee, D. J., Gevonden, M. J., Westerink, J. H., & De Geus, E. J. C. (2021). Validity of electrodermal activity-based measures of sympathetic nervous system activity from a wrist-worn device. International Journal of Psychophysiology, 168, 52–64. https://doi.org/10.1016/j.ijpsycho.2021.08.003

van Dooren, M., & Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. Physiology & behavior, 106(2), 298-304. https://doi.org/10.1016/j.physbeh.2012.01.020

van Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M., ... & Noordzij, M. L. (2020). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. Behavior research methods, 52, 607–629. https://doi.org/10.3758/s13428-019-01263-9

Venables, P. H. & Christie, M. J. (1973). Mechanisms, instrumentation, recording techniques, and quantification of responses. In W. F. Prokasy & D. C. Raskin (eds.), Electrodermal Activity in Psychological Research. New York: Academic Press.

Vinkers, C. H., Penning, R., Hellhammer, J., Verster, J. C., Klaessens, J. H., Olivier, B., & Kalkman, C. J. (2013). The effect of stress on core and peripheral body temperature in humans. Stress, 16(5), 520–530. https://doi.org/10.3109/10253890.2013.807243

Yuan, N., Fu, Z., Zhang, H., Piao, L., Xoplaki, E., & Luterbacher, J. (2015). Detrended partial-cross-correlation analysis: a new method for analyzing correlations in complex system. Scientific reports, 5(1), 8143. https://doi.org/10.1038/srep08143

Zhu, L., Ng, P. C., Yu, Y., Wang, Y., Spachos, P., Hatzinakos, D., & Plataniotis, K. N. (2022, May). Feasibility study of stress detection with machine learning through eda from wearable devices. In ICC 2022-IEEE International Conference on Communications. p. 4800–4805. IEEE. https://doi.org/10.1109/ICC45855.2022.9838970

Appendices

1. Supportive materials

1.1 Information Sheet

Project Title: Validaton Study of the Empatica EmbracePlus Wristband for Measuring Electrodermal Activity Utilizing a Virtual Reality Protocol
Ethics Code: 240530

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Purpose of the research
Wearable devices are increasingly used to measure our physiology, but it remains unclear how well they do that. This study uses a Virtual Reality simulation to assess the validity of signals from a wearable device compared with reference devices. The signals reflect the change of sweat level on the skin in response to a VR stimuli.

Procedure
During the main task, participants will wear three wearable devices: a medical wristband and two research-grade devices. Additionally, participants will engage in a VR simulation resembling real-life scenarios. A set of short surveys will be completed before and after the main task. The data collected from devices will be used to compare signal differences and will be analysed anonymously.

Devices involved
BioNomadix sensor, VU-AMS sensor, and EmbracePlus wristband; Oculus VR system.

Benefits of participating
This research project has been reviewed and approved by the BMS Ethics Committee. On the benefit side, participants will experience the innovative experiment procedure in VR environment. As participants, you will have the pleasure of knowing that you have made a contribution to our understanding of the relationship between signals from two types of devices. If you wish you can take a picture or a screenshot of your signals, which will also be explained to you after the study is completed. There will be no other direct benefits from being involved in the study.

Risks of participating
On the risk side, there would be tripping risk for the participants. As participants enter the task wearing a VR helmet and there is a plank on the floor, a researcher will accompany and navigate you to ensure your safety when it is in need.
The devices utilized in this study adhere to the standard methodology commonly employed by most wearables available in the market. This methodology typically involves applying an imperceptible microcurrent to the fingers or wrist. The safety of this approach is well-established, supported by

numerous research studies and successful implementations over recent decades.
The stressor or stimuli utilized in this experiment is a normal acute stressor, resulting in an acute response that typically resolves within several minutes. While for most participants this experience is not harmful, it's important to acknowledge that a small percentage of individuals may find the task overly stressful. Therefore, we want to emphasize to each participant that you are free to withdraw from the experiment at any time. This information will be reiterated in the instructions given to participants immediately before the onset of the stressor.

Procedures for withdrawal from the study
You can inform the experimenter about the withdrawal at any time before, during or after the participation by email or oral words. Please note, the possibility to withdraw data is no longer possible if published or data have been anonymized.

What happens at the end of the study?
You will get debriefed on the details of the purpose of the study and your physiological reaction will be explained to you. You are free to take a picture of the signal collected and ask any questions you might have about the procedure you underwent or the study overall. Your data is then stored securely and anonymously for subsequent processing.

Personal data
All the personal data will be anonymized. You will be assigned an appointment number and your demographic information as well as all the data gathered during the experiment will be stored and analysed under that number. Informed consent forms containing personal information (signatures) will be stored securely within the UT infrastructure.

Confidentiality and anonymisation
No identifiable information will be collected for analysis purposes and all other personal information (e.g., medical history) will be anonymized and treated as confidential. An individual's responses would not be used in any way that would allow for their identification.

Retention period for the research data
Ten years

Contact details .
Qu, Y. (y.qu-3@student.utwente.nl)
If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-hss@utwente.nl

1.2 Consent Form

Consent Form for [*Validaton Study of the Empatica EmbracePlus Wristband for Measuring Electrodermal Activity Utilizing a Virtual Reality Protocol*]
YOU WILL BE GIVEN A COPY OF THIS INFORMED CONSENT FORM

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated [DD/MM/YYYY], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.	☐	☐
Risks associated with participating in the study		
I understand that taking part in the study involves wearing three wearable devices, completing a VR task, and responding to a set of short surveys.	☐	☐
I understand that taking part in the study involves the following risks: tripping risk due to wearing a VR helmet when moving around in the lab; discomfort of a short and acute stressful experience caused by the stimuli generated by VR; motion sickness from wearing a VR helmet.	☐	☐
Use of the information in the study		
I understand that information I provide will be anonymously processed for the research purpose of a master's thesis as well as for subsequent analysis for a future research project.	☐	☐
I understand that personal information collected about me that can identify me, such as [e.g. my name], will not be shared beyond the study team.	☐	☐
Future use and reuse of the information by others		
I give permission for the data that I provide to be archived in Areda, the UT Archive for Research Data, so it can be used for future research and learning, excluding commercial use.	☐	☐

Signatures

____

____

Name of participant [printed] Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

__________________________ ___________________ _________

Researcher name [printed] Signature Date

Study contact details for further information:

Qu, Y.

y.qu-3@student.utwente.nl

Contact Information for Questions about Your Rights as a Research Participant

If you have questions about your rights as a research participant, or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher(s), please contact the Secretary of the Ethics Committee/domain Humanities & Social Sciences of the Faculty of Behavioural, Management and Social Sciences at the University of Twente by ethicscommittee-hss@utwente.nl

1.3 Questionnaire

Q1

Q1 Please fill in the numeric ID code

Q2

Q2 Please fill in your age

Q3

Q3 What is your ethnic group?

- Black / African / Caribbean background
- Asian background
- White background
- Multiple ethnic background
- Other ethnic group, please describe

Q4

Q4 Please select the sex assigned at birth

- Male
- Female
- Intersex

Q5

Q5 Do you menstruate?

- ○ Yes
- ○ No

Q6

Q6 Are you in the two weeks following the end of last menstrual cycle?

- ○ Yes
- ○ No

Q7

Q7 Have you been diagnosed with any psychiatric disorders?

- ○ No
- ○ Yes

Q8

Q8 If yes, please indicate the specific diagnosis; If you feel not comfortable to do so, please fill in "NA".

Q9

Q9 Have you been diagnosed with any following diseases? If no, please skip this question.

- ☐ Heart diseases
- ☐ Epilepsy
- ☐ High blood pressure
- ☐ Obesity (BMI>=30)

Q10

Q10 Have you been prescribed any medications within the past two weeks (psychiatric medications, belta-blockers, hormonal contraceptives, etl)?

- ○ No
- ○ Yes

Q11

Q11 If yes, please name the medication; If you feel not comfortable to do so, please fill in "NA".

Q12

Q12 Have you used VR before?

- No
- Yes

Q13

Q13 Did you do any of the following: drinking coffee or alcohol, smoking, exercising 3 hours prior to attending the study?

- No
- Yes

Q14

Q14 Did you have at least 6-hour sleep last night?

- No
- Yes

Please inform the researcher that you have reached this page.

Q15

Q15 On a scale of 0-100 how stressed do you feel right now?

	0	10	20	30	40	50	60	70	80	90	100
Stress level											

Page Break

Q16

Q16 On a scale of 0-100 how stressed did you feel during the task?

	0	10	20	30	40	50	60	70	80	90	100
Stress level											

Page Break

Q17

Q17 On a scale of 0-100 how stressed do you feel right now?

	0	10	20	30	40	50	60	70	80	90	100
Stress level											

Page Break

Q18

Q18 In the last month, how often have you been upset because of something that happened unexpectedly?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q19

Q19 In the last month, how often have you felt that you were unable to control the important things in your life?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q20

Q20 In the last month, how often have you felt nervous and "stressed"?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q21

Q21 In the last month, how often have you felt confident about your ability to handle your personal problems?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q22

Q22 In the last month, how often have you felt that things were going your way?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q23

Q23 In the last month, how often have you found that you could not cope with all the things that you had to do?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q24

Q24 In the last month, how often have you been able to control irritations in your life?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q25

Q25 In the last month, how often have you felt that you were on top of things?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q26

Q26 In the last month, how often have you been angered because of things that were outside of your control?

- 0 = Never
- 1 = Almost Never
- 2 = Sometimes
- 3 = Fairly Often
- 4 = Very Often

Q27

Q27 In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?

- ◯ 0 = Never
- ◯ 1 = Almost Never
- ◯ 2 = Sometimes
- ◯ 3 = Fairly Often
- ◯ 4 = Very Often

Q28

Q28 Have you played Richie's Plank Experience prior to attending this experiment? If you have experienced other VR application similar with Richie's Plank Experience, please indicate the name of it:

- ☐ No
- ☐ Yes
- ☐ Other:

Q29

Q29 On a scale of 0-100, to what extent are you afraid of heights?

	0	10	20	30	40	50	60	70	80	90	100
fear of heights											

+ Add page break

Q30

Q30 Have you been diagnosed with acrophobia (fear of heights) or basophobia (fear of falling)?

- ◯ Yes
- ◯ No
- ◯ If yes, please indicate the diagnosis:

1.4 Pilot report

Participant Overview

<u>1. Participant A</u>

Myopia: 500 degrees

Condition: Did not wear glasses under the VR helmet due to incompatibility.

Experience: Reported unclear vision in the game, yet still felt a significant sense of fear.

<u>2. Participant B</u>

Myopia: 200 degrees

Condition: Participated in the VR experience without corrective lenses.

Experience: Reported that the vision was adequate despite not wearing glasses.

<u>3. Participant C</u>

Condition: No specific visual impairments or issues reported.

Experience: Successfully completed the experiment without any complications.

<u>4. Participant D (Withdrawal)</u>

Myopia: 500 degrees

Condition: Wore small-sized glasses during the experiment.

Experience: Having experienced the game previously, reported feelings of boredom and intentionally jumped off the plank.

<u>5. Participant E (Withdrawal)</u>

Condition: No visual impairments reported.

Experience: Initially showed interest in the challenge, but upon exposure to the simulated high altitude, withdrew immediately after refusing to step out of the elevator.

In summary, Participants E experienced fear or discomfort, leading to withdrawal from the experiment. Participant D's prior experience with the game and perceived boredom led to early withdrawal. The attrition rate is 40% (2/5).

2. Experiment materials

2.1 Instruction

"Welcome to my study on the validation of a wristband. During this experiment, your sweat levels on the skin of your hands and wrists in response to a VR task will be recorded using three measurement devices, as displayed on the table. During the experiment, you will wear an VR helmet on your head, two wearable devices on your non-dominant hand, and two portable devices on your dominant hand and waist. Then, you will experience a naturalistic scenario in a VR game. Now, I will apply the devices. Before we begin, please remove any jewelry from your hands or wrists to ensure data quality.

Could you tell me which hand is your dominant hand? (If the answer is the right hand:) I will apply this sensor (EmbracePlus) on your left wrist. Could you help me tighten the wristband? Is it secure enough? Are you comfortable with the tightness? Okay, now you will first engage in a warm up task to ensure signal quality. Please come with me to the stairs. Please walk down and up one floor three times. I will wait here and count for you. Well done! Let's go back to the lab.

Please take a seat here. First, please read through this information sheet. If you decide to participate in this experiment, please sign both copies of the informed consent form. One copy is for you. I will now apply the second sensor (BioNomadix) to your left wrist and fingers. First, I will attach two electrodes to your index and middle fingers (if finger placement). This isotonic and harmless gel is used to fill the gap between the skin and sensors to guarantee the quality of the signals. Please show me your left palm with your fingers spread apart. Please turn your palm over; I will apply this belt around your wrist and connect the electrodes. Is the tightness comfortable? Okay, I will now apply the third sensor (VU-AMS). Please show me your right palm. Thanks! I'm going to use the paper tape to stabilize all the sensors. I will apply this strap around your waist. Please help me press this device on your left hip. Is the tightness comfortable? Okay, now let's check the signal and make sure it works. Please take a deep breath, hold 2 seconds, and out. Thanks. Again, please take a deep breath, hold 2 seconds, and out. Thanks. Before the experiment, please fill out this short questionnaire. When you reach the last question of this page, please inform me, and do not click on the "next" button for now. Thanks!

We are ready to start the experiment. Please stand here. During the experiment, please try to talk less and move less. First, please do three squats and three jumps with your arms slightly away from your body, like this. Now, please stand upright with your arms beside you, without moving or talking, for three minutes. Slight movements and observing the environment are acceptable, just please try to avoid hand movements and pressing electrodes. I will make a timer and let you know when this condition ends. Are you comfortable? Okay, let's start. The time is up. Please rate your stress level at this moment. I'm going to help you wear the VR device. Is the tightness comfortable? Is the sound volume alright? Here is a right-hand controller. Are you on the street now? Please stretch your right hand to the "I Agree" button and press it. Okay, now please turn right. Can you see the tree across the street? Please face the tree and stand for three minutes. I will let you know when this condition ends. Turning upper body to observe the environment is okay, just please try to avoid strong movements, especially hand movements. Are you comfortable? Okay, let's start. The time is up. Thank you. Please turn around. Can you see the elevator? Please walk into the elevator. (if real plank: I'm going to navigate you slowly to the elevator; please watch out to avoid the plank in reality.) Now please take a look at the buttons on the right and press the highest 'Plank' button. I'm removing the controller for you. You have arrived at the plank floor. Please remember it is okay to report discomfort or withdraw at any time. If it's possible, please walk out to the plank and slowly approach the edge (but not too close to the edge). Now please stand for three minutes. Now, please slowly walk back to the elevator and press the lowest button to the ground floor. I'm removing the controller for you. Please walk back to the street. (I'm going to navigate you again to avoid the plank on the floor.) Now please stand in the street for another three minutes. I'm going to help you remove the VR helmet. Now please stand for the final three minutes. Now, please do three squats and jumps again. Please rate your stress level during the plank session and at this moment. Congratulations! You have finished the experiment. I'm now going to help you remove all the remaining devices. Please help by filling out the last few questions.

Thank you for your participation. This experiment was designed to test the validity of EDA signals from the wristband compared to reference devices. Debriefing is as following:

1. Before the experiment, you were asked to walk up and down three times because the wristband's function depends on sweat on the skin. This task ensures that every participant has some sweat on their wrist when attending the main sessions.
2. At the beginning and end of the experiment, you were asked to do three squats and jumps. This task is for data synchronization using movement trackers.
3. EDA signals are easily affected by breathing and movement, which is why you were asked to move and talk less during the experiment.
4. The stimuli, which induce stress, were not specified at the beginning of the experiment because disclosing them could reduce their effect and create an expectation effect.
5. The questionnaire includes three parts. Ten questions from the Perceived Stress Scale measure your perception of stress during the last month. Three stress rating questions are for subjective appraisal of stress during the experiment. Both add to the interpretation of the objective stress indicators, the physiological signals obtained with the wearable devices. The rest of the questionnaire collects demographic and medical information to understand the sample composition.
6. The monitoring target in this study is electrodermal activity (EDA), which measures sweat gland activity using electrical methods. For simplicity, I have used the term "sweat level" to refer to the targeted biomarker.

You can inform me about the registration of the study report, a change of decision, or other questions through the email address provided in the consent form. Feel free to contact me. Have a good day."

2.2 Allocation of four settings

Participant	R_Plank	D_Hand	Placement	Temperature
1	Yes	Right	Eminence	No
3	No	Left	Eminence	No
4	No	Right	Eminence	No
5	No	Right	Eminence	No
6	Yes	Right	Eminence	Yes
7	No	Left	Eminence	Yes
9	No	Right	Eminence	Yes
12	Yes	Right	Finger	Yes
13	Yes	Right	Finger	Yes
15	No	Right	Finger	Yes
16	Yes	Right	Finger	Yes
17	No	Right	Finger	Yes
18	Yes	Right	Finger	Yes
19	Yes	Left	Finger	Yes
21	Yes	Right	Finger	Yes
22	Yes	Right	Finger	Yes
24	Yes	Left	Finger	Yes
25	No	Left	Eminence	No
26	No	Right	Eminence	No
27	No	Right	Eminence	Yes
28	No	Left	Finger	Yes

[a]R_Plank: setting with or without a real plank (Yes vs. No). [b]D_Hand: dominant hand (Right vs. Left). [c]Placement: electrode placement of reference devices (Finger vs. Eminence). [d]Temperature: recording of the ambient temperature (Yes vs. No).

2.3 Protocol

Phases	Estimated Duration	Desicription	Check list			
			N	Steps	Items	Check
Preparation (researcher)	10 minutes	Check time zone and time server on computer and cell phone. Refresh the time synchronization by clicking Date & time settings > Additional settings > Sync now.	1	Time	time zone	
					time server	
		Check the connection between thermometer and the cell phone. Put thermometer on the desk where is close to the playing area of the VR game.	2	Temperature	connection	
					position	
		Take out Biopac main module, EDA transmitter and receiver, accelerometer, and electrode, and electrolyte paste out of the storage boxes. Connect the power and Ethernet wires for the main module. Connect accelerometer with the main module.	3	Biopac	power supply	
					accelerometer	
		Turn on power. Launch Acqknowledge. Set configuration. Calibrate the accelerometer.	4	Acqknowledge	configuration	
		Sign in the Care Lab app on cell phone with one of the five ready generated login identities for participants.	5	Care app	login	
		Check left battery and connection with the cell phone app.	6	EmbracePlus	battery	
					connection	
		Assemble batteries and the flash card for VU-AMS with noting the power on sound effect.	7	VU-AMS	battery	
					power-on sound	
		Set the VR device and game's environment factors including elevator direction, floor level, and playing area. Set the position of real plank for half of the participants. Turn off the computer sound.	8	VR	elevator direction	
					floor calibration	
					playing area	
					(set real plank)	
					sound off	
Before the experiment	30 minutes	Welcome the participants, introduce the project, and provide the informed consent forms.	1	Opening	welcoming	
					introduction	
					consent form	
		After enrolled, the participants wears EmbracePlus and completes 'Stairs' task, then wears the rest devices assisted by the researcher.	2	Set up	EmbracePlus wearing	
					stair task	

					BioNomadix wearing	
					VU-AMS wearing	
		The researcher confirms the signal conditions in Acqknowledge, Care Lab, and VU-DAMS. Deep breathing test used to check signal collection is about asking the participant to take a deep breath, hold two seconds, and exhale when observing the real time signals in Acqknowledge and VU-DAMS respectively. For VU-DAMS, the researcher checks configurations and set time to the computer time before starting the recording. Then the researcher checks both Care Lab app and the interface of the EmbracePlus to confirm the connection state.	3	Check signals	Acqknoledge	
					Care Lab app	
					VU-DAMS	
		Then, a randomization app is provided for participants to generate a 6 digit identity number to fill in the survey, or participants from SONA subject pool could use their appointment number. Next, the researcher guides them to fill in a short questionnaire for demographics collection. Then participants answer the 14 items on the first page.	4	Survey	Q1 - Q14	
Experiment	20 minutes	The researcher introduces the main sessions and guides the participant to move and talk less during the experiment. Upon the researcher makes the first event marker 'start' with the instruction 'Experiment starts', first, the participant is required to squat and jump three times following the researcher's demonstration.	1	Dynamic1	event_marker_start	
					squats and hops	
		Then, the participant is engaged in the first standing session.The experimenter makes the event marker 'baseline1'and a 3-min timer, and then observes and makes notes when necessary during the session and informs the end of session when the time is up.	2	Baseline_1	event_marker_baseline1	
					timer	
		After the first session, a rating question is offered for them to rate the perceived stress level at that moment.Then, the researcher helps them wear the HMD device and confirms the sound volume.	3	Dynamic2	survey (Q15)	
					HMD wearing	
					sound volume	
		Then, the participants were required to stand on the street in VR for another 3 minutes. The experimenter makes the event marker 'baseline2'and another 3-min timer, and then observes and makes notes when necessary during the session and informs the end of session when the time is up.	4	Baseline_2	event_marker_baseline2	
					timer	
		After that, the researcher passes the controller to the participant again, and guides them to walk into the elevator and navigate them to avoid the plank on the floor. Then the participant is guided to press the floor button and go up to the plank floor. The experimenter makes event marker 'going up' when hearing the door-close sound effect, then remove the controller.	5	Dynamic3	passing controller	
					navigation assistance	
					event_marker_going_up	

					door-close sound	
		The researcher makes event marker 'door open' when hearing the sound effect of door's opening, and guides participants to walk out the elevator and approach to the edge of the plank.	6	Dynamic4	event_marker_door_open	
					door-open sound	
		Participants then stand there for another 3 minutes. The researcher makes event marker 'plank' and another 3-min timer, and then observes and makes notes when necessary during the session and informs the end of session when the time is up.	7	Plank	event_marker_plank	
					timer	
		Next, the participants are guided to walk back into the lift. The researcher passes the controller to them.	8	Dynamic5	navigation assistance	
					passing controller	
		Then they press the floor button and go down. The researcher makes event marker 'going down' when hearing the door-close sound effect and then takes the controller away. The researcher makes event marker 'door open' when hearing the sound effect of door's opening, and guides participants to walk out the elevator and back to the street.	9	Dynamic6	event_marker_going_dow n	
					door-close sound	
					event_marker_door_open	
					door open sound	
		Participants stand on the street for another 3 minutes. The researcher makes event marker 'recovery1' and another 3-min timer, and then observes and makes notes when necessary during the session and informs the end of session when the time is up.	10	Recovery_1	event_marker_recovery1	
					timer	
		Then, the researcher helps the participant to remove the HMD device.	11	Dynamic7	HDM removal	
		Participants stand in the lab for another 3 minutes. The researcher makes event marker 'recovery2' and another 3-min timer, and then observes and makes notes when necessary during the session and informs the end of session when the time is up.	12	Recovery_2	event_marker_recovery2	
					timer	
		After that, another two rating question are offered for participants to rate the perceived stress level during the plank session and at that moment. At last, the participant is required to do three squats and jumps again. Then the experiment ends with the researcher making the last marker 'end'. In total, the researcher makes 11 event markers as shown in the check list.	13	Dynamic8	survey (Q17&18)	
					squats and hops	
					event_marker_end	
After the experiment	10 minutes	After the main sessions, the participant answers the rest of the questionnaire. The researcher debriefs and help participants remove devices. If desired, the signal will be explained. Participants who are interested in the study report register for receiving it after its	1	Survey	Q19-Q30	
			2	Closure	debriefing	
					removal of devices	

		completion.			(graph explanation)	
					result report	
Close up (researcher)	60 minutes	Log out the EmbracePlus account in Care Lab app, export data from all devices, using software including Acqknowledge, VU-DAMS,and Cyberduck.	1	Export data	Acqknowledge	
					VU-DAMS	
					Care Lab logout	
					Cyberduck	
		Save raw data and a back-up compressed file in university secured cloud drive. Keep informed consent forms appropriately.	2	Save data	secured cloud drive	
					consent forms	
		Write logs for data documentation, recording details like date, time, temperature, and so on.	3	Log	experiment details	
		Put all devices back to places, delete data from lab devices, shut down the computer, and tidy and close the lab.	4	Cleaning up	delete data	
					devices storage	
					close the lab	

2.4 Log

Experiment Log 1 (reported by Yin Qu) Participant1

Login ID index	1	Temperature	21.8℃ 11:30
Appointment	11:30-12:30	Humidity	56% 11:30
Placement	Eminence	Real plank	Yes
Dominant hand: Right			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: 11:11:45
Data documentation:
Compressed file: -.zip (size: 34.2 MB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part01.acq (25233KB); Part01_ (5FS file 17412KB); EmbracePlus_folder (9.80MB).

Experiment Log 3 (reported by Yin Qu) Participant3

Login ID index	1	Temperature	22.5℃ 10:05
Appointment	10:05-11:05	Humidity	47% 10:05
Placement	Eminence	Real plank	No
Dominant hand: Left hand writer (right controller), applied on left hand.			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: 09:53:39
Data documentation:
Compressed file: -.zip (size: 63.1 MB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part01 (acq file 34089KB); Part01_ (5FS file 13073KB); Part01_ (edf file 287KB); EmbracePlus_folder (17.3MB).

Experiment Log 4 (reported by Yin Qu) Participant4

Login ID index	2	Temperature	23.3℃ 11:30
Appointment	11:30-12:30	Humidity	47% 11:30
Placement	Eminence	Real plank	No
Dominant hand: Right			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: 09:53:39
Data documentation:
Compressed file: -.zip (size: 63.1 MB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part02 (acq file 34155KB); Part02_ (5FS file 14481KB); Part02_ (edf file 318KB); EmbracePlus_folder (17.3MB).

Experiment Log 5 (reported by Yin Qu) Participant5

Login ID index	3	Temperature	24.2℃ 16:00
Appointment	16:00-17:10	Humidity	50% 16:00
Placement	Eminence	Real plank	No
Dominant hand: Right			

Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 09:53:39 **Data documentation:** Compressed file: -.zip (size: 63.1 MB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part03 (acq file 33850KB); Part03_ (5FS file 13961KB); Part03_ (edf file 307KB); EmbracePlus_ folder (17.3MB).

Experiment Log 6 (reported by Yin Qu) Participant6

Login ID index	1	Temperature	20.6℃ 15:50
Appointment	15:50-16:50	Humidity	57% 15:50
Placement	Eminence	Real plank	Yes
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: - **Data documentation:** Compressed file: -.zip (size: 20.4 MB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 36911KB); Part01_ (5FS file 15883KB); Part01_ (edf file 349KB); EmbracePlus_folder (4.63MB)			

Experiment Log 7 (reported by Yin Qu) Participant7

Login ID index	1	Temperature	20.8℃ 10:20
Appointment	09:50-10:50	Humidity	57% 10:20
Placement	Eminence	Real plank	No
Dominant hand: Left hand (left controller), applied on right hand.			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 10:08:35 **Data documentation:** Compressed file: -.zip (size: 52.5 MB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 28812KB); Part01_ (5FS file 13101KB); EmbracePlus_folder (8.86MB).			

Experiment Log 9 (reported by Yin Qu) Participant9

Login ID index	3	Temperature	20.8℃ 15:00
Appointment	15:00-16:00	Humidity	64% 15:00
Placement	Eminence	Real plank	No
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 10:08:35 **Data documentation:** Compressed file: -.zip (size: 52.5 MB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part03 (acq file 32551KB); Part03_ (5FS file 17816KB); EmbracePlus_folder (8.86MB).			

Experiment Log 12 (reported by Yin Qu) Participant12

Login ID index	3	Temperature	22.6°C 15:00
Appointment	15:00-16:00	Humidity	48% 15:00
Placement	Finger	Real plank	Yes
Dominant hand: Right			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: -
Data documentation:
Compressed file: -.zip (size: 89363KB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part03 (acq file 50543KB); Part03_ (5FS file 19929KB); EmbracePlus_folder (233MB).

Experiment Log 13 (reported by Yin Qu) Participant13

Login ID index	4	Temperature	22.8°C 16:40
Appointment	16:40-17:40	Humidity	46% 16:40
Placement	Finger	Real plank	Yes
Dominant hand: Right			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: -
Data documentation:
Compressed file: -.zip (size: 89363KB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part04 (acq file 29490KB); Part04_ (5FS file 12410KB); EmbracePlus_folder (233MB).

Experiment Log 15 (reported by Yin Qu) Participant15

Login ID index	2	Temperature	22.8°C 13:30
Appointment	13:30-14:30	Humidity	46% 13:30
Placement	Finger	Real plank	No
Dominant hand: Right			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: 11:14:54
Data documentation:
Compressed file: -.zip (size: 29633KB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files:Part02 (acq file 30050KB);Part02_ (5FS file 12349KB); Part02_ (edf file 271KB); EmbracePlus_folder (5.39MB).

Experiment Log 16 (reported by Yin Qu) Participant16

Login ID index	1	Temperature	20.8°C 15:00
Appointment	15:00-16:00 (extended to 16:35)	Humidity	53% 15:00
Placement	Finger	Real plank	Yes
Dominant hand:Right			

Date and time: -iPhone time zone: Amsterdam,Netherlands,- -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 15:26:04 **Data documentation:** Compressed file: -.zip (size: 26676KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 41186KB); Part01_ (5FS file 834KB); Part01_ (5FS file 9661KB); Part01_ (5FS file 32311KB); EmbracePlus_folder (5.7MB).

Experiment Log 17 (reported by Yin Qu) Participant17

Login ID index	1	Temperature	19.7℃ 11:10
Appointment	11:10-12:10 (extended to 13:10)	Humidity	47% 11:10
Placement	Finger	Real plank	No
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 11:01:11 **Data documentation:** Compressed file: -.zip (size: 35062KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 31054KB); Part01_ (5FS file 17543KB); Part01_ (5FS file 12176KB); EmbracePlus_folder (4.42MB).			

Experiment Log 20 (reported by Yin Qu) Participant18

Login ID index	1	Temperature	20.9℃ 10:10
Appointment	10:10-11:10 (extended to 11:30)	Humidity	64% 10:10
Placement	Finger	Real plank	Yes
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 11:01:11 **Data documentation:** Compressed file: -.zip (size: 38695KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 27078KB); Part01_ (5FS file 11345KB); EmbracePlus_folder (8.84MB).			

Experiment Log 21 (reported by Yin Qu) Participant19

Login ID index	2	Temperature	20.6℃ 13:30
Appointment	13:30-14:30 (extended to 14:40)	Humidity	69% 13:30
Placement	Finger	Real plank	Yes
Dominant hand: Left hand writer (right controller), applied on right hand.			

Date and time: -iPhone time zone: Amsterdam,Netherlands,- -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 11:01:11 **Data documentation:** Compressed file: -.zip (size: 38695KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part02 (acq file 41498KB); Part02_(5FS file 16306KB); EmbracePlus_folder (8.84MB).

Experiment Log 24 (reported by Yin Qu) Participant21

Login ID index	1	Temperature	25.7℃ 10:00
Appointment	10:00-11:00	Humidity	53% 10:00
Placement	Finger	Real plank	Yes
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 10:24:26 **Data documentation:** Compressed file: -.zip (size: 59381KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 32696KB); Part01_ (5FS file 13508KB); EmbracePlus_folder (153 MB).			

Experiment Log 25 (reported by Yin Qu) Participant22

Login ID index	3	Temperature	21.1℃ 13:10
Appointment	13:10-14:10	Humidity	59% 13:10
Placement	Finger	Real plank	Yes
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 10:24:26 **Data documentation:** Compressed file: -.zip (size: 59381KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part02 (acq file 36523KB); Part02_ (5FS file 13905KB); EmbracePlus_folder (153 MB).			

Experiment Log 28 (reported by Yin Qu) Participant24

Login ID index	1	Temperature	21.4℃ 12:35
Appointment	12:35-13:35	Humidity	65% 12:35
Placement	Finger	Real plank	Yes
Dominant hand: Left hand writer (right controller), applied on right hand.			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 13:05:02 **Data documentation:** Compressed file: -.zip (size: 19349KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 34244KB); Part01_ (5FS file 13686KB); EmbracePlus_folder (3.35 MB).			

Experiment Log 29 (reported by Yin Qu) Participant25

Login ID index	1	Temperature	Forgot setting up
Appointment	11:10-12:10	Humidity	Forgot setting up
Placement	Eminence	Real plank	No
Dominant hand: Left hand writer (left controller), applied on right hand.			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 09:45:07 **Data documentation:** Compressed file: -.zip (size: 35346KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 30392KB); Part01_ (5FS file 13465KB); EmbracePlus_folder (34.5 MB).			

Experiment Log 30 (reported by Yin Qu) Participant26

Login ID index	2	Temperature	Forgot setting up
Appointment	15:30-16:50	Humidity	Forgot setting up
Placement	Eminence	Real plank	No
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands,- -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 09:45:07 **Data documentation:** Compressed file: -.zip (size: 35346KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part02 (acq file 31985KB); Part02_ (5FS file 13754KB); EmbracePlus_folder (34.5 MB).			

Experiment Log 31 (reported by Yin Qu) Participant27

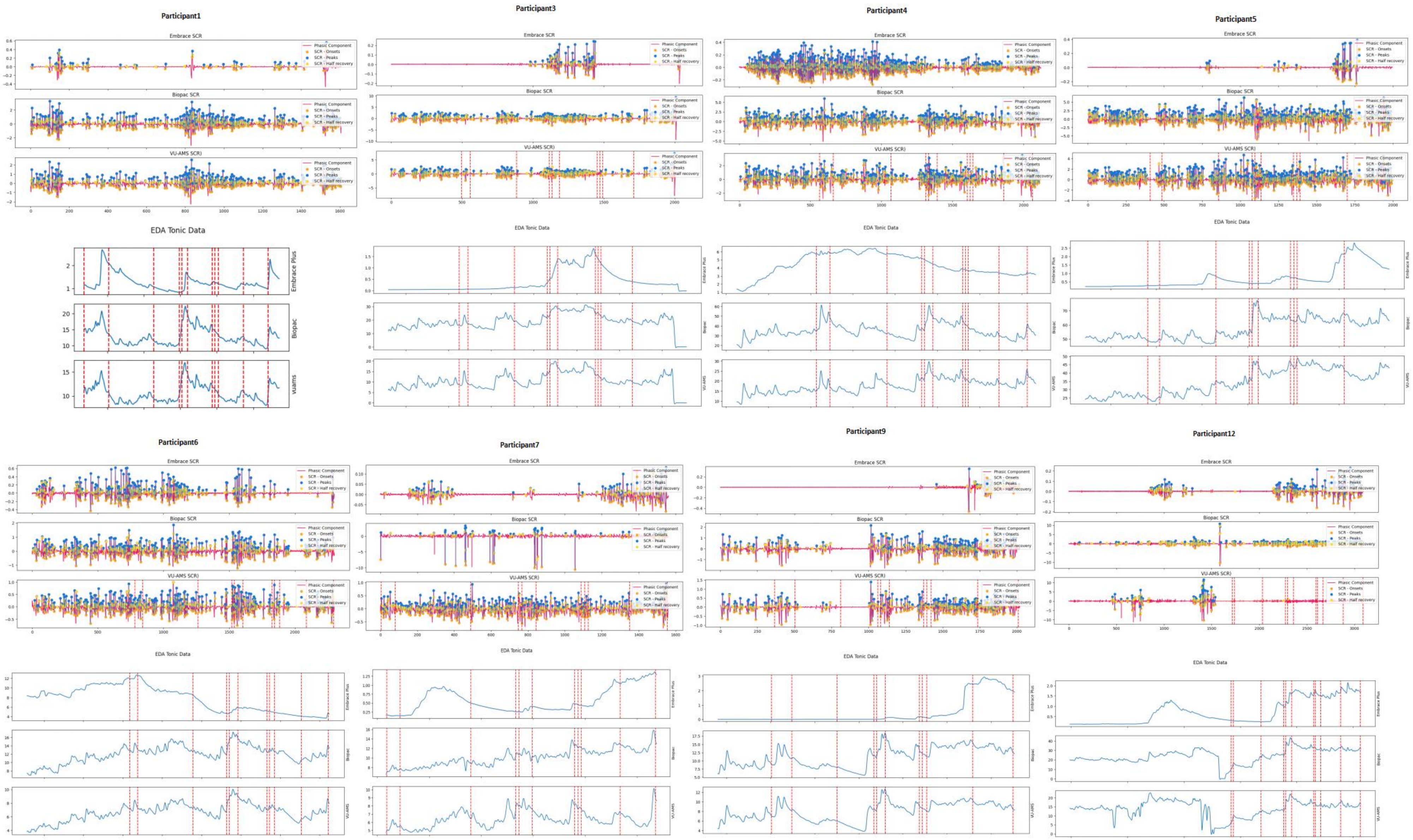
Login ID index	1	Temperature	21°C 12:00
Appointment	12:00-13:00	Humidity	53% 12:00
Placement	Eminence	Real plank	No
Dominant hand: Right			
Date and time: -iPhone time zone: Amsterdam,Netherlands, - -Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna Time server: time.windows.com Synchronized at: 12:08:45 **Data documentation:** Compressed file: -.zip (size: 55573KB) Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study Included files: Part01 (acq file 31837KB); Part01_ (5FS file 13349KB); EmbracePlus_folder (10.6 MB).			

Experiment Log 32 (reported by Yin Qu) Participant28

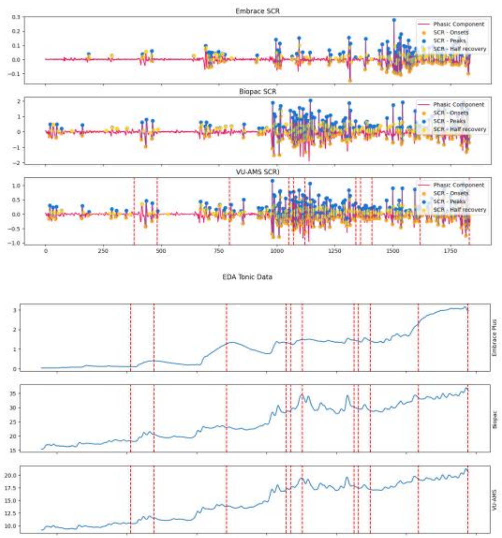
Login ID index	2	Temperature	21.8°C 15:00
Appointment	14:00-15:00	Humidity	47% 15:00
Placement	Finger	Real plank	No
Dominant hand: left-dominant, right-writer. Embraceplus applied on left hand. Because participant said he applies his smartwatch on left hand.			

Date and time:
-iPhone time zone: Amsterdam,Netherlands, -
-Computer time zone: UTC+01:00 Amsterdam,Berlin,Bern,Rome,Stockholm,Vienna
Time server: time.windows.com
Synchronized at: 12:08:45
Data documentation:
Compressed file: .zip (size: 55573KB)
Save in: \\ad.utwente.nl\BMS\bmslab\Projects\Validation_Study
Included files: Part02 (acq file 32496KB); Part02_ (5FS file 12233KB); EmbracePlus_folder (10.6 MB).

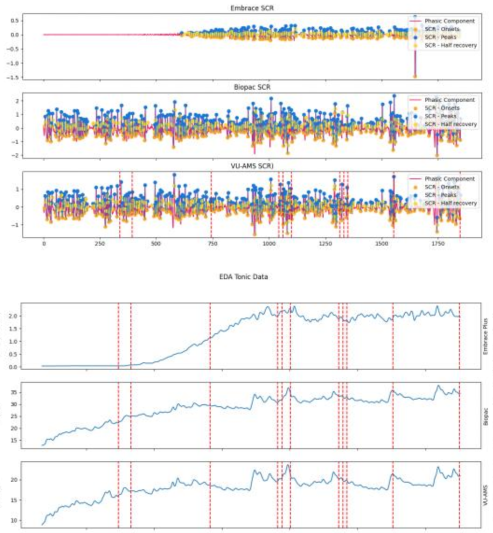
2.5 Visual inspection of tonic and phasic components



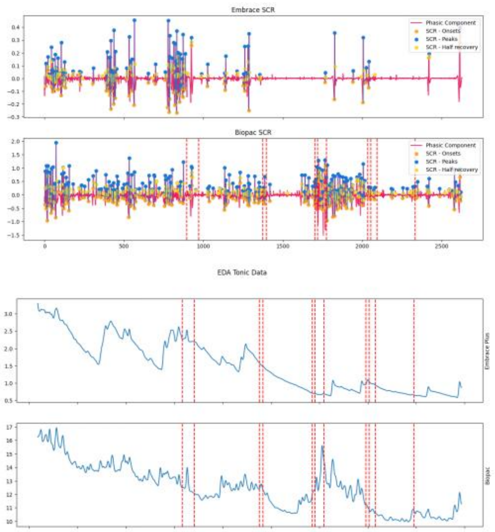
Participant13



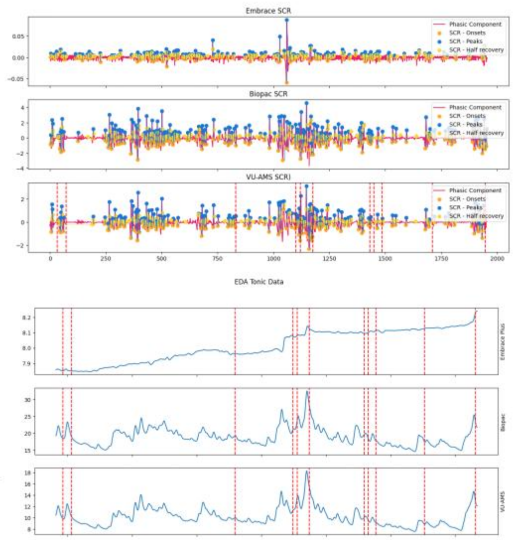
Participant15



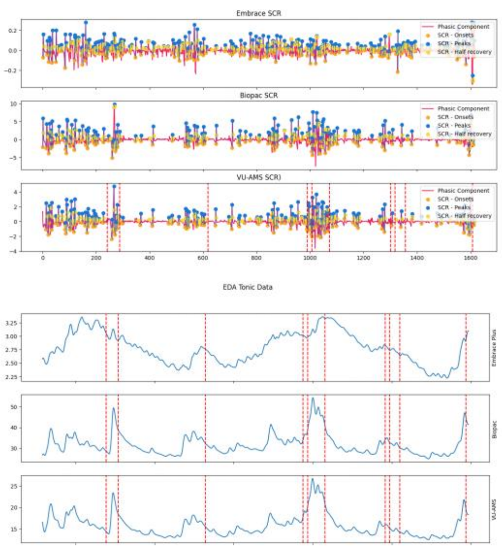
Participant16



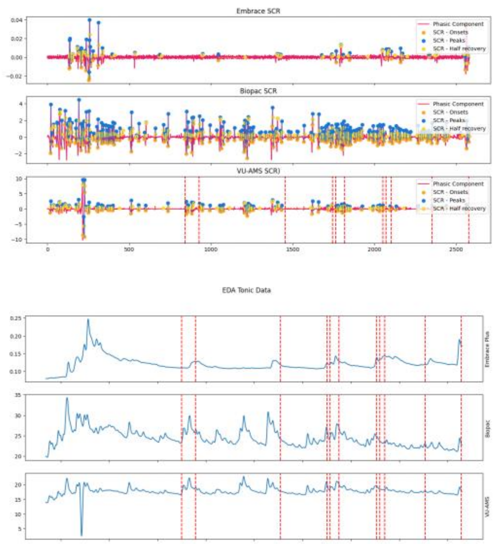
Participant17



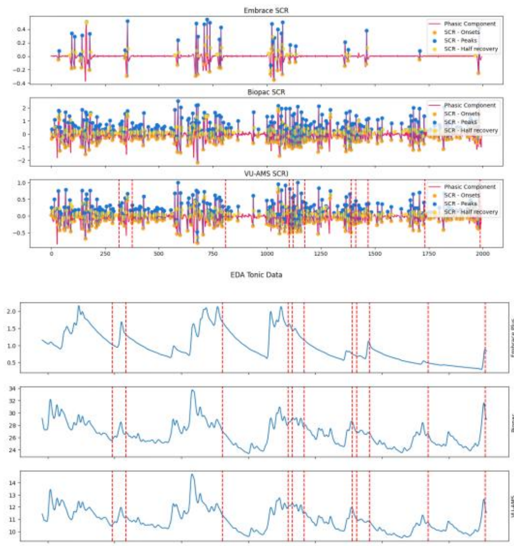
Participant18



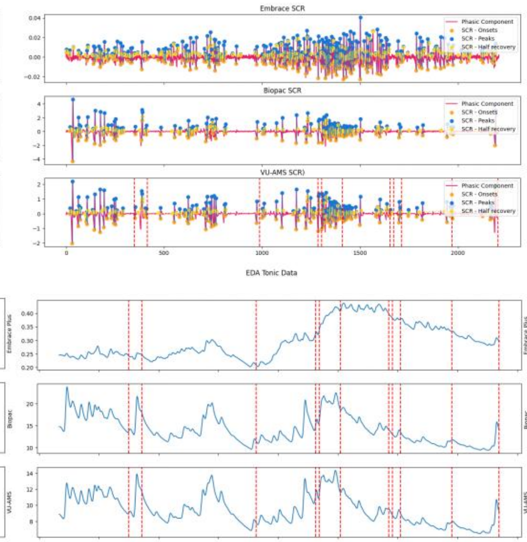
Participant19



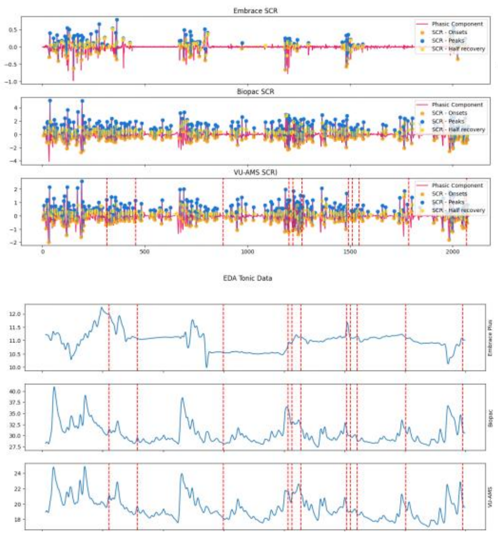
Participant21



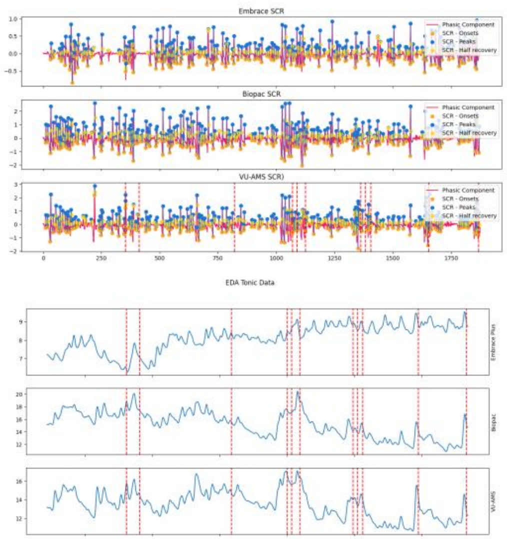
Participant22



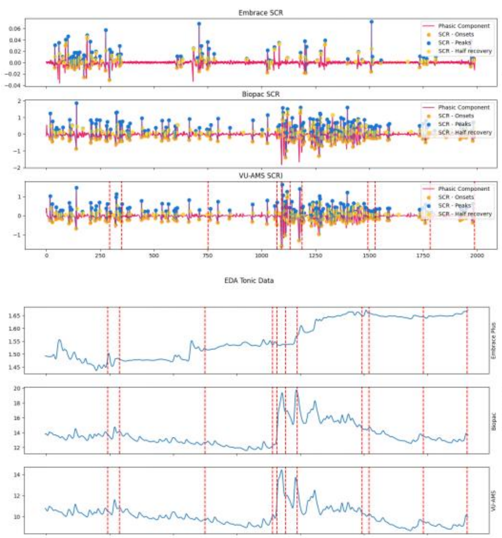
Participant24



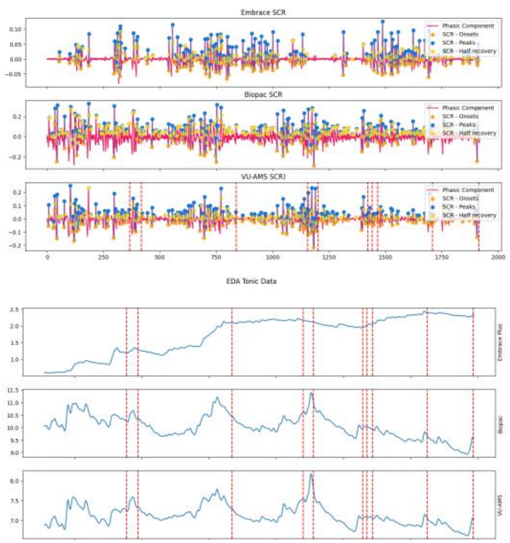
Participant25



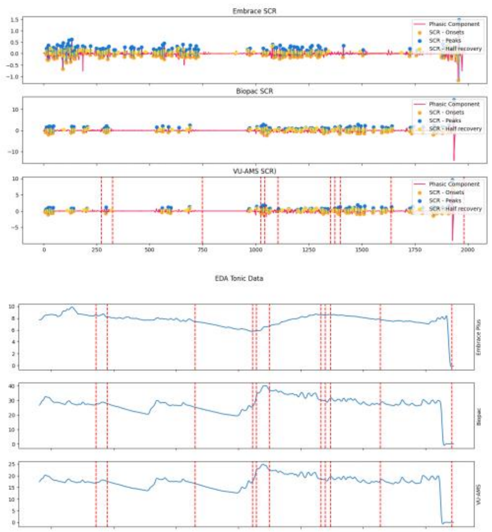
Participant26



Participant27

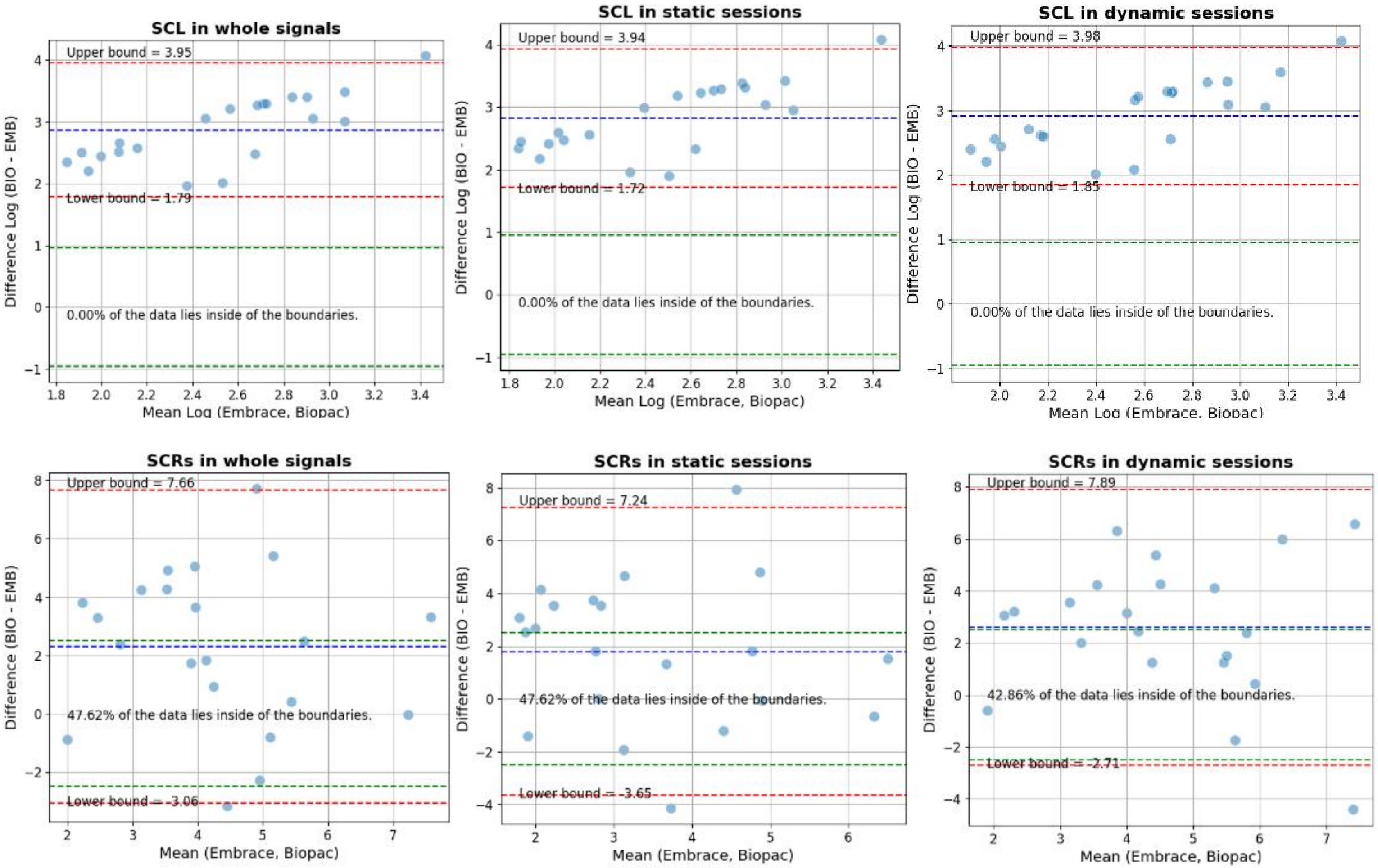


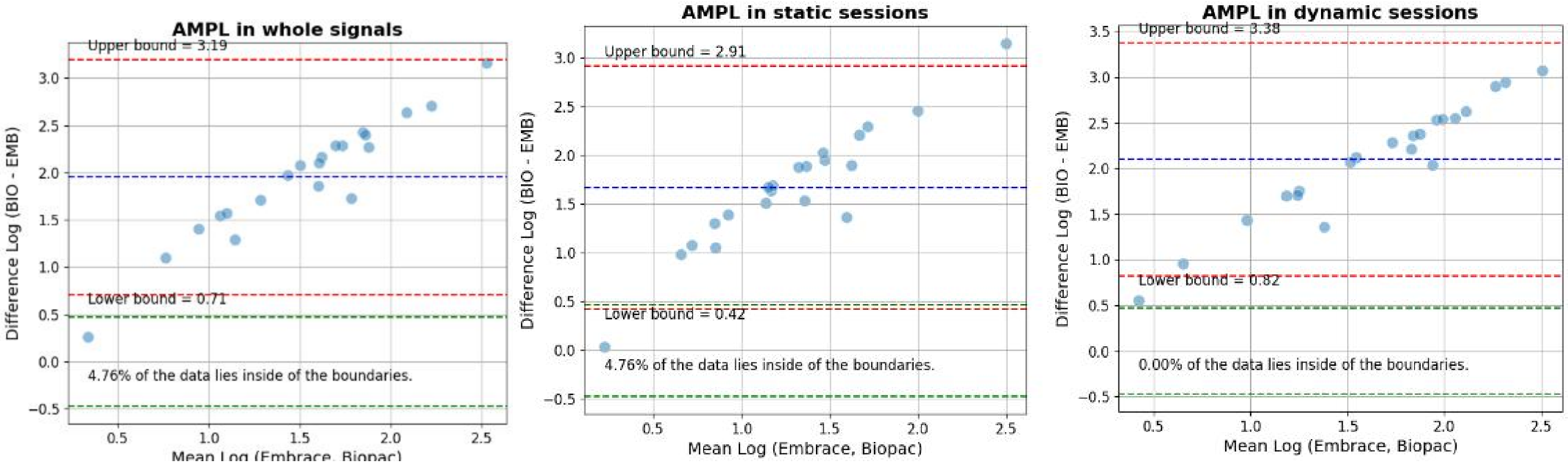
Participant28



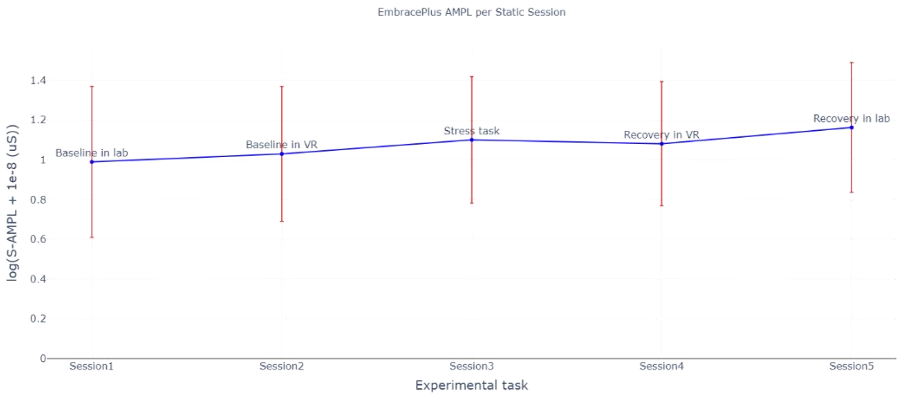
3. Data analysis materials

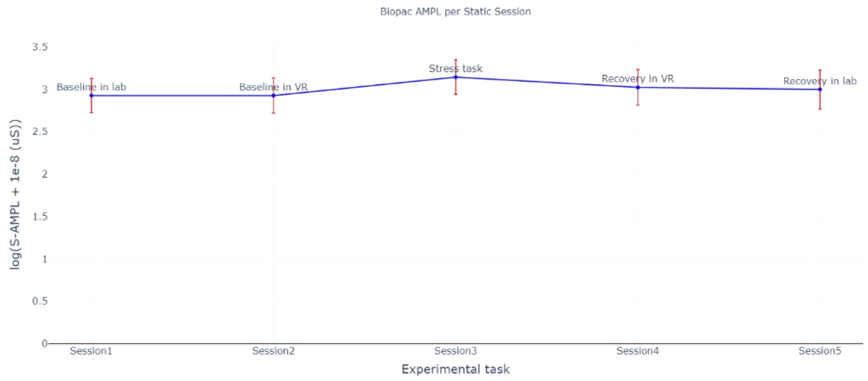
3.1 Bland–Altman plots of three parameter in whole signals, signals within static sessions, and signals within dynamic sessions

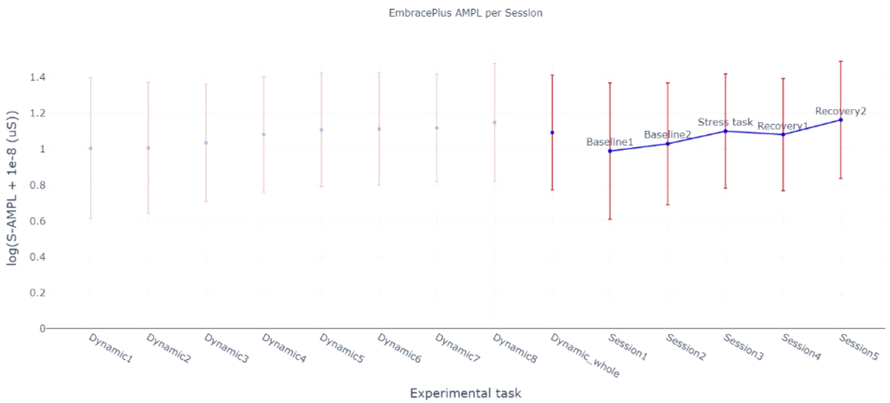


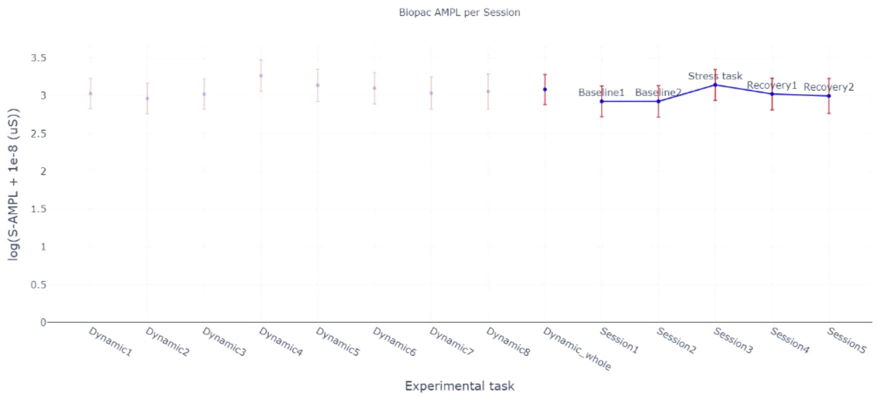

AMPL in whole signals
Upper bound = 3.19
Lower bound = 0.71
4.76% of the data lies inside of the boundaries.
Difference Log (BIO - EMB)
Mean Log (Embrace, Biopac)
AMPL in static sessions
Upper bound = 2.91
Lower bound = 0.42
4.76% of the data lies inside of the boundaries.
Difference Log (BIO - EMB)
Mean Log (Embrace, Biopac)
AMPL in dynamic sessions
Upper bound = 3.38
Lower bound = 0.82
0.00% of the data lies inside of the boundaries.
Difference Log (BIO - EMB)
Mean Log (Embrace, Biopac)

3.2 Event detection plots of SCL

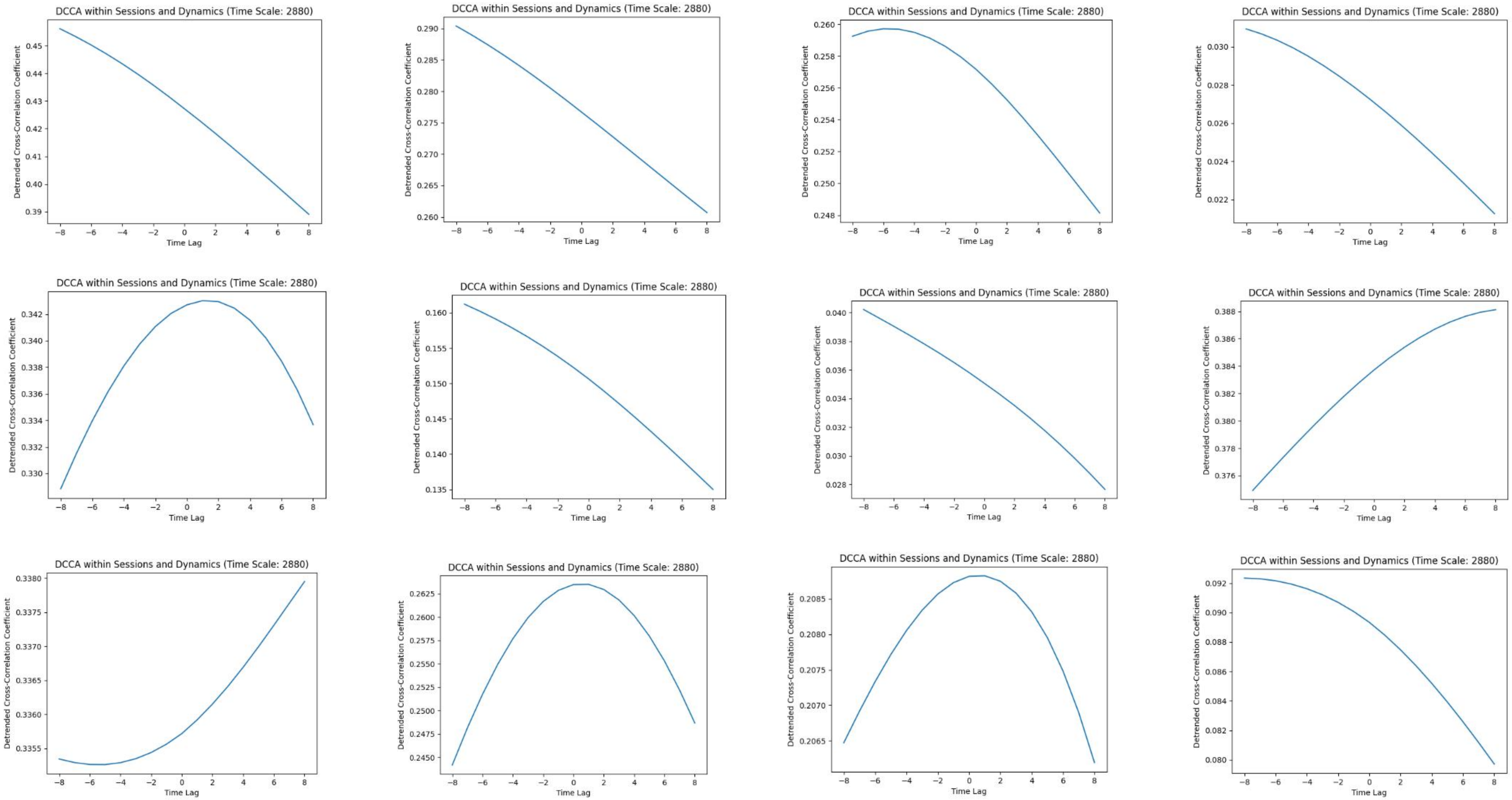


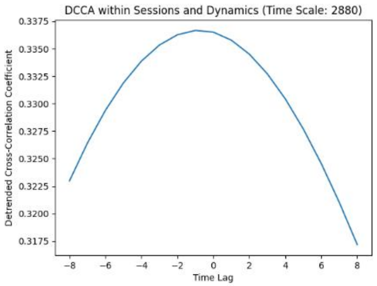


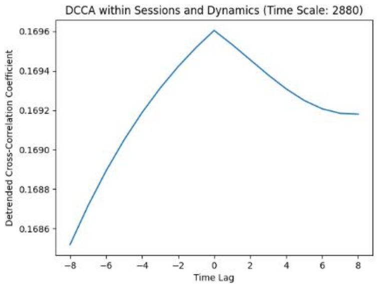


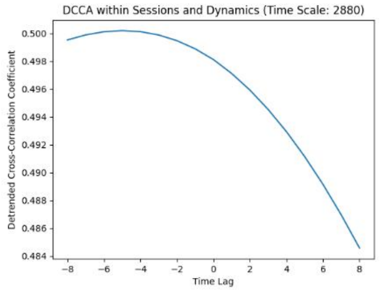


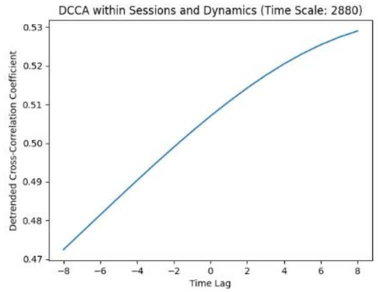
3.3 Time lagged DCCA of each participant

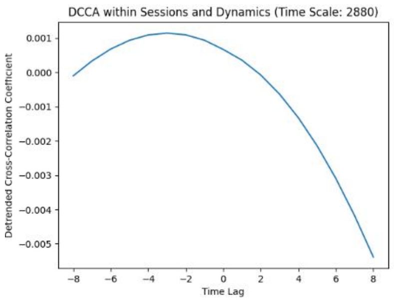


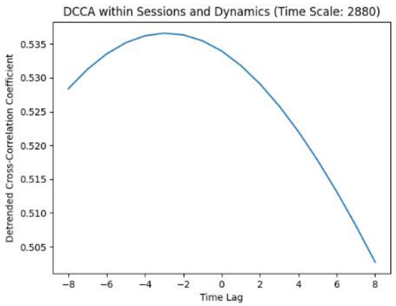

DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.3375
0.3350
0.3325
0.3300
0.3275
0.3250
0.3225
0.3200
0.3175
−8 −6 −4 −2 0 2 4 6 8
Time Lag

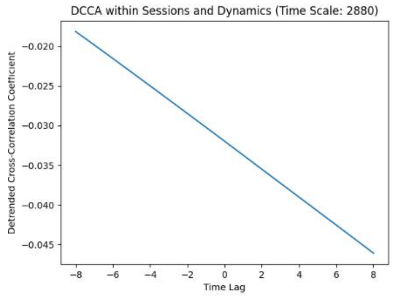

DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.1696
0.1694
0.1692
0.1690
0.1688
0.1686
−8 −6 −4 −2 0 2 4 6 8
Time Lag

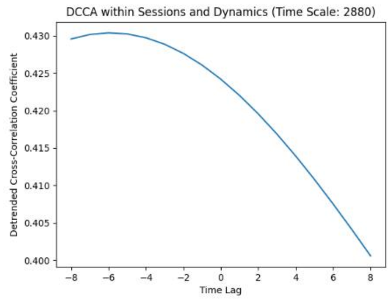

DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.500
0.498
0.496
0.494
0.492
0.490
0.488
0.486
0.484
−8 −6 −4 −2 0 2 4 6 8
Time Lag

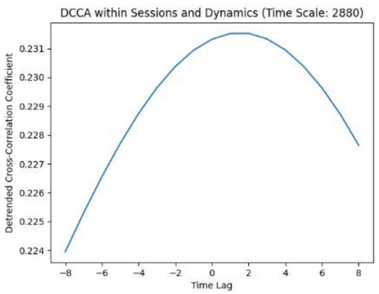

DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.53
0.52
0.51
0.50
0.49
0.48
0.47
−8 −6 −4 −2 0 2 4 6 8
Time Lag


DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.001
0.000
−0.001
−0.002
−0.003
−0.004
−0.005
−8 −6 −4 −2 0 2 4 6 8
Time Lag


DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.535
0.530
0.525
0.520
0.515
0.510
0.505
−8 −6 −4 −2 0 2 4 6 8
Time Lag


DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
−0.020
−0.025
−0.030
−0.035
−0.040
−0.045
−8 −6 −4 −2 0 2 4 6 8
Time Lag


DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.430
0.425
0.420
0.415
0.410
0.405
0.400
−8 −6 −4 −2 0 2 4 6 8
Time Lag


DCCA within Sessions and Dynamics (Time Scale: 2880)
Detrended Cross-Correlation Coefficient
0.231
0.230
0.229
0.228
0.227
0.226
0.225
0.224
−8 −6 −4 −2 0 2 4 6 8
Time Lag

3.4 DPCCA coefficients among six variables

Participant	V1-V2	V1-V3	V1-V4	V1-V5	V1-V6	V2-V3	V2-V4	V2-V5	V2-V6	V3-V4	V3-V5	V3-V6	V4-V5	V4-V6	V5-V6
1	0.44	-0.54	-0.44	0.13	-0.27	0.42	0.16	0.38	-0.06	0.03	0.06	-0.48	0.09	-0.11	0.27
2	0.03	-0.47	0.18	-0.01	-0.13	-0.10	0.02	-0.14	-0.68	-0.24	0.04	-0.02	0.04	0.15	-0.14
3	0.37	0.17	0.04	-0.01	0.08	0.02	0.02	0.57	-0.21	0.48	0.00	-0.20	-0.05	-0.15	0.21
4	0.66	0.49	0.54	-0.30	0.26	-0.43	-0.22	0.57	-0.61	-0.31	0.26	-0.08	0.06	0.05	0.48
5	0.07	0.69	0.11	0.09	0.09	0.34	-0.36	0.45	-0.75	0.29	-0.22	0.14	0.14	-0.22	0.44
6	0.28	-0.39	-0.04	-0.31	-0.15	0.24	0.09	0.72	-0.22	-0.23	-0.20	-0.37	-0.07	0.05	0.17
7	0.24	-0.24	-0.03	0.29	-0.06	0.24	-0.12	0.21	-0.49	0.04	0.20	-0.06	0.32	0.15	0.16
8	0.45	0.30	0.23	-0.25	0.19	0.02	0.04	0.28	-0.41	-0.21	0.12	0.21	-0.11	0.02	0.17
9	0.43	0.31	-0.14	0.56	0.08	-0.07	-0.17	0.03	-0.16	-0.15	-0.33	-0.37	0.05	0.09	-0.05
10	0.17	-0.08	0.11	0.23	-0.04	0.08	0.01	0.01	-0.53	0.45	-0.22	-0.19	0.44	0.05	0.11
11	0.53	0.21	0.31	-0.27	0.21	-0.15	-0.38	0.31	-0.32	0.30	-0.20	-0.65	0.20	0.27	0.04
12	-0.15	-0.64	-0.23	0.35	0.03	-0.36	-0.12	0.71	-0.22	-0.19	0.52	-0.14	0.14	0.05	0.17
13	0.04	-0.14	0.10	-0.07	-0.12	-0.23	0.03	0.37	-0.33	-0.53	0.22	0.28	0.27	0.46	0.04
14	0.25	0.28	0.04	0.17	0.16	0.37	0.40	0.03	-0.50	-0.22	-0.23	-0.08	0.56	0.27	-0.04
Mean	0.27	0.00	0.06	0.04	0.02	0.03	-0.04	0.32	-0.39	-0.04	0.00	-0.14	0.15	0.08	0.15
Min	-0.15	-0.64	-0.44	-0.31	-0.27	-0.43	-0.38	-0.14	-0.75	-0.53	-0.33	-0.65	-0.11	-0.22	-0.14
Max	0.66	0.69	0.54	0.56	0.26	0.42	0.40	0.72	-0.06	0.48	0.52	0.28	0.56	0.46	0.48

Note. V1 = Raw signals from EmbracePlus. V2 = Raw signals from Biopac. V3 = Wrist temperature. V4 = Ambient temperature. V5 = Dummy variable of stressor. V6 = Dummy variable of static periods.

3.5 Python scripts for data analysis

```
<p>import pandas as pd
import neurokit2 as nk
from scipy.io import loadmat
from dcca import detrended_correlation
import matplotlib.pyplot as plt
import numpy as np
from plotnine import *
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.graph_objects as go
from scipy.stats import shapiro</p>
<h1>Define the sampling rates</h1>
<p>sampling_rate_embrace_eda = 4 # Hz
sampling_rate_biopac = 2000
#-------------------STEP 1------------------------</p>
<h1>Step 1: Import Processed EmbracePlus and Biopac EDA data</h1>
<h1><s><s><s>EmbracePlus</s></s></s></h1>
<p>embrace_data = pd.read_csv(f&quot;{file_path}&quot;)
embrace_data.set_index('timestamp', inplace=True)
embrace_data.index = pd.to_datetime(embrace_data.index)
Start_time_embrace = embrace_data.index[0]</p>
<h1><s><s><s>Biopac</s></s></s></h1>
<p>eda_mat_file = f&quot;{file_path}&quot;
mat_data_eda = loadmat(eda_mat_file)
biopac_eda_data = mat_data_eda['EDA_biopac'].squeeze()
start_timestamp_bio = mat_data_eda['start_time'].squeeze()
Start_time_biopac = pd.to_datetime(start_timestamp_bio)
biopac_index = pd.date_range(start=Start_time_biopac, periods=len(biopac_eda_data), 
freq=f&quot;{1/sampling_rate_biopac}S&quot;)
biopac_data = pd.DataFrame({'EDA': biopac_eda_data}, index=biopac_index)
#2 Read the processed event timestamps file
input_file = f&quot;{file_path}&quot;
with open(input_file, 'r') as file:
lines = file.readlines()
event_timestamps_filtered = pd.to_datetime([line.strip() for line in lines])
#-------------------STEP 2------------------------
#STEP 2: Resample to 16 Hz
common_sampling_rate = 16
#---------Function---------</p>
<h1>Resamples the EDA signal to the specified sampling frequency and recreates the timestamps based 
on the starting time and the new sampling frequency.</h1>
<p>def resample_with_timestamps(df, column_name, device_name, starting_time, 
original_sampling_freq, desired_sampling_freq):
#Resample the EDA signal using NeuroKit
```

```
resampled_signal = nk.signal_resample(df[column_name], sampling_rate=original_sampling_freq,
desired_sampling_rate=desired_sampling_freq)
#Recreate the timestamps based on the starting time and the new sampling frequency
new_timestamps = pd.date_range(start=starting_time, periods=len(resampled_signal),
freq=f&quot;{1/desired_sampling_freq}s&quot;)
#Create a new DataFrame with the resampled signal and the new timestamps
resampled_df = pd.DataFrame({f'EDA_{device_name}': resampled_signal}, index=new_timestamps)
resampled_df.rename(columns={column_name:f'EDA_{device_name}'}, inplace=True )
return resampled_df</p>
<h1><s><s><s>EmbracePlus</s></s></s></h1>
<p>resampled_embrace = resample_with_timestamps(embrace_data, &quot;eda&quot;,
&quot;embrace_plus&quot;, Start_time_embrace, sampling_rate_embrace_eda,
common_sampling_rate)</p>
<h1><s><s><s>Biopac</s></s></s></h1>
<p>resampled_biopac = resample_with_timestamps(biopac_data, &quot;EDA&quot;, &quot;biopac&quot;,
Start_time_biopac, sampling_rate_biopac, common_sampling_rate)
#--------------------STEP 3-------------------------
#STEP 3: Align the signals by generating a common index
start_time = max(resampled_embrace.index.min(), resampled_biopac.index.min())
end_time = min(resampled_embrace.index.max(), resampled_biopac.index.max())
#Reindexing before merging the signals into one dataframe (since all timestamps are different)
exp_data_embrace = resampled_embrace.loc[start_time:end_time]
exp_data_biopac = resampled_biopac.loc[start_time:end_time]</p>
<h1>Generate a common index</h1>
<p>common_index = pd.date_range(start=start_time, periods=len(exp_data_embrace),
freq='62.5ms')#for frequency 16Hz
exp_data_embrace = exp_data_embrace.reset_index()
exp_data_biopac = exp_data_biopac.reset_index()
exp_data_embrace = exp_data_embrace.drop(columns=[&quot;index&quot;])
exp_data_biopac = exp_data_biopac.drop(columns=[&quot;index&quot;])
data_aligned = pd.concat([exp_data_embrace, exp_data_biopac], axis=1)
data_aligned.index = common_index
data_aligned_cut = data_aligned[event_timestamps_filtered[0]:event_timestamps_filtered[10]]
#--------------------STEP 4-------------------------</p>
<h1>DCCA (Signal comparison)</h1>
<p>def detrended_correlation_for_a_range_of_lags(x, y, time_scale, time_lag_range):
correlations = []
time_lags = []
for time_lag in range(time_lag_range[0], time_lag_range[1] + 1):
corr = detrended_correlation(x, y, time_scale, time_lag)
correlations.append(corr)
time_lags.append(time_lag)
return correlations, time_lags</p>
<p>def combine_and_sort_dccas(dccas1, lags1, dccas2, lags2):</p>
<h1>Combine lags and dccas</h1>
<p>lags = np.concatenate((lags1, -np.array(lags2)))
dccas = np.concatenate((dccas1, dccas2))</p>
```

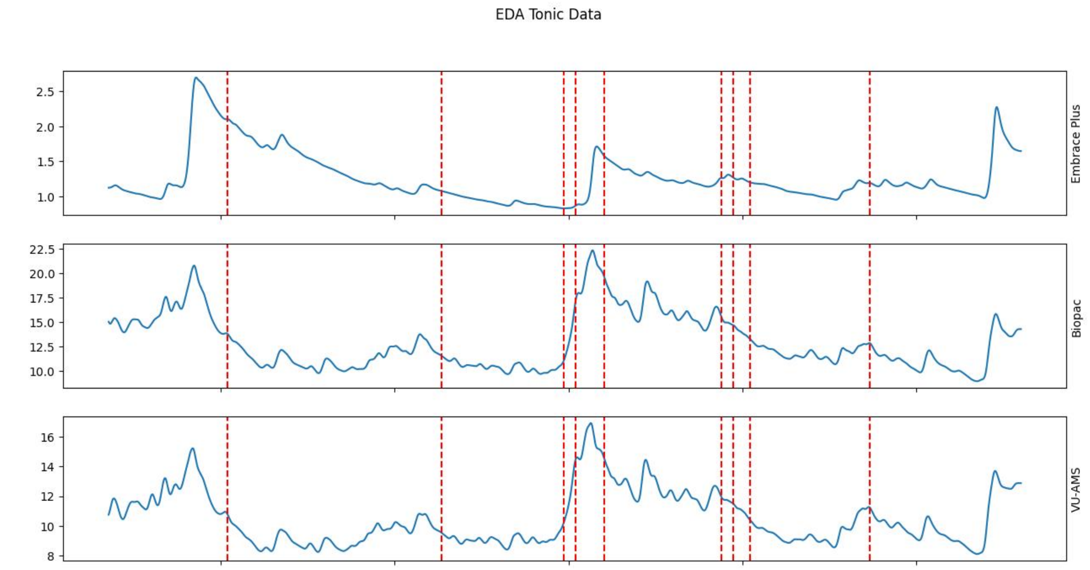
```
<h1>Create DataFrame</h1>
<p>df = pd.DataFrame({'lag': lags, 'dcca': dccas})</p>
<h1>Sort by lag</h1>
<p>df = df.sort_values(by='lag')
return df</p>
<h1>Prepare two arrays for calculation</h1>
<p>column_embrace = data_aligned_cut['EDA_embrace_plus']
embrace_array = column_embrace.to_numpy()
column_biopac = data_aligned_cut['EDA_biopac']
biopac_array = column_biopac.to_numpy()
time_scale = 2880 # 16<em>60</em>3 = sample number of 3 minutes
time_lag_range1 = (0,8)
dccas1, lags1 = detrended_correlation_for_a_range_of_lags(biopac_array, embrace_array, time_scale,
time_lag_range1)
time_lag_range2 = (1,8)
dccas2, lags2 = detrended_correlation_for_a_range_of_lags(embrace_array, biopac_array, time_scale,
time_lag_range2)
combined_df = combine_and_sort_dccas(dccas1, lags1, dccas2, lags2)</p>
<h1>Plot the Detrended Cross-Correlation Coefficients</h1>
<p>plt.plot(combined_df['lag'], combined_df['dcca'])
plt.xlabel(&quot;Time Lag&quot;)
plt.ylabel(&quot;Detrended Cross-Correlation Coefficient&quot;)
plt.title(f&quot;DCCA within Sessions and Dynamics (Time Scale: {time_scale})&quot;)
plt.show()
max_index = combined_df['dcca'].idxmax()
max_dcca = combined_df.loc[max_index, 'dcca']
max_dcca_lag = combined_df.loc[max_index, 'lag']
#-------------------STEP 5------------------------</p>
<h1>Function for Bland-Altman plots (Parameter comparison)</h1>
<p>def bland_altman_plot_symmetriclog(df, data_name, accep_limit):
def symmetric_log_transform(x):
epsilon = 1e-8
return np.sign(x) * np.log1p(np.abs(x) + epsilon)</p>
<pre><code>df['mean_log'] = df['mean'].apply(symmetric_log_transform)
df['diff_log'] = df['diff'].apply(symmetric_log_transform)
accep_limit_log =symmetric_log_transform(accep_limit)
sd_diff = df['diff_log'].std()
mean_diff = df['diff_log'].mean()
plt.figure(figsize=(10, 6))
sns.scatterplot(x=df['mean_log'], y=df['diff_log'], data=df, s=100, alpha=0.5)
plt.axhline(mean_diff, color='blue', linestyle='--')
plt.axhline(mean_diff + 1.96 * sd_diff, color='red', linestyle='--')
plt.axhline(mean_diff - 1.96 * sd_diff, color='red', linestyle='--')
plt.axhline(accep_limit_log, color='green', linestyle='--')
plt.axhline(-accep_limit_log, color='green', linestyle='--')
plt.text(df['mean_log'].min(), mean_diff + 1.96 * sd_diff + 0.1,
        f'Upper bound = {mean_diff + 1.96 * sd_diff:.2f}', fontsize=12)
```

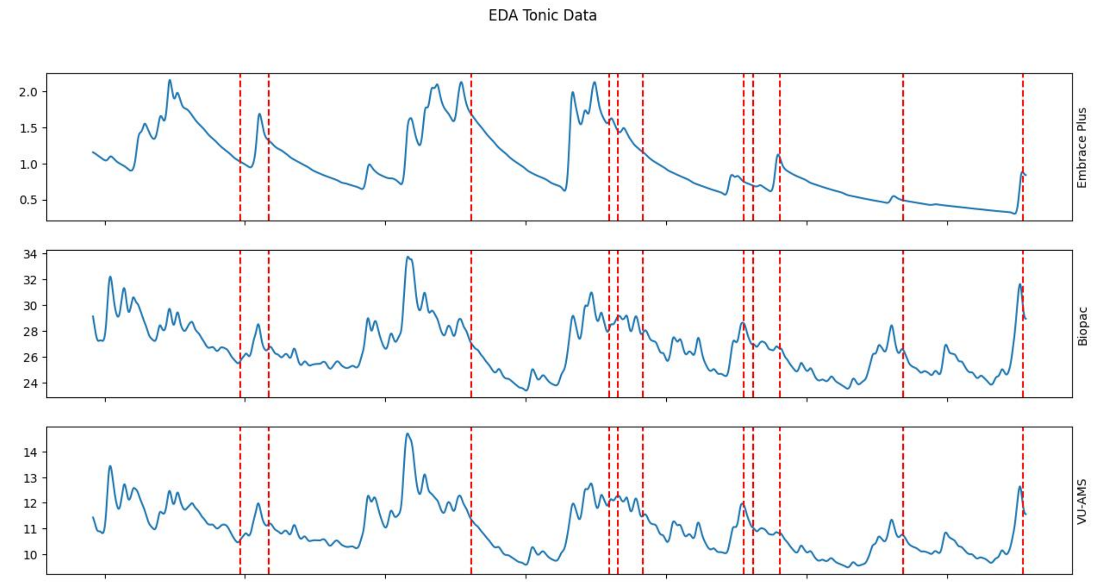
```
plt.text(df['mean_log'].min(), mean_diff - 1.96 * sd_diff - 0.1,
        f'Lower bound = {mean_diff - 1.96 * sd_diff:.2f}', fontsize=12)
inside_bounds = (df['diff_log'] &gt; -accep_limit_log) &amp; (df['diff_log'] &lt; accep_limit_log)
pct_inside = inside_bounds.mean() * 100
print(pct_inside)
plt.text(df['mean_log'].min(), -0.2,
        f'{pct_inside:.2f}% of the data lies inside of the boundaries.', fontsize=12)
plt.title(f'{data_name}', fontsize=16, fontweight='bold')
plt.xlabel('Mean Log (Embrace, Biopac)', fontsize=14)
plt.ylabel('Difference Log (BIO - EMB)', fontsize=14)
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)
plt.grid(True)
plt.show()
</code></pre>
<p>def bland_altman_plot_nolog(df, data_name, accep_limit):
sd_diff = df['diff'].std()
mean_diff = df['diff'].mean()
plt.figure(figsize=(10, 6))
sns.scatterplot(x=df['mean'], y=df['diff'], data=df, s=100, alpha=0.5)
plt.axhline(mean_diff, color='blue', linestyle='--')
plt.axhline(mean_diff + 1.96 * sd_diff, color='red', linestyle='--')
plt.axhline(mean_diff - 1.96 * sd_diff, color='red', linestyle='--')
plt.axhline(accep_limit, color='green', linestyle='--')
plt.axhline(-accep_limit, color='green', linestyle='--')
plt.text(df['mean'].min(), mean_diff + 1.96 * sd_diff + 0.1,
f'Upper bound = {mean_diff + 1.96 * sd_diff:.2f}', fontsize=12)
plt.text(df['mean'].min(), mean_diff - 1.96 * sd_diff - 0.1,
f'Lower bound = {mean_diff - 1.96 * sd_diff:.2f}', fontsize=12)
inside_bounds = (df['diff'] &gt; -accep_limit) &amp; (df['diff'] &lt; accep_limit)
pct_inside = inside_bounds.mean() * 100
print(pct_inside)
plt.text(df['mean'].min(), -0.2,
f'{pct_inside:.2f}% of the data lies inside of the boundaries.', fontsize=12)
plt.title(f'{data_name}', fontsize=16, fontweight='bold')
plt.xlabel('Mean (Embrace, Biopac)', fontsize=14)
plt.ylabel('Difference (BIO - EMB)', fontsize=14)
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)
plt.grid(True)
plt.show()</p>
<p>def calculate_shapiro_and_plot(df, df_name, acceptable_limit):
array = df['diff']
statistic, pvalue = shapiro(array)
print(statistic)
alpha = 0.05<br>
if pvalue &gt; alpha:
```

```
bland_altman_plot_nolog(df, df_name, acceptable_limit)
else:
bland_altman_plot_symmetriclog(df, df_name, acceptable_limit)</p>
<p>#-------------------STEP 6-------------------------</p>
<h1>Function for error bar plot (Event detection comparison)</h1>
<p>def vertical_data_Eventdetection(df):
df = pd.DataFrame(df)
melted_df = df.melt(id_vars='part_id', var_name='Categories', value_name='Values')
melted_df = melted_df.sort_values(by=['part_id', 'Categories'])
return melted_df</p>
<p>def event_detection_static_plot(df, df_name, boundary=None):
def symmetric_log_transform(x):
epsilon = 1e-8
return np.sign(x) * np.log1p(np.abs(x) + epsilon)
df['Values_log'] = df['Values'].apply(symmetric_log_transform)
df_grouped = df.groupby(&quot;Categories&quot;).agg(
mean=(&quot;Values_log&quot;, &quot;mean&quot;),
sd=(&quot;Values_log&quot;, &quot;std&quot;),
count=(&quot;Values_log&quot;, &quot;count&quot;)
).reset_index()
df_grouped['SE'] = df_grouped['sd'] / np.sqrt(df_grouped['count'])
fig = go.Figure()
fig.add_hline(y=0, line_color=&quot;black&quot;, line_width=1)
if boundary is not None:
fig.add_hline(y=boundary, line_color=&quot;green&quot;, line_width=1)
fig.add_hline(y=-boundary, line_color=&quot;green&quot;, line_width=1)
# Add error bars with SE and customize line transparency
for _, row in df_grouped.iterrows():
category = row[&quot;Categories&quot;]
color = 'blue'
error_color = 'red'<br>
fig.add_trace(
go.Scatter(
x=[category],
y=[row[&quot;mean&quot;]],
mode='markers',
marker=dict(color=color),
error_y=dict(
type='data',
array=[row[&quot;SE&quot;] * 1.96],  # Assuming 95% confidence interval
visible=True,
color=error_color,
thickness=1.5,
width=3,
)
)
)
```

```
# Connect the points between Session1-5
fig.add_trace(
go.Scatter(
x=df_grouped[&quot;Categories&quot;],
y=df_grouped[&quot;mean&quot;],
mode='lines+markers',
line=dict(color='blue', width=2),
marker=dict(color='blue'),
opacity=1.0
)
)
fig.update_xaxes(title_text=&quot;Experimental task&quot;)
fig.update_yaxes(
title_text=r&quot;log(S-AMPL + 1e-8 (&quot; + str('u') + &quot;S))&quot;<br>
)
fig.update_layout(
title=df_name,
font=dict(size=16),<br>
legend_title_text=&quot;&quot;,<br>
xaxis_title_font_size=20,
yaxis_title_font_size=20,
title_font_size=16,
title_x=0.5,  # Center title horizontally
plot_bgcolor='rgba(0,0,0,0)',
paper_bgcolor='rgba(0,0,0,0)'
)
fig.update_xaxes(tickvals=df_grouped[&quot;Categories&quot;],
ticktext=df_grouped[&quot;Categories&quot;])
fig.show()</p>
```

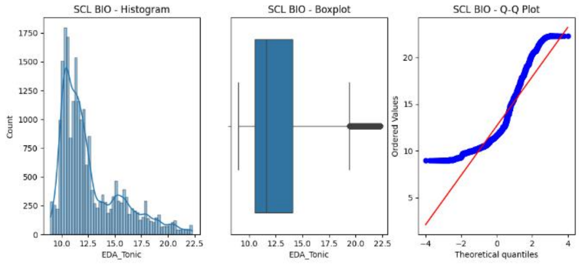
3.6 Two examples of visual resemblance



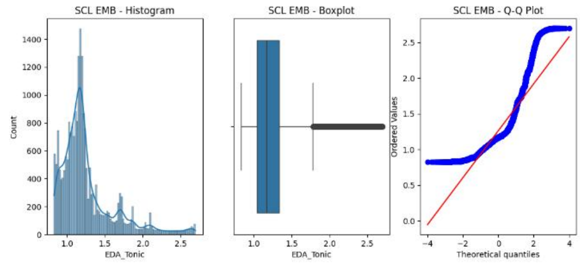


3.7 Visual inspection of SCL's distribution and normality per participant and device (5 examples)

Participant 1 (Biopac vs. EmbracePlus)

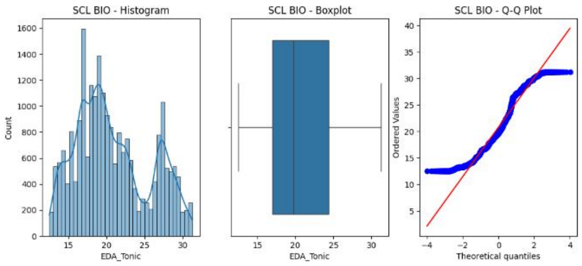


3% of the data lies outside of the boundaries.

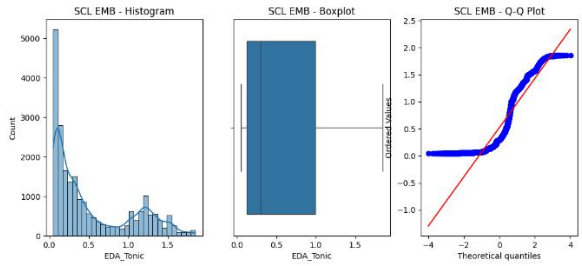


9% of the data lies outside of the boundaries.

Participant 3 (Biopac vs. EmbracePlus)

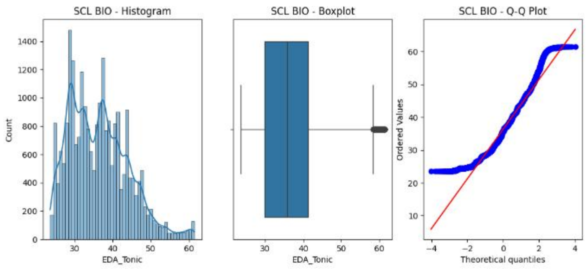


0% of the data lies outside of the boundaries.

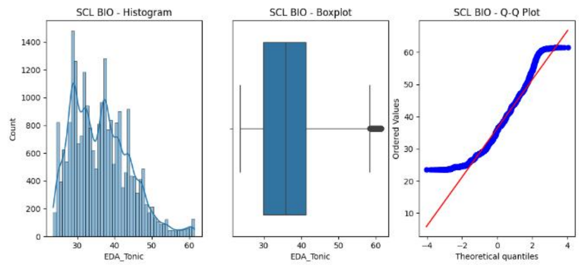


0% of the data lies outside of the boundaries.

Participant 4 (Biopac vs. EmbracePlus)

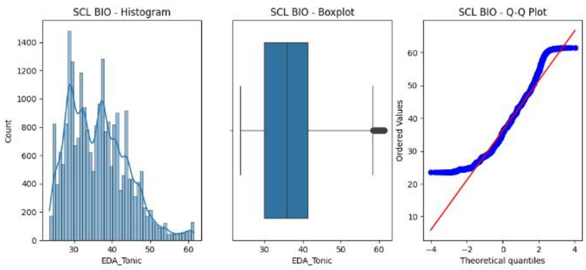


1% of the data lies outside of the boundaries.

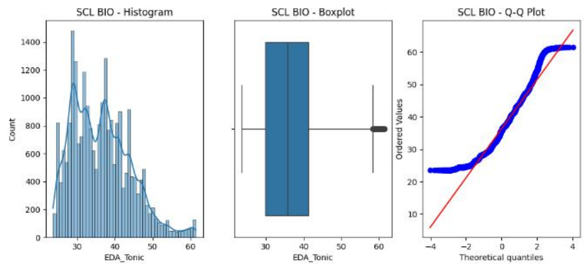


1% of the data lies outside of the boundaries.

Participant 5 (Biopac vs. EmbracePlus)

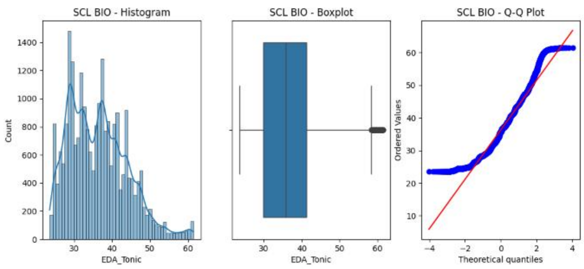


1% of the data lies outside of the boundaries.

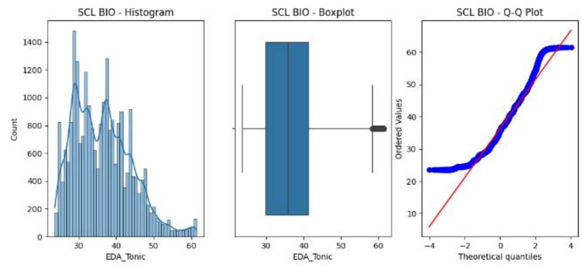


1% of the data lies outside of the boundaries.

Participant 6 (Biopac vs. EmbracePlus)



1% of the data lies outside of the boundaries.



1% of the data lies outside of the boundaries.