Validating Accuracy of Embrace Plus in Measuring Electrodermal Activity During VR Stress Induction

Theresa Homann (s2997118)

Department of Psychology, University of Twente

202000381: Bachelor thesis Positive Clinical Psychology and Technology

Magdalena Sikora, Thomas Vaessen

18.06.2025

Word count: 6099

Abstract

Wearables have become widely used tools in ambulatory psychophysiological research and stress research. However, concerns remain about their accuracy in measuring stress-related responses, especially electrodermal activity (EDA), since the wrist placement can hinder accurate measurements due to a lower density of sweat glands. Therefore, multiple studies have tried to validate these devices, finding issues in the sensitivity for smaller stressors. Using Virtual Reality (VR) for this validation offers a safe, standardised way to test stress responses, especially when using more acute stressors. Therefore, this study aims to validate the accuracy of the Embrace Plus in measuring EDA compared to a golden standard using the three-level framework from Van Lier et al. (2019) including the signal level (overall shape and timing), the parameter level (accuracy of measured parameters) and the event level (across scenarios). The participants were exposed to the VR version of the TSST and a virtual height exposure scenario. As expected, at the signal level, agreement between the devices was low (*M=0.05*). At the parameter level, the Embrace Plus underestimated all three parameters, with the number of SCRs being the most accurate, with a moderate agreement, against previous expectations. For the event-level, responses for the phases were detected by the wearable, but the responses were significantly less prominent. All in all, the results suggest that the Embrace Plus consistently underestimates EDA and cannot accurately detect all trends. These results may be caused by a misalignment of the devices, caused by a temporal difference between the two.

Validating Accuracy of Embrace Plus in Measuring Electrodermal Activity During VR Stress Induction.

Over the past few years, emotional stress has increased worldwide, with around 30-50% of the population reportedly experiencing psychological stress (Piao et al., 2024). These numbers highlight the urgent need to address this global issue. One advancement in this field is the development of wearable devices that measure physiological and behavioural stress responses (Roos & Slavich, 2023). Monitoring physiological activities like electrodermal activity (EDA) can improve individuals' stress management skills by improving their ability to control certain stress responses (De Witte et al., 2019; Dillon et al., 2016). While these wearables can measure multiple stress responses, EDA is one of the most reliable physiological parameters to measure when looking at stress responses (Hanshans et al., 2024) and is seen as a valuable biomarker for stress management (Klimek et al., 2023). However, the increasing use of wearable technology raises the question of whether these devices are similarly accurate in measuring stress as gold-standard devices, which are validated in accuracy. Studies have found that wrist-worn wearable devices might suffer from a reduced quality due to fewer responses, a lower number of sweat glands on the wrist, and potential movement artefacts (Menghini et al., 2019; Van Lier et al., 2019). Therefore, testing the accuracy of these devices in controlled settings is essential. Virtual reality (VR) is a technology able to create stress-inducing scenarios that simulate real-life experiences while being standardizable and controllable (Halbig & Latoschik, 2021; Van Dammen et al., 2022).

While stress is a natural human response to challenges or threats, and can even be helpful in small amounts, too much of it can cause both physical and mental health problems such as anxiety, depression, heart diseases and high blood pressure (World Health Organization [WHO], n.d.; Mayo Clinic, n.d.). There has been some research into interventions that are effectively able to improve stress management by using biofeedback from EDA (De Witte et al., 2019). An invention with smartphone games applying biofeedback to teach users stress management skills showed that monitoring one’s EDA level makes it easier for individuals to control their stress responses and can help teach relaxation techniques (Dillon et al., 2016).

EDA refers to fluctuations in the electrical conductance of the skin caused by sweat gland activity, influenced by the sympathetic nervous system (SNS) (Texas Instruments [TI], 2023). Activity in the SNS is an important indicator of stress levels (De Geus & Gevonden, 2024), and to analyse the activity in the SNS, Boucsein (2012) defines two activities that are

measured in EDA. The slow tonic shifts in the basal skin conductance level (SCL) and faster phasic transient events are both influenced by emotional stress. They argue that, firstly, from the tonic shifts, the mean SCL of a certain defined period is often looked at for analysing the general arousal state. Secondly, for the phasic component, the focus lies on the skin conductance responses (SCR). For this component, commonly analysed features include the number of SCRs and the total SCR amplitude. The number of SCRs analysed in a given time window acts as a frequency-based measure of activity in the sympathetic nervous system, providing insights about how often responses are triggered in a certain condition. The total SCR amplitude sums the amplitudes of all detected SCRs in a certain period and integrates both the frequency and the size of responses. This parameter is a quantification of the overall magnitude of phasic activity and can be used to compare the overall reactivity across conditions or devices. Both components of EDA are widely used in studies measuring emotional arousal (Lang et al., 1993).

Since the development of wearable devices, interventions aiming to improve stress management have become even more accessible and with only low costs (De Witte et al., 2019). Due to easier accessibility and generally larger acceptance of participants, wearables like watches are very prominent on the market (De Geus & Gevonden, 2024). Wearable technology has shown to be an innovation for research in psychophysiology, especially in assessing stress reactivity (Van der Mee et al., 2021). If wearable EDA sensors could achieve similar accuracy comparable to other measurement techniques, they could potentially improve our ability to monitor chronic stress in daily life (Klimek et al., 2023). Another advantage of these devices could be to make monitoring of patients easier and improve stress management techniques (Cahill et al., 2007, as cited in Klimek et al., 2023).

However, the accuracy of these devices is still questionable (Mühlen et al., 2021). De Geus and Gevonden (2024) list several issues with EDA recording of wearable devices, including a smaller sweat gland density at the wrist, leading to a limited level of electrodermal activity and a more prominent thermoregulatory sweating than emotional sweating. Additionally, sensor placement remains a critical factor, as even slightly different placements could already disturb the signal processing (Klimek et al., 2023). There is, however, not enough validation of wearables to date, and therefore, it is very important to accurately validate wrist-worn EDA sensors and make their usage as easy as possible.

Van Lier et al. (2019) developed a framework to validate wearable devices using a comprehensive protocol at three levels. Firstly, the Signal level determines whether the

overall shape and timing of the EDA signals from the wearable match that of the gold standard device. Next, the Parameter Level aims to assess agreement between the obtained parameters and is used to check whether the data from the wearable is useful in measuring stress responses. Lastly, the event level is used to determine whether the device can distinguish between different conditions. Van Lier et al. (2019) used their framework to validate the accuracy of the Embrace Plus and found that the total amplitude of skin conductance responses was valid at the event level. This was particularly true when assessing strong, sustained stressors. However, there were also some limitations of the wearable found such as the limited sensitivity of the devices causing them to not reliably detect minor stressors. Additionally, the application of the framework found that there are multiple factors that can affect the accuracy of the measures, such as the placement of the sensor and environmental factors.

Secondly, van der Mee et al. (2021) used wrist-worn devices with dry electrodes and compared the measurement with traditional palm-based wet electrodes using multilevel modelling. The findings indicate that wrist-based EDA measurements are feasible, but these devices tend to show lower absolute skin conductance levels (SCL) and non-specific skin conductance response (ns.SCR) frequencies than palm-based measures (Van der Mee et al., 2021). Moreover, when comparing the within-subject correlations, correlations between the two measures were only modestly significant (Van der Mee et al., 2021). This indicates that wrist-worn devices may have limitations in accurately capturing rather subtle changes in sympathetic nervous system activity (Van der Mee et al., 2021). This outlines the need to compare different levels of stress.

While the aforementioned studies were able to test the accuracy of wearable devices in experimental conditions, Klimek et al. (2023) argue that lab data and outcomes are often not translatable to daily life situations. Combined with the aforementioned issue of sensitivity, this suggests a need for more immersive and acute stress scenarios to simulate real-world stress responses to better assess the accuracy of wearables. Virtual reality (VR) offers a promising solution because it can create a measurable, stress-inducing environment that helps comparative validation of wearable and traditional EDA measures (Van Dammen et al., 2022). The advantages of a VR setting lie in the variability of possible stressful situations to place the participant in (Van Dammen et al., 2022). Using VR instead of regular, lab-based stress induction ensures that one can follow ethical guidelines when putting people in stressful or life-threatening situations, which are acute stressors that are deeply rooted in our physiology (Shirtcliff et al., 2024). Depending on the reaction that is to be measured, one can tailor the

experience in the VR setting to get the most accurate results and ensure standardizability of these stress measures (Van Dammen et al., 2022

This research uses VR-based stress simulation to assess the Embrace Plus watch's accuracy in measuring EDA. VR will simulate diverse stress scenarios, including social pressure and height-induced fear, as well as baseline and relaxation recordings. Signals from the Embrace Plus will be compared to a gold standard reference device using the three-level validation protocol of van Lier et al. (2019). The primary research question guiding this study is: *How accurately does the Embrace Plus watch measure the physiological stress response EDA compared to gold-standard laboratory measures across signal, parameter and event levels, as outlined by van Lier et al. (2019) in VR-based stress experiments?*

Based on the aforementioned findings in previous studies, it is hypothesised that the Embrace watch may not measure more subtle stress responses accurately, but strong changes in the sympathetic nervous system and can report only intense stressors. Additionally, it is suggested that the smart watch underestimates SCL and ns.SCR frequency, while the measurement of the total amplitude of the SCRs will be the most accurate when compared to the BIOPAC.

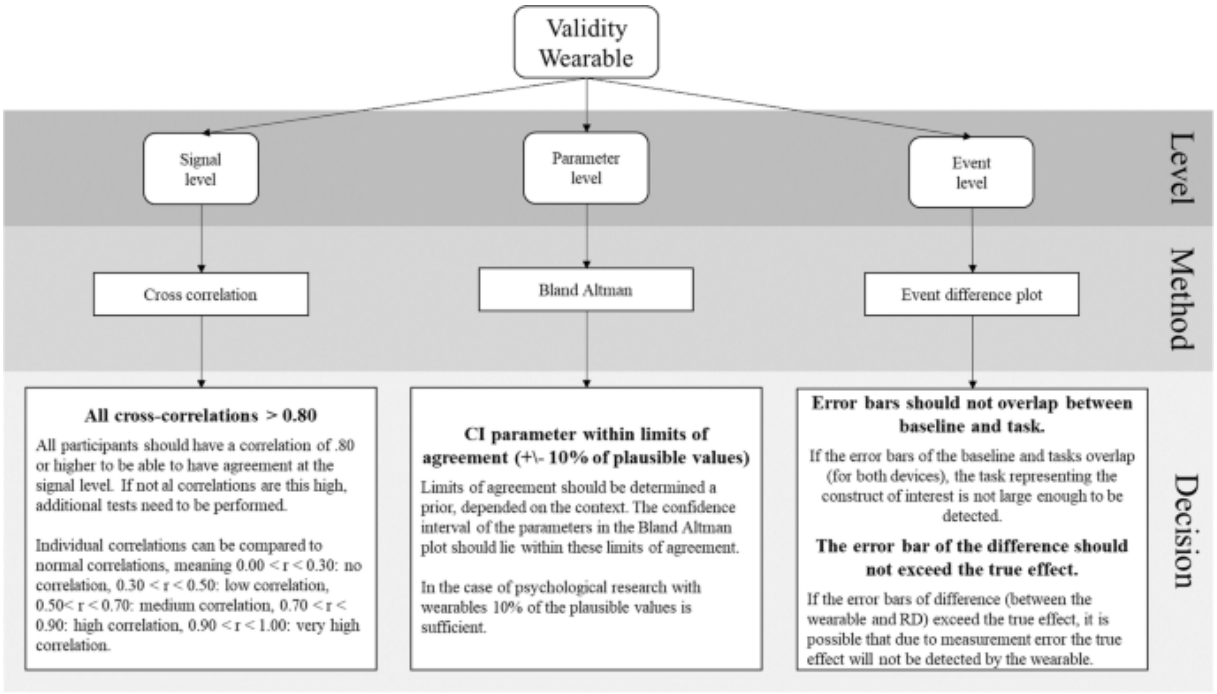
Methods

Design

This study has a within-subject experimental design and tests the accuracy of the Embrace Plus in measuring EDA using a three-level validation framework as proposed by van Lier et al. (2019). For an overview of the protocol, see Figure 1. The participants were exposed to two different VR-based stress-inducing scenarios introduced in a random order and a relaxation scenario. During these scenarios, their EDA responses were measured using the Embrace Plus watch and the BIOPAC MP160 System as a gold-standard lab device (Empatica, 2019; BIOPAC Systems, Inc., 1989). These signals were later analysed across the signal, parameter and event level to assess the validity of the wearable device.

Figure 1

Overview of the analysis protocol by Van Lier et al. (2019)



Participants

Participants were recruited using convenience sampling and were directly approached by the researchers or took part using the Sona system, from which they gained 3 Sona credits. One of the inclusion criteria was for them to be above the age of 18. Additionally, the participants had to be able to speak and understand English fluently. Participants had to satisfy some relevant health criteria because otherwise, the physiological responses could interfere with the data. They were excluded if they had a diagnosed heart disease or if they were taking medication that influences the autonomic nervous system. These medications include Beta-blockers, anti-sweating medications (acetylcholine blockers), heart rhythm stabilisers, tricyclic antidepressants, serotonin-norepinephrine reuptake inhibitors or selective serotonin reuptake inhibitors. Anti-sweating medications, for example, could change the EDA response of the participant and therefore affect the data and hinder the detection of stressors.

Materials

Wearable

In this study, the Embrace Plus watch was used as the main wearable data collecting device. The Embrace Plus watch is a wearable device that measures Electrodermal activity (EDA), skin temperature, heart rate and heart rate variability using a photoplethysmogram (PPG), and physical activity using an accelerometer (Empatica, 2019). This paper focusses on the measures of EDA only.

Gold standard

The BIOPAC MP160 system was used to measure an accurate, gold standard signal of EDA, ICG, ECG and movement (BIOPAC Systems, Inc., 1989). These values are compared to those of the wearable devices. For this paper, however, only the EDA signals are relevant. EDA was measured using Bionomadix EDA sensors and a PPGED-R receiver (BIOPAC Systems, Inc., 2010). The software AcqKnowledge was used to collect and save the data from the BIOPAC.

Virtual Reality (VR)

In this study, the Meta Quest 3 headset was used to put the participant in different VR scenarios (Meta, 2023). Meta Link was used to connect the VR headset to the computer so that the researchers could tailor the scenario that the participant is put in. Moreover, Unity was used to synchronise the VR scenarios with the Acqknowledge program. It did automatic timestamps for every part of the study.

Procedure

As beforementioned, this paper was part of a bigger study, and for a concise and clear overview of this part of the study, only the relevant parts are mentioned. Before coming to the lab, participants were asked to sign the informed consent form and to fill in a brief questionnaire about some background information, including gender, age and ethnicity, and their dominant hand. When arriving at the lab, participants were asked if they had any more questions about the informed consent form or the study itself. After that, they were informed that they could at any point stop the experiment in case they were getting uncomfortable. To start with the study, the Embrace Plus watch was placed on the non-dominant wrist above the wrist joint. Next, the BIOPAC EDA electrodes were attached to their non-dominant hand on the middle and ring fingers. These were dry EDA electrodes, and isotonic gel was applied to them to ensure a good signal. The electrodes were attached to the intermediate phalanges and secured using tape (see Figure 2). In order to build up some sweat so that the electrodes work better, the participants were asked to go up and down the stairs five times. Once all measures were attached to the participant, the participant was asked to do a set of movements for data synchronisation. These included putting the arms five times up and down and bending them from side to side.

Baseline measures

After that, baseline measures were taken. The participant was asked to read aloud a neutral text for three minutes and to fill out some additional questionnaires about sleep and physical activity; however, these were not relevant for this study. Next, the participant was asked to stand, sit and lie down still for two minutes, also regarded as a baseline recording. Once the participant was done with the baseline measures, they were placed in VR, and again, there were baseline measures taken while standing and sitting while wearing the VR headset for two minutes. After these measures, the participant was exposed to two different stress-inducing scenarios in VR. The order of these scenarios was randomised.

Trier Social Stress Test (TSST)

One of them was to resemble social stress. For this, a VR version of the Trier Social Stress Test (TSST) (Kirschbaum et al., 1993) was used. This test is used as the gold standard for assessing acute stress in a laboratory setting (Allen et al., 2017). The TSST was previously tested in a VR setting and was able to ensure greater efficiency and standardisation of this test (Shiban et al., 2016). Both the in vivo and the VR version of this TSST are proven to reliably evoke physiological stress responses (Shiban et al., 2016). The TSST is done by initiating a job interview scenario in which the person is asked to prepare for five minutes to hold a five-minute speech for their dream job, talking about their strengths and weaknesses. The participant is then asked to hold this five-minute speech in front of three judges. After the five minutes, the participant is asked to continuously subtract the number 7 from 1022 for five minutes, in which every mistake or if they are too slow leads to them having to start over from 1022. During both the speech part and the math portion, the experimenter was able to trigger prerecorded audio phrases that the VR character then said to the participant. These included phrases like “incorrect, start over from 1022”. These recordings made sure that the experimenter could somewhat guide the interaction to make sure that the participant continued speaking and generate more stress in the participant.

Plank task

The other stress-inducing scenario aimed to induce fear of height. The participant was asked to step into an elevator that went up a couple of floors, where they were asked to step onto a plank that was at the side of the building at a height equivalent to 80 meters. The participant was then asked to stand at the end of the plank for two minutes and look down at the road underneath. After these two minutes, the participant was told to go back to the elevator. Previous studies suggest that these kinds of fear-of-height-inducing scenarios are efficient in the induction of stress responses (Finseth et al., 2018). Before the first stress-

inducing scenario and after each of the stressors, the participant was asked to stand or sit still for two minutes as part of a recovery and to rate their stress level from 1-10 on a Likert-like scale slider. This ranking was also done after the first sitting baseline measure.

Relaxation Scenario

After the stress-inducing scenarios, the participant was put in a relaxation, nature scenario during which they were asked to touch butterflies on the ground for five minutes. This exercise is done in order to include movements to test the wearables' accuracy in measuring EDA when movement is introduced. Once the exercise was completed, the participant was asked one last time to rate their stress level and was taken out of VR. Lastly, the participant had to repeat the set of synchronisation movements. If the participant was interested in their data, the EDA signal was shown and explained to them in relation to the different tasks they underwent.

Figure 2

Attachment EDA electrodes

Data Analysis

To test how accurately the Embrace Plus watch measures EDA compared to a gold standard device, the validation analysis protocol by Van Lier et al. (2019) was used, which advises to assess validity on the signal, parameter and event level for the most optimal

validity overview. At first, the data was visually checked to see whether there were any failed measures for both the Embrace Plus watch and the BIOPAC.

Signal Level

For the signal level, a cross-correlation analysis was used in order to compare the EDA measures from the wearable device with the gold-standard device in terms of whether the shape and timing of the two match. First, the data of the BIOPAC was down-sampled from 2000 Hz, and the data of the Embrace Plus was up-sampled from 4 Hz so that they have the same frequency of 16 Hz. Next, the data was normalised so that it is better comparable without losing any important information. Moreover, the data was detrended to ensure that the focus is on shape and not long-term shifts, making it stationary. For this step, the slow baseline drifts were removed. After that, the actual cross-correlation was computed across a range of lags from -8 to +8. Each of the lags is representative of a delay of ± 0.5 seconds. Lastly, the highest cross-correlation for all participants was identified, and these were plotted in a histogram. This was done in order to have an overview of the optimal cross-correlations.

Parameter level

The Parameter level was analysed by looking at the key values extracted from the EDA signal. These include the mean of the skin conductance level (SCL), the number of Skin Conductance responses per minute (SCR count), and the total SCR amplitude. All of these measures were compared between the wearable and the gold-standard device to see how closely they match. For this, the parameters were calculated over defined time windows for both devices, starting from the Baseline Reading until after the nature task. Additionally, log Bland-Altman plots were generated in order to visualise any agreement and to detect potential biases between the two measures. The boundaries for the Bland-Altman plots were set based on the paper from van Lier et al. (2019). For the mean SCL, the boundary is ± 0.96, for the SCR count ±1.25 and for the SCR total amplitude ±0.47.

Event level

Lastly, the event level checks whether the Embrace Plus can detect any physiological changes across different scenarios or events, similarly to the BIOPAC. These include, for example, stress and relaxation. First, the different scenarios were defined, including the TSST part, which includes a Baseline seated, the TSST anticipation phase, the TSST speech phase, the mental arithmetic and lastly the recovery from the TSST. Secondly, the plank task included a standing baseline, the plank anticipation phase, the walk on the plank, the baseline

of standing on the plank and the recovery phase. The mean SCL was chosen as the stress parameter due to its easy comparability. Next, it was checked for normality and any missing data. Lastly, the effect was visualised by creating event difference plots of the data. Firstly, a line plot with the log means and the SE per task, including a line through the zero y-axis, with each individual represented in a line. And secondly, a line plot including the log differences between the Embrace Plus and BIOPAC for each person represented by a line. Both the mean and the SE were included per task.

Results

There were 20 participants taking part in this study, of which 2 were excluded because data was missing due to technical issues for participant 1, and participant 9 was feeling nauseous before entering VR. Additionally, participants 2 and 10 were removed from the data set due to errors in data collection. Lastly, participants 14 and 16 were excluded because of data from the BIOPAC close to 0 or even negative and participants 3, 8 and 18 due to missing data from the Embrace Plus. All of these were removed from the data set, resulting in 11 final participants (*male=7; female=4*). The age of the participants ranged from 20 to 26 (*M=23.27; SD=2.1*). Most participants were white (10), and one participant was East Asian.

Signal level

For 10 of the participants, the correlation is considered very low, below 0.25, and only for one participant the correlation was high, above 0.8 (see Figure 3). Out of all participants, seven even had a negative correlation. Therefore, the mean maximum cross-correlation coefficient between the signals from the Embrace Plus watch and the BIOPAC was $r = 0.05$ ($SD = 0.27$). This suggests very low average agreement (>0.25) in signal and timing between the two devices. For the optimal lag, there was also only one participant, who had the optimal lag at 0, with all other participants at -8 or 8. The average for an optimal lag was 2.91 samples ($SD = 7,4$), which corresponds to a delay of about 0.18 seconds ($SD = 0.46$). For a histogram of optimal lags across participants, see Figure 4. For a visual representation of the cross-correlation in Figure 5, the raw data from Participant 7 with the highest correlation ($r = 0.80$) is plotted. It is visible that while the levels are lower, there is some agreement in the temporal patterns and shape of the signal. Figure 6 shows participant 20 with the lowest correlation ($r = -0.15$).

Figure 3

Histogram of maximum cross-correlations

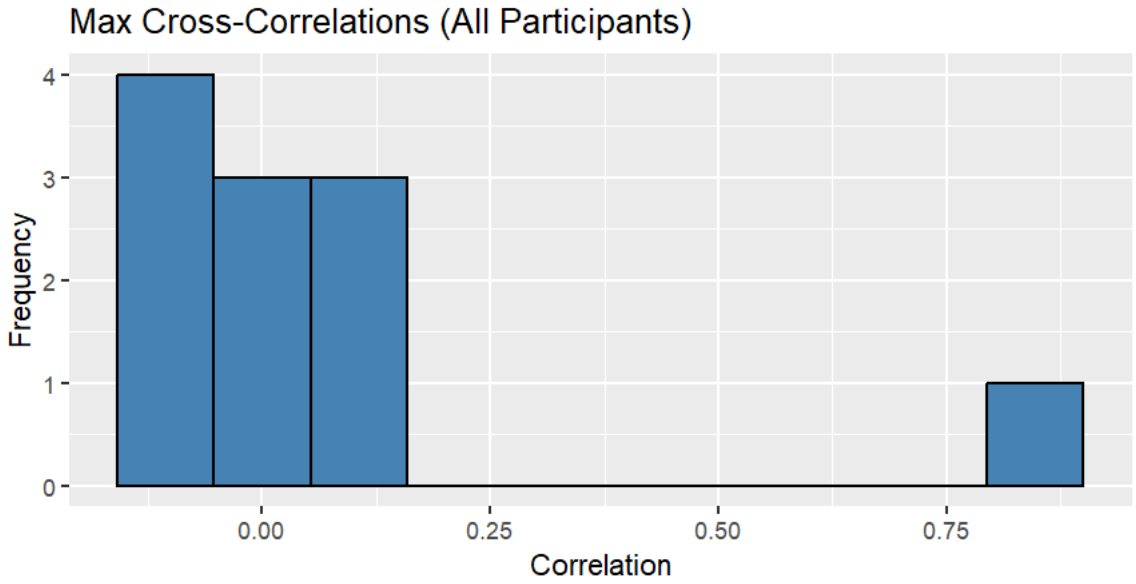


Figure 4

Histogram of Optimal Lags (in samples)

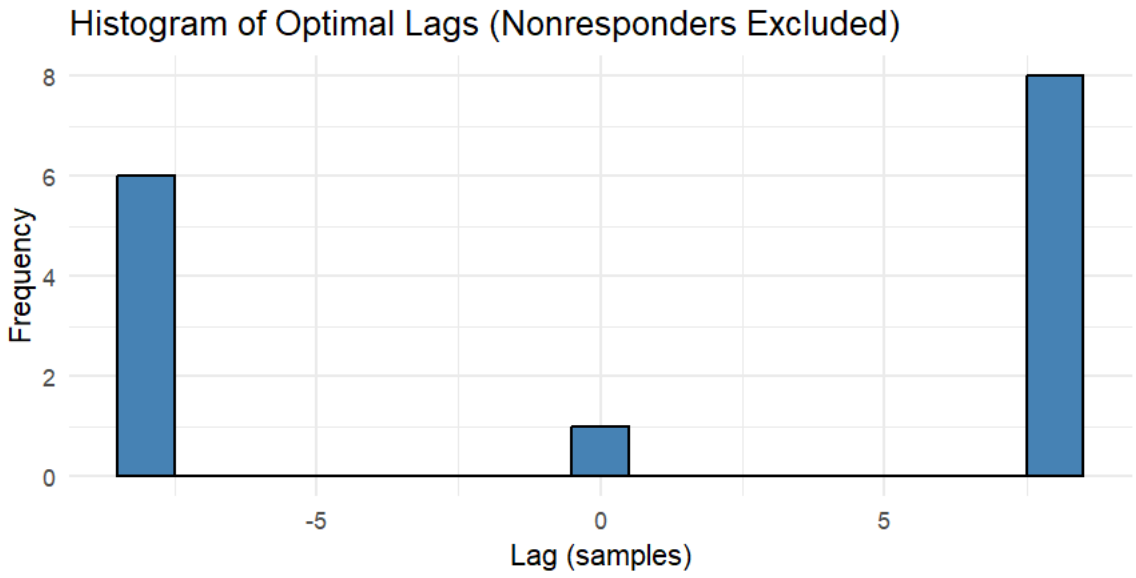


Figure 5

Cross-correlation plot based on the raw signals of the participant with the highest correlation (Participant 7, r=0.80)

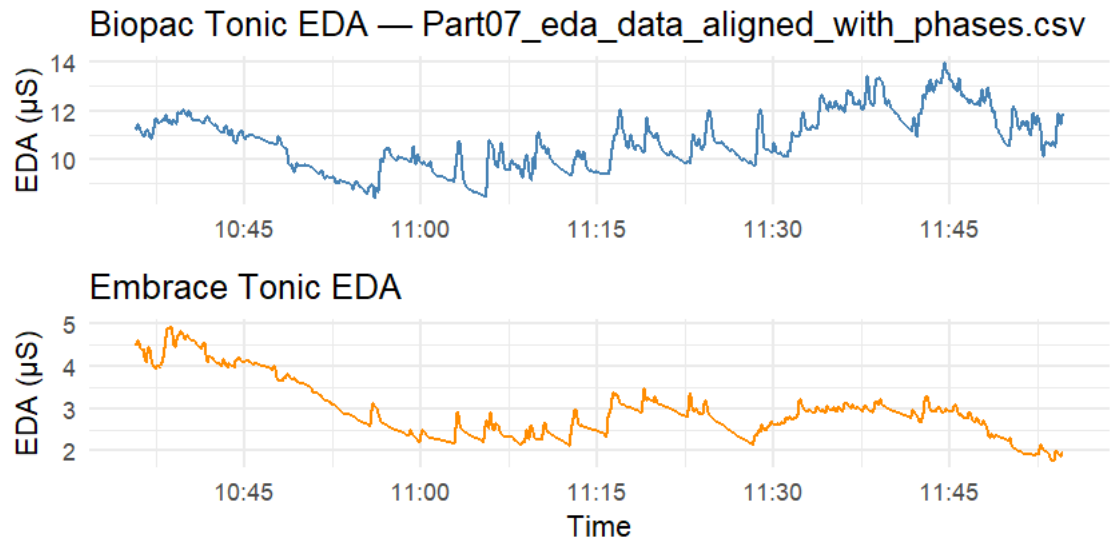
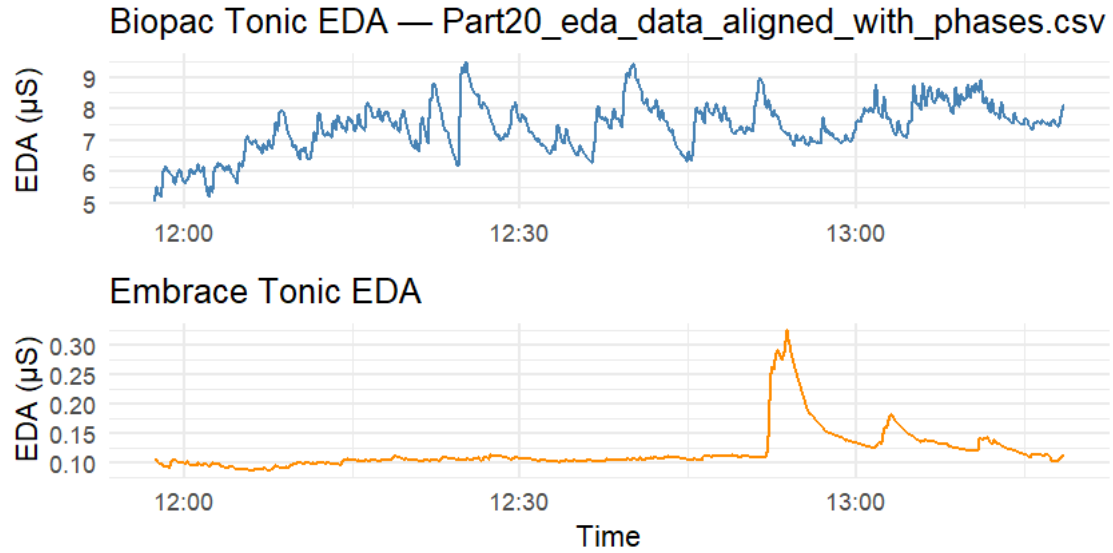


Figure 6

Cross-correlation plot based on the raw signals of the participant with the lowest correlation (Participant 20, r=-0.15)



Parameter level

For the Parameter level, log Bland-Altman plots were computed for the mean SCL (Figure 7), the number of SCRs (Figure 8) and the total SCR amplitude (Figure 9). In order to

only use data in which the participant was exposed to stress-related scenarios, the data was only considered from the beginning of the baseline reading until the end of the Nature task, excluding the preparation and ending. The plot for the mean SCL shows a mean difference of around -1.7, indicating that the Embrace Plus watch underestimates the SCL levels (see Figure 7). The 95% limits of agreement (LoA) ranged from -3.04 to -0.38. This exceeds the predefined acceptable range of ±0.96. Only 18.2% (n=2) of participants fell within this range, suggesting a poor agreement. In general, all participants lay below zero, which indicates an overall underestimation of SCL values of the Embrace Plus compared to BIOPAC. When looking at the number of SCRs, the Embrace watch detects fewer SCRs than the BIOPAC (see Figure 8). The mean difference is around -2.3 with a 95% LoA from +1.71 to -6.32. Therefore, the Embrace Plus underestimates the number of SCRs when compared to BIOPAC. However, about 50% (*n=6*) of the participants lay within the predefined limits of agreement, indicating moderate agreement. Lastly, the total SCR amplitude shows a consistent underestimation by the Embrace Plus watch with a bias of around -2.7, with all data points lying below 0 (see Figure 9). This indicates that the Embrace Plus reports much smaller total amplitudes than BIOPAC. For an overview of the raw mean and standard deviation for each parameter per device, see Table 1.

Figure 7

Bland-Altman plot for the log difference of the mean SCL between BIOPAC and the Embrace Plus

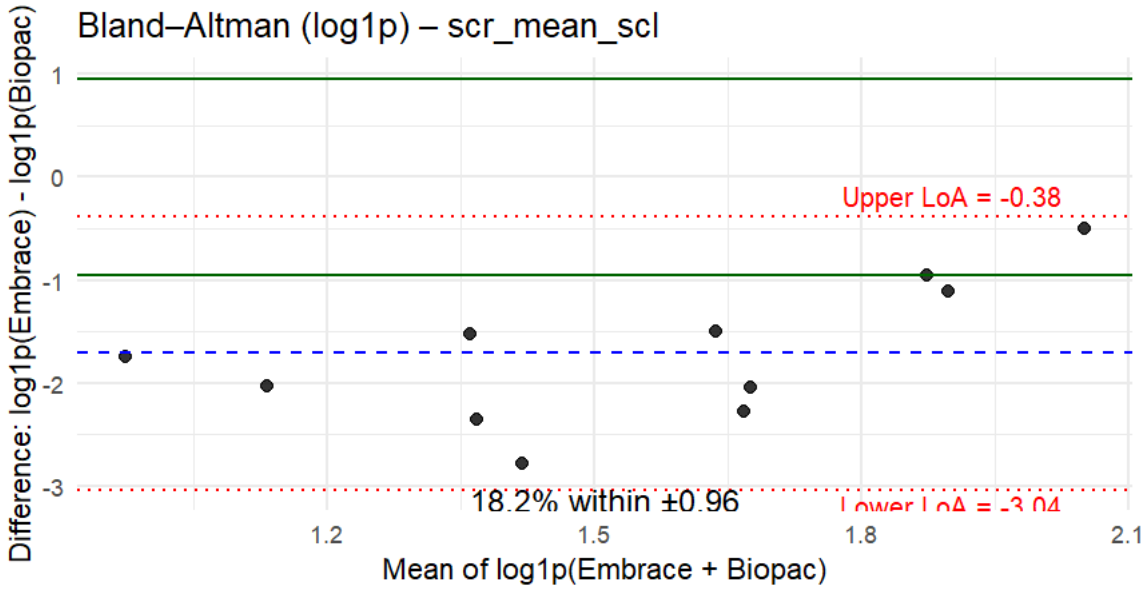


Figure 8

Bland-Altman plot for the log difference of the number of SCRs between BIOPAC and the Embrace Plus

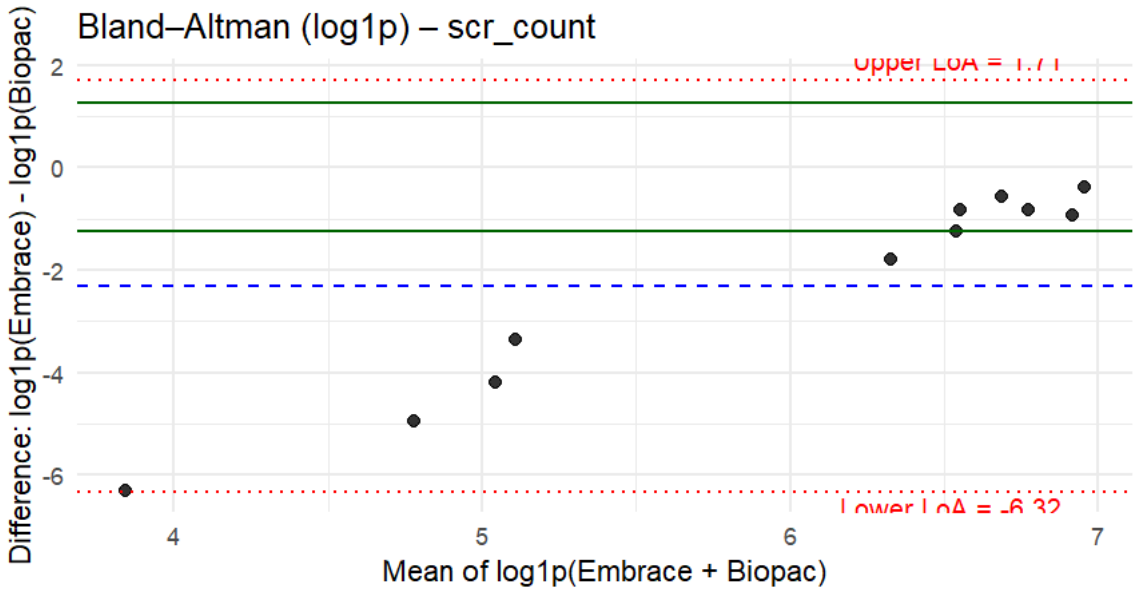


Figure 9

Bland-Altman plot for the log difference of the total SCR amplitude between BIOPAC and the Embrace Plus

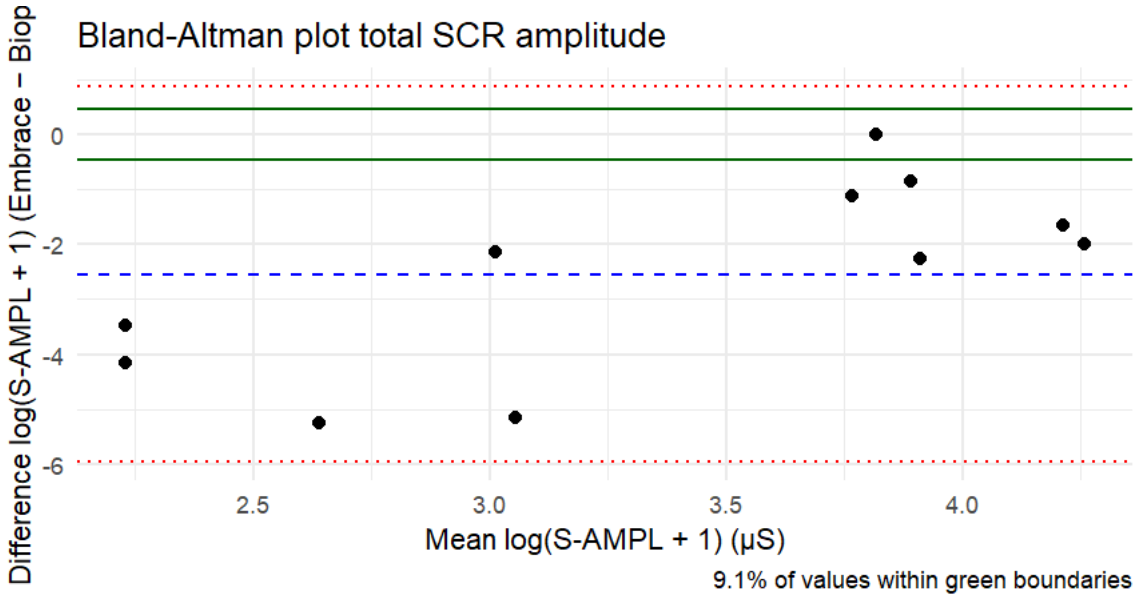


Table 1

Mean and SD values of raw signals for all participants from BIOPAC and Embrace Plus between the start of Baselines until the end of the Nature task

Parameter	Mean BIOPAC	SD BIOPAC	Mean Embrace Plus	SD Embrace Pus
Mean SCL (μS)	10.5	3.37	1.39	1.61
Number of SCRs	1238	201	345	306
Total SCR amplitude (μS)	159	69.3	21.6	21.7

Event level

When looking at the event level, the average SCL recorded by the BIOPAC during the TSST task show that there is a slight incline starting from the anticipation phase with the peak during the mental arithmetic phase of the task (Figure 10). However, this incline is relatively small, and the confidence intervals overlap (Table 2). The Figure also shows a drop for the TSST recovery phase. For a summary of the mean, SD and Confidence Interval for both devices, see Table 2. The same measures from the Embrace Plus watch show a larger peak during the mental arithmetic task with similarly high levels for the recovery phase, indicating that there might be a time delay in measurements (Figure 11). The Figure shows clearly that there are multiple participants showing very low mean SCL when measuring with the watch. The data shows very large standard deviations when compared to the mean and large differences between participants (Table 2). For the speech phase of the study, the Embrace data even has a negative lower CI, which is very unusual. According to Empatica, these negative values are possible to be seen as very low physiological skin conductance signals and can occur rarely in the range [-0.5μS,0μS] (Empatica, n.d.).

Figure 10

TSST: Biopac Average SCL per Phase

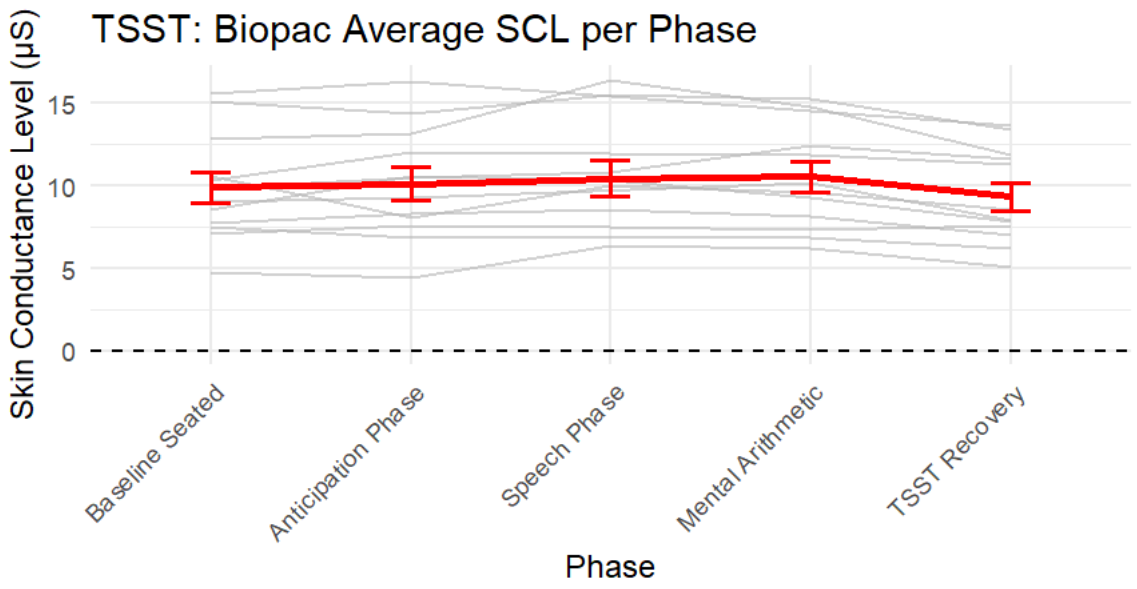


Figure 11

TSST: Embrace Plus Average SCL per Phase

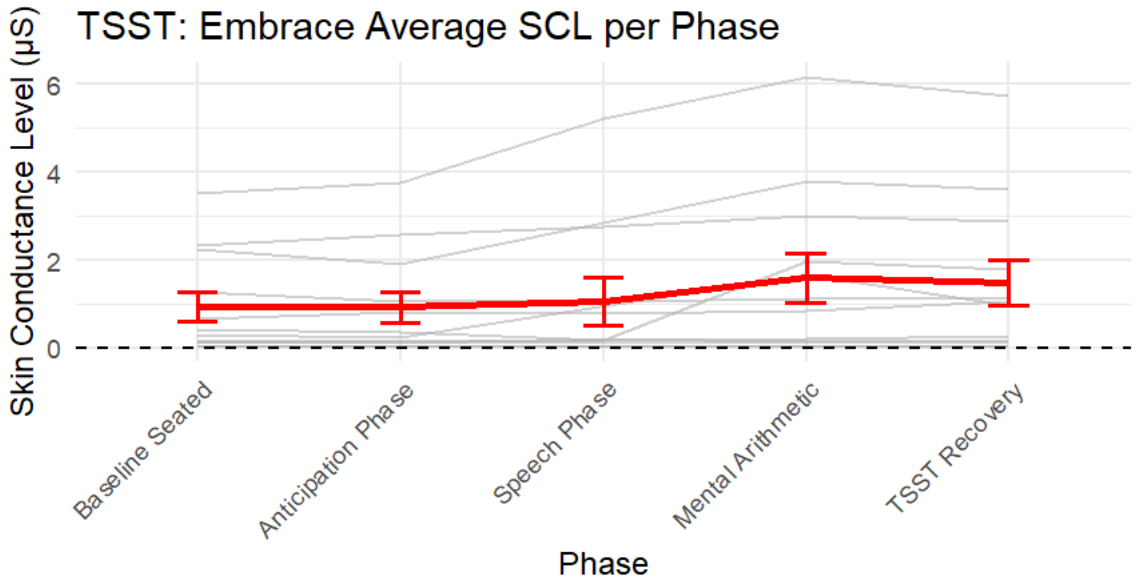


Table 2

TSST summary of Mean, SD and Confidence Intervals (CI) from SCL per Phase for BIOPAC and Embrace Plus

	BIOPAC			Embrace Plus		
Phase	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**
Baseline seated	9.91	3.24	7.85/12.0	0.934	1.16	0.197/1.67
Anticipation phase TSST	10.1	3.4	7.97/12.3	0.929	1.21	0.161/1.7
Speech phase	10.4	3.42	7.99/12.9	1.06	1.69	-0.151/2.27
Mental Arithmetic	10.6	3.2	7.99/12.6	1.59	1.91	0.381/2.8
TSST Recovery	9.33	2.9	7.49/11.2	1.48	1.77	0.353/2.6

For the Plank task, the BIOPAC data shows an increase starting from the anticipation phase, with the highest average at the stage when walking the plank (Figure 12). The mean SCL levels decrease from the plank baseline until the recovery phase. However, the confidence intervals overlap largely, indicating that this increase is not true for every participant (Table 3). The Embrace Plus plot shows an increase in SCL levels for the anticipation phase, which afterwards decreases slightly (Figure 13). However, again, the CI overlaps through all phases, which shows that there is not necessarily a large increase for all participants (Table 3). In general, the plots for the Embrace Plus do not show a comparable large increase in stress response to the BIOPAC plots. For a summary of the Mean, SD and Confidence Interval per device, see Table 3.

Figure 12

Plank: Biopac Average SCL per Phase

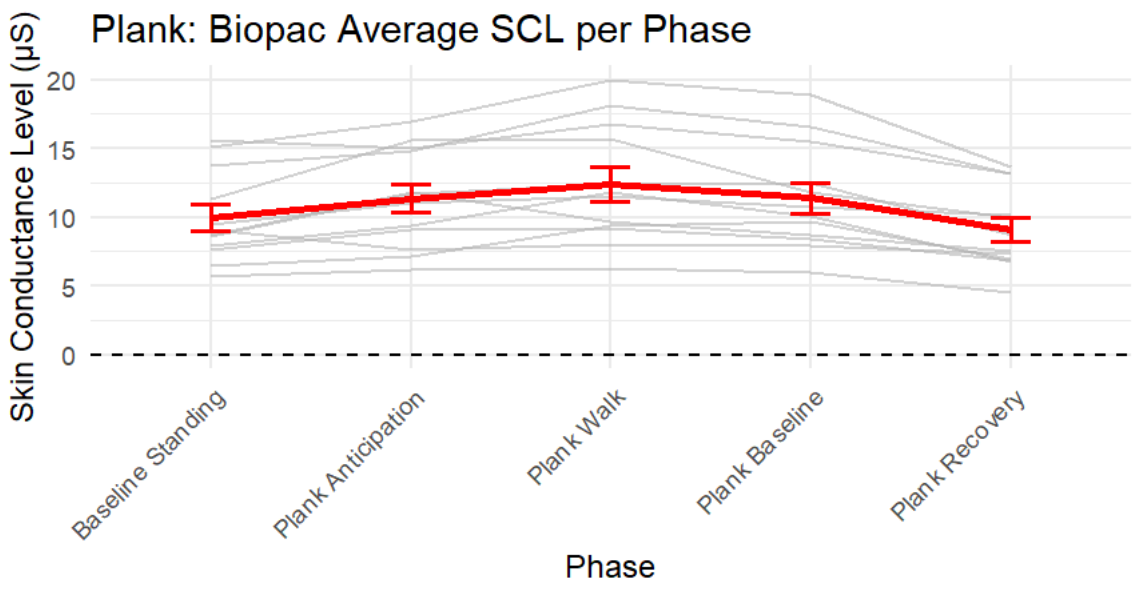


Figure 13

Plank: Embrace Plus Average SCL per Phase

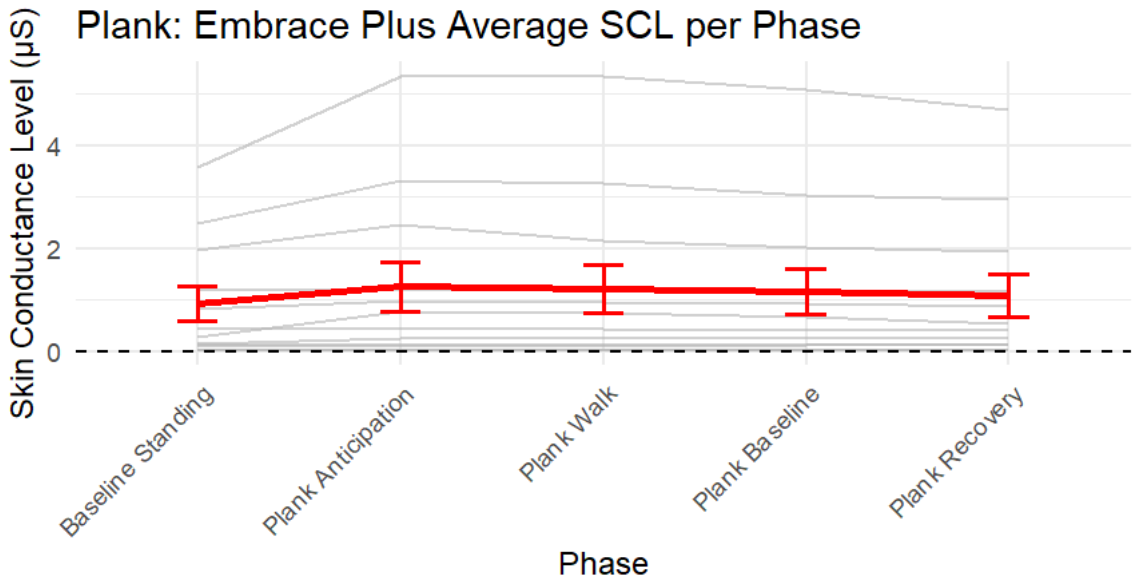


Table 3

Plank task summary of Mean, SD and Confidence Intervals (CI) from SCL per Phase for BIOPAC and Embrace Plus

	BIOPAC			Embrace Plus		
Phase	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**
Baseline Standing	9.95	3.28	7.86/12.0	0.931	1.16	0.194/1.67
Plank Anticipation	11.4	3.61	9.07/13.7	1.25	1.66	0.2/2.3
Plank Walk	12.4	4.31	9.66/15.1	1.22	1.63	0.181/2.25
Plank Baseline	11.4	3.85	8.95/13.8	1.15	1.54	0.176/2.13
Plank Recovery	9.07	2.98	7.17/11.0	1.09	1.44	0.174/2.01

When looking at the difference between the two devices, it is visible that the BIOPAC records consistently higher SCL values than the Embrace Plus watch (Figures 14, 15). The difference between the devices is higher than the considered acceptable range of ± 50%, following van Lier et al. (2019). This indicates that the Embrace watch does not accurately measure SCL data. The red line indicates the average between participants and shows that there is no large change in the level of disagreement across phases.

Figure 14

TSST: Log Ratio Difference in SCL between BIOPAC and Embrace

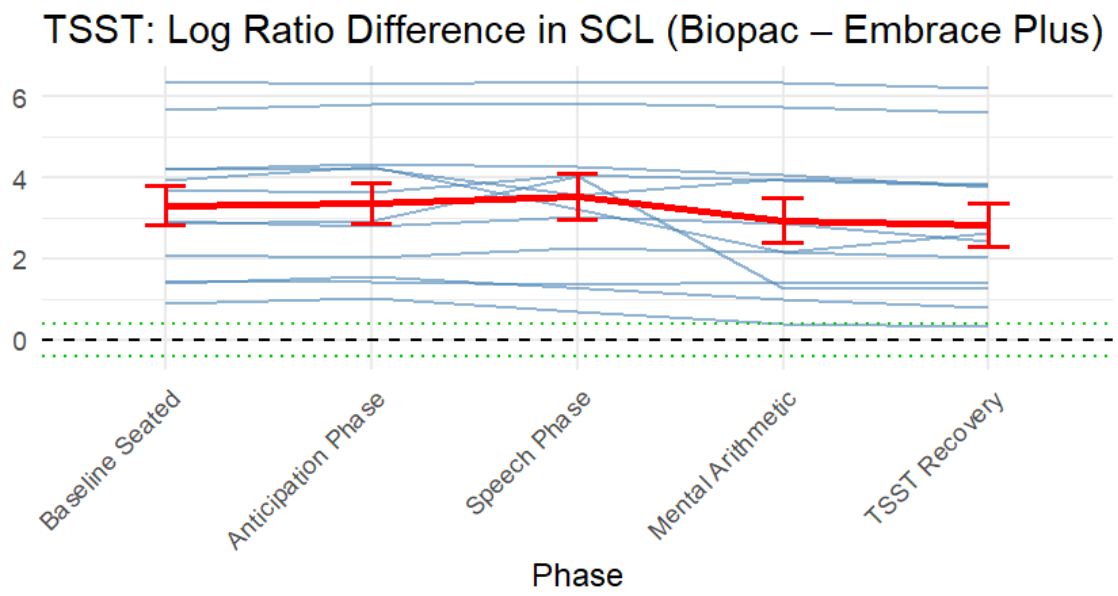
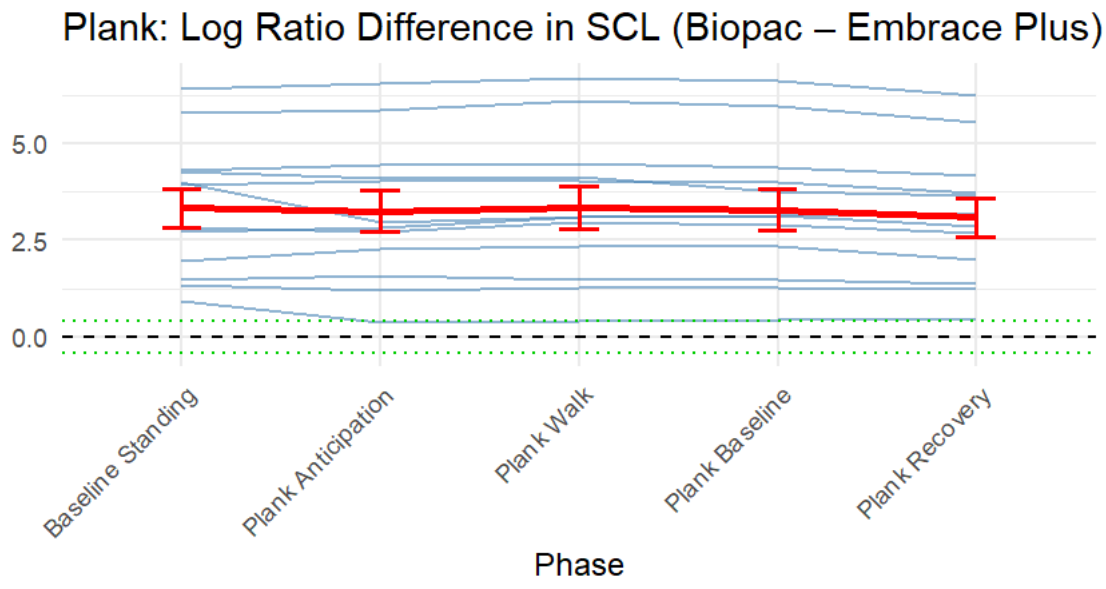


Figure 15

Plank task: Log Ratio Difference in SCL between BIOPAC and Embrace Plus



Discussion

This study aimed to evaluate the accuracy of the Embrace Plus watch in measuring EDA compared to a gold-standard device across signal, parameter and event levels, using VR-based stress-inducing scenarios. The results show that while the watch captures some trends in physiological changes, it consistently underestimates EDA values across all levels. This

underestimation was found in low-cross correlation with the gold-standard device, the lower mean SCL, and lower counts and amplitudes of SCRs. Previous studies found similar underestimations, and it was commonly attributed to factors like a slower sweat gland density and greater thermoregulatory sweating than emotional sweating on the wrist, and possible movement-related artefacts (Van Lier et al., 2019; De Geus & Gevonden, 2024; Menghini et al., 2019).

The signal-level analysis revealed a poor agreement in the overall shape and timing of EDA signals between the two devices. The maximum cross-correlation was very low. Seven participants even had a negative cross-correlation. Only one of the participants showed a high cross-correlation, indicating that signal synchrony is not achieved. These findings align with previous findings about the limited temporal resolution and sensitivity of wrist-worn sensors (Van der Mee et al., 2021). This may be due to wrong sensor placement, a limited sensitivity of the Embrace Plus or the lack of inducing severe stressors from the experiment (Klimek et al., 2023; Van der Mee et al., 2021). Additionally, most optimal lags were far from zero, suggesting a temporal misalignment or phase shift. This means that the Embrace Plus signal reached its peaks or changes in EDA several seconds before or after the BIOPAC. Van Lier et al. (2019) found similar trends across participants showing both positive and negative optimal lags. This may be due to the large difference in frequency of the two devices when measuring EDA (Menghini et al., 2019). The lower sampling rate of the Embrace Plus may have caused this misalignment.

When looking at the Parameter level, the Embrace Plus significantly underestimated the mean skin conductance level (SCL) compared to the BIOPAC. The number of skin conductance responses (SCRs) was also lower on the Embrace Plus than BIOPAC; however, the underestimation was less severe, indicating moderate agreement. Lastly, the total SCR amplitude was again underestimated by the Embrace Plus. These patterns mostly align with previous findings that wearable EDA sensors tend to detect lower mean SCL and number of SCRs (Van Lier et al., 2019; Van der Mee et al., 2021). These findings, however, do not align with the hypothesis posed at the beginning of this study of the total amplitude being the most accurate measurement of the Embrace Plus (Van Lier et al., 2019). Research suggests that this underestimation may be caused by emotional sweating being less pronounced on the wrist (De Geus & Gevonden, 2024).

At the event level, both the TSST and the Plank task were not able to induce significant increases in physiological responses. There seemed to be a slight increase in SCL

during stress phases and decreases during recovery; however, due to the error bars overlapping, this finding is not significant. The Embrace Plus was able to reflect some of these trends, while it still reported lower values, especially during peak stress phases. The log ratio difference plots confirmed a systematic underestimation across all phases, and the differences often exceeded the acceptable range. These results suggest that there is a need for more acute stress-inducing scenarios in order to test the expected accuracy under more acute stressors (Van Lier et al., 2019).

Some additional factors may account for the large differences in measured data between the two devices. Another reason could be that the filtering of the signal may have attenuated peaks and reduced response counts. Lastly, the data between the two watches may be completely misaligned, leading to a mismatch between the results of both devices. One possibility of checking for a misalignment would be to realign based on the data from the accelerometer of both devices.

Strengths of the study

One of the key strengths of this study is the exposure to multiple different conditions, including not just stress and relaxation scenarios, but multiple baselines and especially the inclusion of movement and speaking in the different parts of this study. In general, the relatively long duration of this study ensured a variety of data for each participant that can give a clear picture of different levels of stress. Additionally, the use of VR ensured an immersive and stress-inducing environment, ensuring consistent and controllable exposure across participants.

Limitations

Despite the strengths of this study, it also faced some limitations. For one, the sample size is relatively small, especially after having to exclude multiple participants due to technical difficulties or not getting accurate signals, especially from the Embrace Plus (*N=11*). When compared to the advised sample size of 55 by Van Lier et al. (2019), the small sample size can limit generalizability. Additionally, the large exclusion of ‘nonresponders’ may introduce certain biases and an underestimation of the variability. Lastly, the scenarios may not have induced strong stress responses in the participants.

Future Research

While the Embrace Plus watch does show potential for detecting broad trends in physiological responses related to stress, it lacks the accuracy and sensitivity required for

clinical or detailed psychological research. It might therefore be suitable for within-person general stress tracking, but not for measuring absolute values. For future research, it would be good to have a larger study with a larger sample size. It might also be helpful to check for reliability when performing the same tasks multiple times per participant.

Conclusion

In conclusion, the Embrace Plus does not accurately measure the physiological stress response EDA. While it might pick up trends for some individuals, it is not accurate in giving exact measurements. Therefore, future research should continue to validate wearables to receive a general level of accuracy that these devices need to achieve to ensure sufficient measurements of EDA to use in inventions and clinical settings or at home. If the devices achieve this accuracy, they may help to improve stress management skills for individuals in a cost-efficient and accessible way.

References

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress*, *6*, 113-126. https://doi.org/10.1016/j.ynstr.2016.11.001

BIOPAC Systems, Inc. (1989). MP160 System with AcqKnowledge. [Apparatus and Software]. https://www.biopac.com/product/mp160-system-with-acqknowledge/

BIOPAC Systems, Inc. (2010). BioNomadix series. [Apparatus]. https://www.biopac.com/wp-content/uploads/BioNomadix-Series.pdf

Boucsein, W. (2012). Electrodermal activity (2nd ed.). New York, NY: Springer. https://doi.org/10.1007/978-1-4614-1126-0

De Geus, E. J. C. & Gevonden, M. J. (2024). Acquisition and Analysis of Ambulatory Autonomic Nervous System Data. In Mehl, M. R., Eid, M., Wrzus C., Harari, G. M. & Ebner-Priemer, U. W. (Eds.). Mobile Sensing in Psychology : Methods and Applications (p. 129-167). *The Guilford Press.*

De Witte, N.A.J., Buyck, I. & Van Daele, T. (2019). Combining Biofeedback with Stress Management Interventions: A Systematic Review of Physiological and Psychological Effects. *Appl Psychophysiology and Biofeedback* **44**, 71–82. https://doi.org/10.1007/s10484-018-09427-7

Dillon, A., Kelly, M., Robertson, I. H., & Robertson, D. A. (2016). Smartphone Applications Utilizing Biofeedback Can Aid Stress Reduction. *Frontiers in Psychology*, *7*, 832. https://doi.org/10.3389/fpsyg.2016.00832

Empatica. (2019). EmbracePlus. [Apparatus]. https://www.empatica.com/en-eu/embraceplus/

Empatica. (n.d.). *Contents of downloaded data files*. Empatica Support. https://support.empatica.com/hc/en-us/articles/360028646091-Contents-of-downloaded-data-files

Finseth, T., Barnett, N., Shirtcliff, E. A., Dorneich, M. C., & Keren, N. (2018). Stress Inducing Demands in Virtual Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. https://doi.org/10.1177/1541931218621466

Halbig, A., & Latoschik, M. E. (2021). A Systematic Review of Physiological Measurements, Factors, Methods, and Applications in Virtual Reality. *Frontiers in Virtual Reality*, *2*, 694567. https://doi.org/10.3389/frvir.2021.694567

Hanshans, C., Amler, T., Zauner, J., & Bröll, L. (2024). Inducing and measuring acute stress in virtual reality: Evaluation of canonical physiological stress markers and measuring methods. *Journal of Environmental Psychology*, *94*, 102107. https://doi.org/10.1016/j.jenvp.2023.102107

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, *28*(1-2), 76–81. https://doi.org/10.1159/000119004

Klimek, A., Mannheim, I., Schouten, G., Wouters, E. J. M., & Peeters, M. W. H. (2023). Wearables measuring electrodermal activity to assess perceived stress in care: a scoping review. *Acta neuropsychiatrica*, 1–11. Advance online publication. https://doi.org/10.1017/neu.2023.19

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Mayo Clinic. (n.d.). *Stress: Signs, symptoms, causes, and stress management*. https://www.mayoclinic.org/healthy-lifestyle/stress-management/in-depth/stress/art-20046037

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). https://doi.org/10.1111/psyp.13441

Meta. (2023). Meta Quest 3. [Apparatus and Software]. https://www.meta.com/nl/en/quest/quest-3/

Piao, X., Xie, J., & Managi, S. (2024). Continuous worsening of population emotional stress globally: universality and variations. *BMC Public Health*, *24*(1). https://doi.org/10.1186/s12889-024-20961-4

Roos, L. G., & Slavich, G. M. (2023). Wearable technologies for health research: Opportunities, limitations, and practical and conceptual considerations. *Brain, Behavior, and Immunity*, *113*, 444-452. https://doi.org/10.1016/j.bbi.2023.08.008

Shiban, Y., Diemer, J., Brandl, S., Zack, R., Mühlberger, A., & Wüst, S. (2016). Trier Social Stress Test in vivo and in virtual reality: Dissociation of response domains. *International Journal of Psychophysiology*, *110*, 47-55. https://doi.org/10.1016/j.ijpsycho.2016.10.008

Shirtcliff, E. A., Finseth, T. T., Winer, E. H., Glahn, D. C., Conrady, R. A., & Drury, S. S. (2024). Virtual stressors with real impact: What virtual reality-based biobehavioral research can teach us about typical and atypical stress responsivity. *Translational Psychiatry*, *14*, 441. https://doi.org/10.1038/s41398-024-03129-x

Texas Instruments [TI]. (2023). Electrodermal activity measurement: Principles and applications (Application Brief No. SBAA556). *Texas Instruments.* https://www.ti.com/lit/ab/sbaa556/sbaa556.pdf

Van Dammen, L., Finseth, T. T., McCurdy, B. H., Barnett, N. P., Conrady, R. A., Leach, A. G., Deick, A. F., Van Steenis, A. L., Gardner, R., Smith, B. L., Kay, A., & 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. D. M. & 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 Lier, H. G., Pieterse, M. E., Garde, A., Postel, M. G., de Haan, H. A., Vollenbroek-Hutten, M. M. R., Schraagen, J. M., & Noordzij, M. L. (2019). A standardized validity assessment protocol for physiological signals from wearable technology: Methodological underpinnings and an application to the E4 biosensor. *Behavior Research Methods*, *52*(2), 607-629. https://doi.org/10.3758/s13428-019-01263-9

World Health Organization [WHO]. (n.d.). *Stress*. https://www.who.int/news-room/questions-and-answers/item/stress

Appendix

Appendix 1

AI statement

During the preparation of this work, I used ChatGPT to improve and understand code for the data analysis. After using this tool/service, we thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome.

Appendix 2

R code

```
#load packages

library(dplyr)

library(ggplot2)

library(signal)

library(readr)

library(tidyr)

library(ggpubr)

library(tidyverse)

library(purrr)

library(vroom)


#open data

setwd("C:\\Users\\there\\Documents\\Data Bachelor Thesis")


# Resamples a signal to 16 Hz using linear interpolation

resample_signal <- function(signal, timestamps, target_hz = 16) {

  time_seq <- seq(from = min(timestamps), to = max(timestamps), by = 1 / target_hz)

  interpolated <- approx(timestamps, signal, xout = time_seq, method = "linear", rule = 2)$y

  return(tibble(time = time_seq, signal = interpolated))

}


# Detrend a signal to remove slow drifts and focus on shape
```

```
# Important for correlation analysis
detrend_signal <- function(x) {
  return(resid(lm(x ~ seq_along(x))))
}

# Extract Skin Conductance Responses from the phasic signal
# Uses a threshold of 0.01 µS
extract_scrs <- function(signal, threshold = 0.01) {
  peaks <- which(diff(sign(diff(signal))) == -2) + 1
  valid_peaks <- signal[peaks] > threshold
  amplitudes <- signal[peaks[valid_peaks]]
  return(list(count = sum(valid_peaks), amp = mean(amplitudes), total = sum(amplitudes)))
}

# --- Main Analysis Loop ---

results <- list()
corr_values <- c()
lag_values <- c()
signal_results <- list()


# Load all CSV files in the working directory
files <- list.files(pattern = "\\.csv$")


# Define plot function ONCE outside loop
plot_raw_tonic_eda <- function(file) {
  df <- read_csv(file, show_col_types = FALSE)
  df$timestamps <- as.POSIXct(df$timestamps)
```

```
  ggplot(df, aes(x = timestamps)) +
    geom_line(aes(y = EDA_biopac_tonic, color = "Biopac")) +
    geom_line(aes(y = EDA_embrace_tonic, color = "Embrace")) +
    labs(
      title = paste("Tonic EDA: Raw Signals for", file),
      x = "Time", y = "EDA (µS)", color = "Device"
    ) +
    theme_minimal()
}

corr_filenames <- c()

# MAIN LOOP
for (f in files) {
  df <- read_csv(f, show_col_types = FALSE)
  df$timestamps <- as.POSIXct(df$timestamps)
  ts <- as.numeric(difftime(df$timestamps, min(df$timestamps), units = "secs"))

  print(paste("Processed:", f))

  # Skip files with insufficient or problematic data
  valid_biopac <- sum(!is.na(df$EDA_biopac_tonic)) >= 2
  valid_embrace <- sum(!is.na(df$EDA_embrace_tonic)) >= 2
  valid_time <- sum(!is.na(ts)) >= 2 && length(unique(ts)) >= 2

  if (!valid_biopac | !valid_embrace | !valid_time) {
    warning(paste("Skipped", f, "- invalid data for interpolation"))
    next
  }
```

```
# --- SCR Extraction ---
scr_embrace <- extract_scrs(df$EDA_embrace_phasic)
scr_biopac <- extract_scrs(df$EDA_biopac_phasic)

# --- Resample and Align ---
biopac <- resample_signal(df$EDA_biopac_tonic, ts)
embrace <- resample_signal(df$EDA_embrace_tonic, ts)

common_time <- intersect(biopac$time, embrace$time)
biopac_interp <- dplyr::filter(biopac, time %in% common_time)
embrace_interp <- dplyr::filter(embrace, time %in% common_time)

# --- Normalize and Detrend ---
b_sig <- detrend_signal(scale(biopac_interp$signal)[,1])
e_sig <- detrend_signal(scale(embrace_interp$signal)[,1])

# --- Cross-Correlation ---
lags <- -8:8
cors <- sapply(lags, function(lag) {
  if (lag < 0) {
    cor(e_sig[1:(length(e_sig)+lag)], b_sig[(1-lag):length(b_sig)])
  } else {
    cor(e_sig[(1+lag):length(e_sig)], b_sig[1:(length(b_sig)-lag)])
  }
})

mean_corr <- mean(cors, na.rm = TRUE)
max_corr <- max(cors, na.rm = TRUE)
```

```
  best_lag <- lags[which.max(cors)]

  print(paste("Best lag:", best_lag, "Max corr:", max_corr))

  # Collect data
  corr_values <- c(corr_values, mean_corr)
  corr_filenames <- c(corr_filenames, f)
  lag_values <- c(lag_values, best_lag)

  signal_results[[f]] <- tibble(
    file = f,
    max_cross_correlation = max_corr,
    optimal_lag_samples = best_lag,
    optimal_lag_seconds = best_lag / 16
  )

  results[[f]] <- tibble(
    file = f,
    scr_count_embrace = scr_embrace$count,
    scr_count_biopac = scr_biopac$count,
    scr_amp_embrace = scr_embrace$amp,
    scr_amp_biopac = scr_biopac$amp,
    scr_total_embrace = scr_embrace$total,
    scr_total_biopac = scr_biopac$total
  )
}
  # --- Store Results ---
  results[[f]] <- tibble(
    file = f,
    scr_count_embrace = scr_embrace$count,
```

```
    scr_count_biopac = scr_biopac$count,
    scr_amp_embrace = scr_embrace$amp,
    scr_amp_biopac = scr_biopac$amp,
    scr_total_embrace = scr_embrace$total,
    scr_total_biopac = scr_biopac$total
  )


# Combine results
signal_summary_df <- bind_rows(signal_results) %>%
  arrange(desc(max_cross_correlation))

print(signal_summary_df)
write_csv(signal_summary_df, "max_cross_correlation_per_participant.csv")

# Summary stats
cat("Number of participants with lag values:", length(lag_values), "\n")
cat("Mean lag:", mean(lag_values), "\n")
cat("SD lag:", sd(lag_values), "\n")
cat("Mean lag (sec):", mean(lag_values)/16, "\n")
cat("SD lag (sec):", sd(lag_values)/16, "\n")

# Histogram
ggplot(data.frame(lag = lag_values), aes(x = lag)) +
  geom_histogram(binwidth = 1, fill = "steelblue", color = "black") +
  labs(
    title = "Histogram of Optimal Lags (All Participants)",
    x = "Lag (samples)",
    y = "Frequency"
  ) +
```

```
  theme_minimal()

# Combine and filter results
df_results <- bind_rows(results)
df_results$nonresponder <- df_results$scr_count_embrace < 1 |
is.na(df_results$scr_amp_embrace)
filtered_results <- dplyr::filter(df_results, !nonresponder)
filtered_corr_values <- corr_values[!df_results$nonresponder]



mean(filtered_corr_values)
sd(filtered_corr_values)

# Find best/worst files using the correct matched filenames
best_idx <- which.max(corr_values)
worst_idx <- which.min(corr_values)

best_file <- corr_filenames[best_idx]
worst_file <- corr_filenames[worst_idx]

cat("Highest mean cross-correlation:", corr_values[best_idx], "from", best_file, "\n")
cat("Lowest mean cross-correlation:", corr_values[worst_idx], "from", worst_file, "\n")

#--------different axis-----
library(patchwork)

plot_raw_tonic_eda <- function(file) {
  df <- read_csv(file, show_col_types = FALSE)
  df$timestamps <- as.POSIXct(df$timestamps)
```

```
  p1 <- ggplot(df, aes(x = timestamps, y = EDA_biopac_tonic)) +
    geom_line(color = "steelblue") +
    labs(title = paste("Biopac Tonic EDA —", file), x = NULL, y = "EDA (µS)") +
    theme_minimal()

  p2 <- ggplot(df, aes(x = timestamps, y = EDA_embrace_tonic)) +
    geom_line(color = "darkorange") +
    labs(title = "Embrace Tonic EDA", x = "Time", y = "EDA (µS)") +
    theme_minimal()

  # Combine vertically
  p1 / p2
}
# Plot best
plot_raw_tonic_eda(best_file)

# Plot worst
plot_raw_tonic_eda(worst_file)

#-------plot all raw data-----
# Loop over only the files that were actually used in analysis
for (f in corr_filenames) {
  plot <- plot_raw_tonic_eda(f)

  ggsave(
    filename = paste0(gsub(".csv", "", f), "_stacked_tonic_plot.png"),
    plot = plot,
    width = 10,
    height = 6
  )
```

```
  cat("Saved stacked plot for", f, "\n")
}

# --- SUMMARY ANALYSIS ---

# Combine all participant results into one data frame
df_results <- bind_rows(results)

# Include everyone — no nonresponder filtering
filtered_results <- df_results
filtered_corr_values <- corr_values

# Plot a histogram of the maximum cross-correlation values across valid participants
# Interpretation: High values mean good signal-level agreement
ggplot(data.frame(correlation = filtered_corr_values), aes(x = correlation)) +
  geom_histogram(bins = 10, fill = "steelblue", color = "black") +
  labs(
    title = "Max Cross-Correlations (All Participants)",
    x = "Correlation",
    y = "Frequency"
  )
#--------plot all participants in one plot------
library(tidyverse)

# Store results per participant
biopac_phase_list <- list()
embrace_phase_list <- list()

for (f in corr_filenames) {
```

```
  df <- read_csv(f, show_col_types = FALSE)
  df$timestamps <- as.POSIXct(df$timestamps)
  df$time_relative <- as.numeric(difftime(df$timestamps, min(df$timestamps), units =
"secs"))

  participant_id <- gsub("_eda_data_aligned_with.*", "", f)

  # Keep only valid rows
  # Skip participant if Event_Phases is missing or only NA
  if (!("Event_Phases" %in% colnames(df)) || all(is.na(df$Event_Phases))) {
    warning(paste("Skipped", f, "- Event_Phases missing or all NA"))
    next
  }

  # Ensure Event_Phases is treated as character
  df$Event_Phases <- as.character(df$Event_Phases)

  # Mean EDA per phase
  biopac_means <- df %>%
    group_by(Event_Phases) %>%
    summarise(mean_eda = mean(EDA_biopac_tonic, na.rm = TRUE), .groups = "drop") %>%
    mutate(participant = participant_id)

  embrace_means <- df %>%
    group_by(Event_Phases) %>%
    summarise(mean_eda = mean(EDA_embrace_tonic, na.rm = TRUE), .groups = "drop")
%>%
    mutate(participant = participant_id)

  biopac_phase_list[[f]] <- biopac_means
```

```
  embrace_phase_list[[f]] <- embrace_means
}

# Combine across participants
biopac_df <- bind_rows(biopac_phase_list)
embrace_df <- bind_rows(embrace_phase_list)

# Biopac
ggplot(biopac_df, aes(x = Event_Phases, y = mean_eda, group = participant, color =
participant)) +
  geom_line(alpha = 0.8) +
  geom_point() +
  labs(
    title = "Biopac Tonic EDA by Event Phase",
    x = "Event Phase",
    y = "Mean Biopac EDA (µS)"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

# Embrace
ggplot(embrace_df, aes(x = Event_Phases, y = mean_eda, group = participant, color =
participant)) +
  geom_line(alpha = 0.8) +
  geom_point() +
  labs(
    title = "Embrace Tonic EDA by Event Phase",
    x = "Event Phase",
    y = "Mean Embrace EDA (µS)"
  ) +
```

```
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))




# Create Bland-Altman plots for each SCR parameter using filtered data with range
# Acceptable log-transformed boundaries
acceptable_bounds <- list(
  scr_mean_scl = log1p(1.6),   # ≈ 0.96
  scr_count    = log1p(2.5),   # ≈ 1.25
  scr_total    = log1p(0.6)    # ≈ 0.47
)

# Function to extract SCR features
extract_scrs <- function(signal, threshold = 0.01) {
  peaks <- which(diff(sign(diff(signal))) == -2) + 1
  valid_peaks <- signal[peaks] > threshold
  amplitudes <- signal[peaks[valid_peaks]]
  list(
    count = sum(valid_peaks),
    total = sum(amplitudes)
  )
}

# Loop through all participant files
files <- list.files(pattern = "eda_data_aligned.*\\.csv$")
results <- list()

for (f in files) {
```

```
df <- read_csv(f, show_col_types = FALSE)
if (!"event_marker" %in% names(df)) next

# Define the time window: Baseline Start to Nature Scene End
i_start <- which(grepl("Baseline Reading Start", df$event_marker, ignore.case = TRUE))[1]
i_end   <- which(grepl("Nature Task end", df$event_marker, ignore.case = TRUE))[1]

if (is.na(i_start) || is.na(i_end) || i_end <= i_start) {
  warning(paste("Skipping", f, "- invalid markers"))
  next
}

segment <- df[i_start:i_end, ]

# Calculate mean SCL from tonic signal
mean_scl_embrace <- mean(segment$EDA_embrace_tonic, na.rm = TRUE)
mean_scl_biopac  <- mean(segment$EDA_biopac_tonic, na.rm = TRUE)

# Extract SCR parameters
scr_e <- extract_scrs(segment$EDA_embrace_phasic)
scr_b <- extract_scrs(segment$EDA_biopac_phasic)

results[[f]] <- tibble(
  file = f,
  scr_mean_scl_embrace = mean_scl_embrace,
  scr_mean_scl_biopac  = mean_scl_biopac,
  scr_count_embrace    = scr_e$count,
  scr_count_biopac     = scr_b$count,
  scr_total_embrace    = scr_e$total,
  scr_total_biopac     = scr_b$total
```

```
  )
}

# Combine all results
df_results <- bind_rows(results) %>%
  dplyr::filter(!is.na(scr_count_embrace), !is.na(scr_count_biopac),
         !is.na(scr_total_embrace), !is.na(scr_total_biopac),
         !is.na(scr_mean_scl_embrace), !is.na(scr_mean_scl_biopac))

# Plot each parameter with Bland–Altman plot (log1p-transformed)
params <- c("scr_mean_scl", "scr_count", "scr_total")

for (param in params) {
  x <- log1p(df_results[[paste0(param, "_embrace")]])
  y <- log1p(df_results[[paste0(param, "_biopac")]])

  df_ba <- tibble(
    mean_value = (x + y) / 2,
    difference = x - y
  ) %>% drop_na()

  mean_diff <- mean(df_ba$difference)
  sd_diff   <- sd(df_ba$difference)
  upper     <- mean_diff + 1.96 * sd_diff
  lower     <- mean_diff - 1.96 * sd_diff
  acc_bound <- acceptable_bounds[[param]]
  pct_within <- mean(abs(df_ba$difference) <= acc_bound) * 100

  p <- ggplot(df_ba, aes(x = mean_value, y = difference)) +
    geom_point(size = 2, alpha = 0.8) +
```

```
  geom_hline(yintercept = mean_diff, linetype = "dashed", color = "blue") +
  geom_hline(yintercept = upper, linetype = "dotted", color = "red") +
  geom_hline(yintercept = lower, linetype = "dotted", color = "red") +
  geom_hline(yintercept = acc_bound, color = "darkgreen") +
  geom_hline(yintercept = -acc_bound, color = "darkgreen") +
  annotate("text", x = max(df_ba$mean_value), y = upper,
           label = sprintf("Upper LoA = %.2f", upper),
           hjust = 1.1, vjust = -0.5, size = 3.5, color = "red") +
  annotate("text", x = max(df_ba$mean_value), y = lower,
           label = sprintf("Lower LoA = %.2f", lower),
           hjust = 1.1, vjust = 1.2, size = 3.5, color = "red") +
  annotate("text", x = mean(range(df_ba$mean_value)), y = min(df_ba$difference),
           label = sprintf("%.1f%% within ±%.2f", pct_within, acc_bound),
           size = 4.2, color = "black", vjust = 2) +
  labs(
    title = paste("Bland–Altman (log1p) –", param),
    x = "Mean of log1p(Embrace + Biopac)",
    y = "Difference: log1p(Embrace) - log1p(Biopac)"
  ) +
  coord_cartesian(ylim = c(min(df_ba$difference, -acc_bound, lower),
                           max(df_ba$difference, acc_bound, upper))) +
  theme_minimal()


 print(p)
}
# Filter to time from "Baseline" through the full "Nature Scene"
# We'll assume data is ordered in time, so we take all rows between first "Baseline" and last
"Nature Scene"


# Parameters
```

```
scr_threshold <- 0.01  # Minimum SCR amplitude (µS)
sampling_rate <- 16     # Hz
data_path <- "C:/Users/there/Documents/Data Bachelor Thesis"

# Get list of CSV files
file_list <- list.files(path = data_path, pattern = "eda_data_aligned_.*\\.csv", full.names = TRUE)

# Peak detection function
get_scr_total_amplitude <- function(signal, threshold) {
  peaks <- findpeaks(signal, minpeakheight = threshold, minpeakdistance = sampling_rate)
  if (is.null(peaks)) return(0)
  sum(peaks[, 1])  # Sum of peak amplitudes
}

# Initialize results table
results <- data.frame(
  participant = character(),
  biopac_total_amplitude = numeric(),
  embrace_total_amplitude = numeric(),
  stringsAsFactors = FALSE
)

# Loop over all files
for (file in file_list) {
  # Load data
  eda_df <- tryCatch(read_csv(file), error = function(e) return(NULL))
  if (is.null(eda_df)) {
    warning(paste("Skipping file due to read error:", file))
    next
```

```
}

# Check for required event markers
start_idx <- which(eda_df$Event_Phases == "Baseline Reading")[1]
end_idx   <- tail(which(eda_df$Event_Phases == "Nature Task"), 1)

if (is.na(start_idx) || is.na(end_idx)) {
  warning(paste("Skipping file due to missing event markers:", file))
  next
}

# Extract signals
analysis_window <- start_idx:end_idx
biopac_phasic   <- eda_df$EDA_biopac_phasic[analysis_window]
embrace_phasic  <- eda_df$EDA_embrace_phasic[analysis_window]

# Compute amplitudes
biopac_amp  <- get_scr_total_amplitude(biopac_phasic, scr_threshold)
embrace_amp <- get_scr_total_amplitude(embrace_phasic, scr_threshold)

# Extract participant ID from filename
participant_id <- gsub(".*eda_data_aligned_(p[0-9]+)\\.csv", "\\1", file)

# Append results
results <- rbind(results, data.frame(
  participant = participant_id,
  biopac_total_amplitude = biopac_amp,
  embrace_total_amplitude = embrace_amp,
  stringsAsFactors = FALSE
))
```

```
}

# Show summary table
print(results)



#total SCR amplitude revised
#------Bland Altman Amplitude--------
library(tidyverse)
library(ggplot2)

# Set your data folder path
folder_path <- "C:/Users/there/Documents/Data Bachelor Thesis"
# Acceptable limit
acceptable_limit <- 0.5

# Load and combine all CSVs, harmonize timestamps
file_list <- list.files(path = folder_path, pattern = "\\.csv$", full.names = TRUE)

eda_data <- file_list %>%
  set_names() %>%
  map_dfr(~ read_csv(.x, col_types = cols(.default = col_guess(), timestamps =
col_character())),
          .id = "file") %>%
  mutate(participant_id = tools::file_path_sans_ext(basename(file)))



results_clean <- results %>%
  filter(embrace_total_amplitude > 0 & biopac_total_amplitude > 0)
# Add log(S-AMPL + 1) transformation
```

```
results_clean <- results %>%
  filter(biopac_total_amplitude > 0, embrace_total_amplitude > 0) %>%
  mutate(
    log_biopac   = log(biopac_total_amplitude + 1),
    log_embrace  = log(embrace_total_amplitude + 1),
    mean_log     = (log_biopac + log_embrace) / 2,
    diff_log     = log_embrace - log_biopac
  )

# Compute Bland–Altman stats
mean_diff <- mean(results_clean$diff_log)
sd_diff   <- sd(results_clean$diff_log)
loa_upper <- mean_diff + 1.96 * sd_diff
loa_lower <- mean_diff - 1.96 * sd_diff

# Define green boundaries based on van Lier
acc_upper <- 0.47
acc_lower <- -0.47

within_accept <- mean(results_clean$diff_log >= acc_lower & results_clean$diff_log <=
acc_upper) * 100

# Plot
ggplot(results_clean, aes(x = mean_log, y = diff_log)) +
  geom_point(size = 2) +
  geom_hline(yintercept = mean_diff, color = "blue", linetype = "dashed") +
  geom_hline(yintercept = c(loa_upper, loa_lower), color = "red", linetype = "dotted") +
  geom_hline(yintercept = c(acc_upper, acc_lower), color = "darkgreen") +
  labs(
    title = "Bland-Altman plot total SCR amplitude",
```

```
    x = "Mean log(S-AMPL + 1) (µS)",
    y = "Difference log(S-AMPL + 1) (Embrace − Biopac)",
    caption = paste0(round(within_accept, 1), "% of values within green boundaries")
  ) +
  theme_minimal()


#table for mean and sd before log
df_results <- bind_rows(results)

# Create summary table
summary_table <- tibble(
  Parameter = c("Mean SCL (µS)", "SCR count", "Total SCR amplitude (µS)"),
  `Embrace Mean` = c(
    mean(df_results$scr_mean_scl_embrace, na.rm = TRUE),
    mean(df_results$scr_count_embrace, na.rm = TRUE),
    mean(df_results$scr_total_embrace, na.rm = TRUE)
  ),
  `Embrace SD` = c(
    sd(df_results$scr_mean_scl_embrace, na.rm = TRUE),
    sd(df_results$scr_count_embrace, na.rm = TRUE),
    sd(df_results$scr_total_embrace, na.rm = TRUE)
  ),
  `Biopac Mean` = c(
    mean(df_results$scr_mean_scl_biopac, na.rm = TRUE),
    mean(df_results$scr_count_biopac, na.rm = TRUE),
    mean(df_results$scr_total_biopac, na.rm = TRUE)
  ),
  `Biopac SD` = c(
    sd(df_results$scr_mean_scl_biopac, na.rm = TRUE),
```

```
    sd(df_results$scr_count_biopac, na.rm = TRUE),
    sd(df_results$scr_total_biopac, na.rm = TRUE)
  )
)

# Print it
print(summary_table)



#------Event analysis----
# Set working directory
setwd("C:/Users/there/Documents/Data Bachelor Thesis")


# 2. Load all files and add participant ID
file_list <- list.files(pattern = "_with_phases\\.csv$")

all_data <- file_list %>%
  map_df(~ read_csv(.x, col_types = cols(timestamps = col_character())) %>%
           mutate(participant = str_remove(.x, "_with_phases\\.csv$")))

TSST <- c("Baseline Seated", "Anticipation Phase", "Speech Phase",
                "Mental Arithmetic", "TSST Recovery")

# Convert wide to long format
eda_long <- all_data %>%
  select(participant, Event_Phases, EDA_biopac_tonic, EDA_embrace_tonic) %>%
  pivot_longer(cols = c(EDA_biopac_tonic, EDA_embrace_tonic),
               names_to = "device",
```

```
            values_to = "tonic_eda") %>%
  mutate(device = case_when(
    device == "EDA_biopac_tonic" ~ "Biopac",
    device == "EDA_embrace_tonic" ~ "Embrace Plus"
  ))

eda_long <- eda_long %>%
  mutate(condition = case_when(
    Event_Phases %in% c("Baseline Seated", "Anticipation Phase", "Speech Phase",
                        "Mental Arithmetic", "TSST Recovery") ~ "TSST",
    Event_Phases %in% c("Baseline Standing", "Plank Anticipation", "Plank Walk",
                        "Plank Baseline", "Plank Recovery") ~ "Plank",
    TRUE ~ "Other"
  ))
# Filter to just TSST and calculate mean SCL per phase × person × device
eda_mean_scl <- eda_long %>%
  dplyr::filter(condition == "TSST") %>%
  group_by(participant, Event_Phases, device) %>%
  summarise(mean_scl = mean(tonic_eda, na.rm = TRUE), .groups = "drop") %>%
  mutate(Event_Phases = factor(Event_Phases, levels = c(
    "Baseline Seated", "Anticipation Phase", "Speech Phase",
    "Mental Arithmetic", "TSST Recovery"
  )))

eda_mean_scl %>%
  dplyr::filter(device == "Biopac") %>%
  ggplot(aes(x = Event_Phases, y = mean_scl, group = participant)) +
  geom_line(color = "gray70", alpha = 0.6, linewidth = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
```

```
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1), color = "red", width =
0.2, linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(title = "TSST: Biopac Average SCL per Phase",
       x = "Phase", y = "Skin Conductance Level (µS)") +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))


eda_mean_scl %>%
  dplyr::filter(device == "Embrace Plus") %>%
  ggplot(aes(x = Event_Phases, y = mean_scl, group = participant)) +
  geom_line(color = "gray70", alpha = 0.6, linewidth = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1), color = "red", width =
0.2, linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed")+
labs(title = "TSST: Embrace Average SCL per Phase",
     x = "Phase", y = "Skin Conductance Level (µS)") +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))


file_list <- list.files(pattern = "_with_phases\\.csv$")


all_data <- file_list %>% +
  labs(title = "TSST: Embrace Plus Average SCL per Phase",
       x = "Phase", y = "Skin Conductance Level (µS)") +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))


#find the outlier in biopac
```

```
# 2. List and read all files

  map_df(~ read_csv(.x, col_types = cols(timestamps = col_character())) %>%
           mutate(participant = str_remove(.x, "_with_phases\\.csv$")))

# 3. Define TSST phases
tsst_phases <- c("Baseline Seated", "Anticipation Phase", "Speech Phase",
                 "Mental Arithmetic", "TSST Recovery")

# 4. Filter for TSST and compute average Biopac tonic per participant
biopac_outliers <- all_data %>%
  dplyr::filter(Event_Phases %in% tsst_phases, !is.na(EDA_biopac_tonic)) %>%
  group_by(participant) %>%
  summarise(mean_biopac_tonic = mean(EDA_biopac_tonic, na.rm = TRUE)) %>%
  arrange(mean_biopac_tonic)

# 5. View the lowest values
print(biopac_outliers)

#Plank part

all_data <- file_list %>%
  map_df(~ read_csv(.x, col_types = cols(timestamps = col_character())) %>%
           mutate(participant = str_remove(.x, "_with_phases\\.csv$")))

# Define Plank phases
plank_phases <- c("Baseline Standing", "Plank Anticipation", "Plank Walk",
                  "Plank Baseline", "Plank Recovery")
```

```
# Filter & summarise for Plank only
eda_plank <- all_data %>%
  dplyr::filter(Event_Phases %in% plank_phases, !is.na(EDA_biopac_tonic),
!is.na(EDA_embrace_tonic)) %>%
  mutate(Event_Phases = factor(Event_Phases, levels = plank_phases))

#plot biopac
eda_biopac_plank <- eda_plank %>%
  group_by(participant, Event_Phases) %>%
  summarise(mean_scl = mean(EDA_biopac_tonic, na.rm = TRUE), .groups = "drop")

ggplot(eda_biopac_plank, aes(x = Event_Phases, y = mean_scl, group = participant)) +
  geom_line(color = "gray70", alpha = 0.6, linewidth = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1),
               width = 0.2, color = "red", linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(title = "Plank: Biopac Average SCL per Phase",
       x = "Phase", y = "Skin Conductance Level (µS)") +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

#plot Embrace
eda_embrace_plank <- eda_plank %>%
  group_by(participant, Event_Phases) %>%
  summarise(mean_scl = mean(EDA_embrace_tonic, na.rm = TRUE), .groups = "drop")

ggplot(eda_embrace_plank, aes(x = Event_Phases, y = mean_scl, group = participant)) +
  geom_line(color = "gray70", alpha = 0.6, linewidth = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
```

```
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1),
               width = 0.2, color = "red", linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(title = "Plank: Embrace Plus Average SCL per Phase",
       x = "Phase", y = "Skin Conductance Level (µS)") +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))



#log difference plot TSST
# Prepare log differences per participant and phase
eda_log_diff <- all_data %>%
  dplyr::filter(Event_Phases %in% c("Baseline Seated", "Anticipation Phase", "Speech
Phase",
                                    "Mental Arithmetic", "TSST Recovery")) %>%
  group_by(participant, Event_Phases) %>%
  summarise(
    log_biopac = mean(log(EDA_biopac_tonic), na.rm = TRUE),
    log_embrace = mean(log(EDA_embrace_tonic), na.rm = TRUE),
    .groups = "drop"
  ) %>%
  mutate(
    log_diff = log_biopac - log_embrace,
    Event_Phases = factor(Event_Phases, levels = c(
      "Baseline Seated", "Anticipation Phase", "Speech Phase",
      "Mental Arithmetic", "TSST Recovery"
    ))
  )
```

```
# Plot with green boundaries (±log(1.5))
ggplot(eda_log_diff, aes(x = Event_Phases, y = log_diff)) +
  geom_line(aes(group = participant), color = "steelblue", alpha = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1),
               width = 0.2, color = "red", linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_hline(yintercept = log(1.5), linetype = "dotted", color = "green3") +
  geom_hline(yintercept = -log(1.5), linetype = "dotted", color = "green3") +
  labs(
    title = "TSST: Log Ratio Difference in SCL (Biopac – Embrace Plus)",
    x = "Phase",
    y = "Log(Biopac) − Log(Embrace Plus)"
  ) +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

#log difference plot Plank task
# Define Plank phases
plank_phases <- c("Baseline Standing", "Plank Anticipation", "Plank Walk",
                  "Plank Baseline", "Plank Recovery")

# Calculate log differences per participant and phase
eda_log_diff_plank <- all_data %>%
  dplyr::filter(Event_Phases %in% plank_phases) %>%
  group_by(participant, Event_Phases) %>%
  summarise(
    log_biopac = mean(log(EDA_biopac_tonic), na.rm = TRUE),
    log_embrace = mean(log(EDA_embrace_tonic), na.rm = TRUE),
    .groups = "drop"
```

```
  ) %>%
  mutate(
    log_diff = log_biopac - log_embrace,
    Event_Phases = factor(Event_Phases, levels = plank_phases)
  )

#plot
ggplot(eda_log_diff_plank, aes(x = Event_Phases, y = log_diff)) +
  geom_line(aes(group = participant), color = "steelblue", alpha = 0.6) +
  stat_summary(fun = mean, geom = "line", aes(group = 1), color = "red", linewidth = 1.2) +
  stat_summary(fun.data = mean_se, geom = "errorbar", aes(group = 1),
               width = 0.2, color = "red", linewidth = 0.8) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  geom_hline(yintercept = log(1.5), linetype = "dotted", color = "green3") +
  geom_hline(yintercept = -log(1.5), linetype = "dotted", color = "green3") +
  labs(
    title = "Plank: Log Ratio Difference in SCL (Biopac – Embrace Plus)",
    x = "Phase",
    y = "Log(Biopac) − Log(Embrace Plus)"
  ) +
  theme_minimal(base_size = 12) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

#--------table for event level----------
# Define TSST phases
tsst_phases <- c("Baseline Seated", "Anticipation Phase", "Speech Phase",
                 "Mental Arithmetic", "TSST Recovery")

# Filter for TSST and valid data
eda_tsst <- all_data %>%
```

```
  dplyr:: filter(Event_Phases %in% tsst_phases,
         !is.na(EDA_biopac_tonic), !is.na(EDA_embrace_tonic)) %>%
  mutate(Event_Phases = factor(Event_Phases, levels = tsst_phases))

# Reuse the same summary function as before
summarise_scl <- function(df, scl_col) {
  df %>%
    group_by(participant, Event_Phases) %>%
    summarise(mean_scl = mean({{scl_col}}, na.rm = TRUE), .groups = "drop") %>%
    group_by(Event_Phases) %>%
    summarise(
      Mean     = mean(mean_scl),
      SD       = sd(mean_scl),
      N        = n(),
      CI_Lower = Mean - qt(0.975, df = N - 1) * SD / sqrt(N),
      CI_Upper = Mean + qt(0.975, df = N - 1) * SD / sqrt(N)
    ) %>%
    arrange(factor(Event_Phases, levels = tsst_phases))
}

# Summary table for Biopac (TSST)
tsst_biopac_summary <- summarise_scl(eda_tsst, EDA_biopac_tonic)
print("TSST Biopac SCL Summary:")
print(tsst_biopac_summary)

# Summary table for Embrace Plus (TSST)
tsst_embrace_summary <- summarise_scl(eda_tsst, EDA_embrace_tonic)
print("TSST Embrace Plus SCL Summary:")
print(tsst_embrace_summary)
```

```
# Define Plank phases
plank_phases <- c("Baseline Standing", "Plank Anticipation", "Plank Walk",
                  "Plank Baseline", "Plank Recovery")

# Filter for plank task and valid data
eda_plank <- all_data %>%
  dplyr::filter(Event_Phases %in% plank_phases,
         !is.na(EDA_biopac_tonic), !is.na(EDA_embrace_tonic)) %>%
  mutate(Event_Phases = factor(Event_Phases, levels = plank_phases))

# Function to calculate summary statistics per phase
summarise_scl <- function(df, scl_col) {
  df %>%
    group_by(participant, Event_Phases) %>%
    summarise(mean_scl = mean({{scl_col}}, na.rm = TRUE), .groups = "drop") %>%
    group_by(Event_Phases) %>%
    summarise(
      Mean     = mean(mean_scl),
      SD       = sd(mean_scl),
      N        = n(),
      CI_Lower = Mean - qt(0.975, df = N - 1) * SD / sqrt(N),
      CI_Upper = Mean + qt(0.975, df = N - 1) * SD / sqrt(N)
    ) %>%
    arrange(factor(Event_Phases, levels = plank_phases))
}

# Summary table for Biopac
biopac_summary <- summarise_scl(eda_plank, EDA_biopac_tonic)
print("Biopac SCL Summary:")
print(biopac_summary)
```

```
# Summary table for Embrace Plus
embrace_summary <- summarise_scl(eda_plank, EDA_embrace_tonic)
print("Embrace Plus SCL Summary:")
print(embrace_summary)
```