MASTER THESIS

YLab Validation Study: Measuring Stress

I.G.M. Jongmans, BSc
S1711725

EXAMINATION COMMITTEE

First Supervisor:
Dr. M. Schmettow

Second Supervisor:
Dr. R.W. Chan

8th of March 2023

FACULTY BEHAVIOUR, MANAGEMENT AND SOCIAL SCIENCES (BMS)

HUMAN FACTORS AND ENGINEERING PSYCHOLOGY (HFE)

LAB

UNIVERSITY OF TWENTE.

Abstract

Purpose

Psychophysiology and the measurement of physiological variables is an objective way to estimate cognitive states such as perceived stress. Available renowned physiological measuring devices are often expensive and lack transparency in how data is processed. Both play a factor in the shortage of hands-on psychophysiological education in some universities. Fortunately, the YLab initiative of the University of Twente has the goal to increase hands-on physiological measurement education. The YLab initiative is creating a platform where students can access relatively inexpensive hardware and free software to build physiological measurement devices. One such device is the YLab initiative's first prototype. This YLab prototype was to be experimentally validated in measuring stress against the renowned Empatica E4.

Method

The experiment included a baseline measurement, a subjective stress scale, and an adjusted Sing-a-Song Stress Task (SSST). During the experiment, both devices recorded the participants' electrodermal activity (EDA) and heart rate (HR). 12 participants were recruited through convenience sampling. All participants were given a sensor placement manual to attach the sensors themselves, which all except one wanted to follow.

Results

The validation of the YLab prototype was unsuccessful thus far, partially because of the novelty of this prototype and the barricades in terms of transparency and compatibility of the E4. Difficulties were encountered with synchronizing the devices and pre-processing the YLab data. As well as, the EDA sensor of the YLab did not seem to be functioning properly. For this multitude of reasons, no systematic correlations of .8 or higher were found between devices or with the subjective stress scale. However, the ECG sensor placement by participants who followed a sensor placement manual did foster a clear signal.

Conclusion

Despite the failed validation of the YLab prototype important insights for future research can be drawn from this study. Recommendations include the need for further defining the ECG signal pre-processing procedure to extract the HR. Furthermore, the sampling frequency of the ECG sensor should be increased. Thus, the first step was taken in investigating the YLab prototype. Therefore, the goal of the YLab initiative has come a little closer to being realized.

Table of Contents

Introduction

Psychophysiology refers to mechanisms in the brain and body that are related to mental processes. In the field of psychology, physiology is studied in relation to behavior, emotions, and perception often via direct experimental manipulation. For example, by introducing a stressful stimulus and measuring changes in physiological variables. Those changes show when and how the body reacts. As some (prolonged) changes in physiology are linked to unhealthy cognitive states or diseases (Torkamani-Azar et al., 2022; Segerstrom & Miller, 2004), insights gathered this way are useful. Understanding psychophysiological mechanisms and knowing which physiological variable values are disadvantageous can help improve (working) conditions and help prevent harm. The importance of measuring physiology can lie in the need for furthering knowledge or supporting decisions etcetera.

Devices currently available often use multiple sensors for continuous and non-intrusive physiological measurements. Depending on the cognitive state or reaction of interest, various sensors can be used to measure physiology. For example, when measuring perceived stress you can measure skin conductance, and heart rate but also cortisol levels in saliva (Marques, Silverman & Sternberg, 2010). If one would want to continuously measure cortisol in saliva that would require the sensor to be in the mouth. This sensor placement would interfere with speaking to a greater extent than e.g., skin conductance measurements on the forearm. During an experiment, (continuous) saliva measurements might be accounted for in terms of discomfort and tasks that need to be performed. However, when one wants to measure physiology in natural scenarios a cortisol measurement is less ideal. Therefore, the choice of the non-intrusive sensor(s) is often key to enabling measuring physiology in a broad range of scenarios, such as in a work environment.

In recent years wearables have been on the rise which can measure perceived stress. The available wearables range from several commercially available smartwatches to the Muse headband gen 2 (Muse, 2022) and the (in Europe) medically certified Empatica E4 wristband (Empatica, 2020b). However, most of these and other validated methods of measuring stress are unavailable, difficult to interpret, and/or too expensive for people outside of research fields (Tsao, Li & Ma, 2019). For example, the algorithms implemented in those devices are often not available for inspection, making it difficult to understand how the physiological signals are processed. Even within (research) universities the number of measurement devices can be limited. This can be to the point that teaching how to measure physiological signals by first-hand experience is rare. Despite that the necessity of such devices is dependent on the research question, the strain on such education is still a serious issue. The shortage of first-hand experience could be tackled by attempting to increase the availability of physiological measurement devices.

Thankfully, the YLab initiative within the University of Twente was established to create an educational platform that helps students gain first-hand experience with physiological measurement

devices. The YLab platform will cover the complete research workflow of creating physiological measuring devices. This workflow includes access to low-budget hardware and software to program sensors, analyze the data, and interpret the results. The students will conduct their own experiments, with the devices they have built and programmed. For the future of physiological education, this will be a crucial step in the right direction. Besides the gain on the educational front, another advantage to being able to create physiological measuring devices makes one less reliant on suppliers of the current measuring equipment. Possibly, the self-made devices can be used as an alternative to the less available store-bought equipment.

The physiological variables related to stress will be the focus of this study. Stress was chosen as it is considered a relevant topic in today's society. Furthermore, stress and the applicable sensors are already well covered in scientific literature, which is more practical in contrast to less researched phenomena. In order to use stress measurements, sufficient background knowledge of stress is essential. For this purpose, the stress topics of concern for the current study will be elaborated on below.

Stress

Stress has received attention in various fields of research and among the general public. For example, how humans experience and perceive stress, what stress is correlated with, and how it is measured (Epel et al., 2018). Stress in psychology has been considered a dynamic process that occurs when an individual appraises situational demands as exceeding available resources (Lazarus in Torkamani-Azar et al., 2022). Epel et al. (2018) state that perceived stress, in line with that definition, has the components of cognitive states and feelings. Because perceived stress is a cognitive state it is common to estimate stress by measuring its effects on individuals. Besides measuring (perceived) stress through physiology it can also be measured subjectively by ways of introspection through e.g. interviews or surveys.

Individuals might react differently cognitively or behaviorally to stress however, inside the body a distinct stress response is widely accepted. This response lies within the "unconsciously" regulating autonomic nervous system (ANS). The ANS controls the functions of a wide range of organs including the heart, skin, digestive tract, and blood vessels. It has two branches the parasympathetic nervous system (PNS) and the sympathetic nervous system (SNS). The PNS is responsible for the conservation of energy, in other words, it is responsible for functioning in a state of relaxation. On the contrary, stress activates the SNS to release hormones like cortisol and adrenaline into the blood (Indikawati & Winiarti, 2020), simultaneously, (among others) sweat production, body temperature, and heart rate (HR) increase (Critchley, 2002; Kodithuwakku Arachchige et al., 2022; Torkamani-Azar et al., 2022). These physiological changes in the SNS are part of the "fight or flight" response. This stress response has evolved to increase our fitness by keeping us from harm.

However, chronic stress is harmful. It has been linked to mental illnesses e.g. depression and anxiety disorders and physical impairments e.g. the immune system (Torkamani-Azar et al., 2022; Segerstrom & Miller, 2004). Beside health decrements, stress has also been associated with reduced task performance, which can be particularly dangerous when e.g. involving surgeons or pilots (Torkamani-Azar et al., 2022). Because of the dangers surrounding stress-induced reduced task performance, Kodithuwakku Arachchige et al. (2022) propose that physiological measurements can be used as a pre-employment screening tool e.g., in selecting air traffic controllers. Propositions like the one from Kodithuwakku Arachchige et al. (2022) could spread to other high-risk occupations i.e., where one person can inflict damage on a large scale.

The findings mentioned above again underline the importance of measuring the physiological reactions to stress for the prevention of harm, health-related or otherwise. Some examples of the available non-intrusive sensors are electrodermal activity (EDA) for the increases in sweat and thus skin conductance (Critchley, 2002), Electrocardiogram (ECG) or Photoplethysmography (PPG) (Shaffer & Ginsberg, 2017) from which HR and heart rate variability can be deduced. When it comes to these sensors various nuances should be considered as well as the devices available that incorporate them.

Electrodermal Activity

Electrodermal activity or electoral dermal activity (EDA) is the variation in the electrical conductance of the skin. EDA of the skin is mediated by sweat produced by the eccrine sweat glands. What is commonly understood is that sweat production is increased when stress is introduced (Boucsein et al., 2012; Critchley, 2002). Essentially, the conductance increases because resistance decreases due to an increase in sweat, as per Ohm's Law below (A = Ampere, V = Volt, and Ω = resistance).

$$A = \frac{V}{\Omega}$$

EDA measurements often have two main components, the slow shifts in basal skin conductance (SCL, a.k.a. tonic EDA) and rapid changes known as skin conductance responses (SCRs, a.k.a. phasic EDA) (Critchley, 2002). Because the SCRs are sensitive to stimuli and stress, SCRs are often of more interest in experiments than the entirety of the EDA recording (Braithwaite et al., 2015; Caruelle et al., 2019). Braithwaite et al. (2015) and Li et al. (2022) state that a baseline measure of 2-4 min can be enough to compare non-specific, event-related responses or unresponsiveness. In contrast, Boucsein et al. (2012) do not find the need for baseline measurements as event-related SCRs do not depend on SCL. Some EDA measurements can record the two components separately, but it is often necessary to filter out the SCRs afterwards. In such cases, higher sampling frequencies are recommended, also because higher frequencies can reduce timing errors (Caruelle et al., 2019; Li et al., 2022).

EDA can be measured by applying an external voltage or without (exosomatic and endosomatic recordings respectively). The current used in exosomatic recordings is applied between two electrodes, preferably placed on a singular palm or sole (Boucsein et al., 2012). That is because those areas have the highest density of eccrine sweat glands which makes them most sensitive during a stress response (Paloniemi et al., 2022). In particular, Boucsein et al. (2012) recommend the thenar and hypothenar parts of the palm. In other words, where the thumb is attached and its counterpart on the other outer side of the palm. Despite that, the wrist is used for EDA measurements to prevent obstruction of daily activities (Paloniemi et al., 2022). However, the wrist is considered suboptimal for various reasons (Caruelle et al., 2019; Li et al., 2022) e.g., because the wrist is more likely to reflect thermoregulation instead of relevant electrodermal phenomena (Boucsein et al., 2012). The drawback of these placements is that the wrist, palm, and fingers are all prone to motion artifacts. Therefore, depending on the activity the non-intrusiveness of a sensor placement has to be considered.

Heart Rate

Critchley (2002) and De Looff et al. (2018) stated that heart rate (HR) rises as a response to stress. This rise in HR is mainly because the heart must facilitate possible movement, which requires blood flow to the muscles involved, for the fight or flight response mentioned earlier (Critchley, 2002). HR (a.k.a. instantaneous HR) is determined by dividing 60 by the Inter Beat Interval (IBI). The IBI is also known as the R-R interval i.e., the time in between consecutive R peaks. R peaks are part of the P-Q-R-S-T wave that represents the electrical activity it takes to depolarize and repolarize the heart. Each letter of the P-Q-R-S-T wave represents a distinct part of the heartbeat measured by an ECG. Depending on the lead i.e., the configuration of the ECG electrodes, used the components of the P-Q-R-S-T wave can be positive or negative. In Figure 1 an example of a positive P-Q-R-S-T wave is given. The highest peak in the Q-R-S complex is the R peak and it signals a heartbeat (Figure 1).

Figure 1

Two consecutive P-Q-R-S-T waves with R-R interval and Q-R-S complex
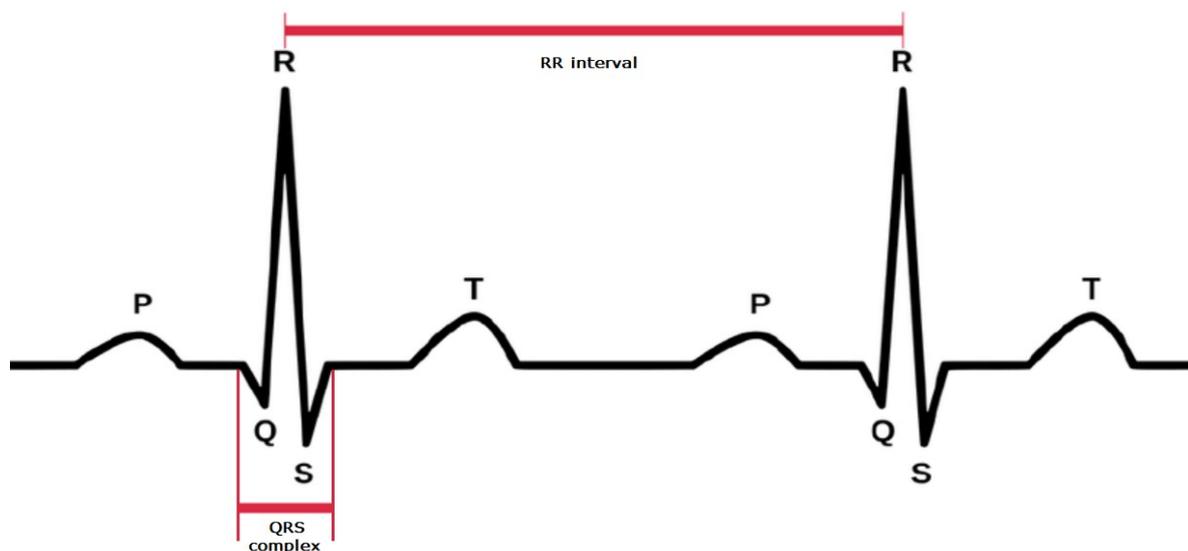
Note. Adapted from "Conceptualizing a real-time remote cardiac health monitoring system" by Page et al., 2017, *In Medical Imaging: Concepts, Methodologies, Tools, and Applications,* p. 177 (https://www.researchgate.net/publication/281176903). IGI Global.

The gold standard to measure HR is to derive it from a 12-Lead, a particular arrangement of 12 electrodes on the torso, ECG reading (Cilhoroz et al., 2020; Kim et al., 2018; Laborde et al., 2017). The (12-Lead) ECG is the preferred method because it detects the depolarization and repolarization of the heart with a high temporal resolution (Shei et al., 2022). However, leads with fewer electrodes are also available and considered reliable (Jauniaux et al., 2020; Madona et al., 2021). For example, chest straps using fewer electrodes (often 2 or more) are also seen as reliable enough to measure HR (Cilhoroz et al., 2020) such as the Polar V800 HR monitor (Krejčí et al., 2018). The placement on the chest area for both sensors is favorable because it is less sensitive to motion artifacts in comparison to placement onto the extremities.

Photoplethysmography

Photoplethysmography (PPG) can be used as an alternative to ECG for measuring heart rate (HR) (Hinde et al., 2021; Shei et al., 2022). PPG sensors utilize light to measure the absorption, reflection, or transmission of light through the skin. These are altered by tissue perfusion i.e., blood flow, indicating heartbeats and consequently HR (Shei et al., 2022).. While the ECG is superior in accuracy Cilhoroz et al. (2020) state that in clinical settings patients have reported the discomfort of the irritating adhesive of the electrodes. The experienced discomfort has led to low patient compliance. Even chest straps are less employed than wrist-worn devices due to their impracticalities (Cosoli et al., 2021). Therefore, Hinde et al. (2021) state that the wrist-worn PPG's accessibility and practicality outweigh the negligible error rates reported. Especially, with the current updated algorithms the PPG has improved as a measuring alternative (Shei et al., 2022).

However, an important nuance remains, the PPG significantly loses accuracy when the measuring conditions are not at rest e.g., when exercising. This can be due to the loss of contact with the skin, motion artifacts or sweat (Laborde et al., 2017; Shei et al., 2022; Wójcikowski & Pankiewicz, 2020). Sometimes manufacturers of smartwatches/wristbands that employ this sensor also warn about the motion handicap of their devices (Empatica, 2020a). Another major drawback of the PPG is that it is not optimized for measuring HR in people of color (Bradford et al., 2022; Shei at al., 2022). The PPG is not powerful enough to accommodate higher pigmentation of the skin, often resulting in erroneous signals or no signal at all (Bradford et al., 2022).

Current study

In this thesis, the YLab prototype and the Empatica E4 will be compared. Current implementations of the YLab initiative's efforts resulted in the YLab prototype. Coined the "YLab", the YLab is a device containing two sensors to measure stress. Those sensors would match the sensors applied in the Empatica E4, a renowned device for personal and professional use (Empatica, 2021).

The professional use of the E4 can be in medical settings for patient monitoring and research settings such as in universities (Chandra et al., 2021). The E4 was chosen as a reference device because of its regular use in research settings.

The current study combines the aim to improve and increase available physiological education and the importance of measuring the physiological stress response. The stress induced in the experiment that will be measured are electro dermal activity (EDA) and heart rate (HR). Between both devices and variables a positive correlation is hypothesized. Furthermore, a subjective stress scale is introduced to investigate if the devices correlate with those values. Again, a positive correlation is hypothesized. During this study, the signal of the YLab was investigated. Attempts were made to produce a suitable data processing procedure/algorithm. This paper aims to describe the pragmatic validation study to assess if the YLab prototype compares to the Empatica E4 in measuring stress.

The research questions include,

RQ1 Explore if the YLab can measure the same EDA as the E4 with a correlation of .8 or higher,

Sub RQ1.1 Explore if the YLab and E4 EDA data correlate with the subjective stress measured,

RQ2 find a way to detect the R peaks in the ECG data of the YLab,

Sub RQ2.1 Explore if the YLab can measure the same HR as the E4 with a correlation of .8 or higher,

Sub RQ2.2 Explore if the YLab and E4 HR data correlate with the subjective stress measured,

Sub RQ2.3 Explore if the placement of the ECG electrodes by participants can foster an accurate signal,

Sub RQ2.4 Explore if the use of linear interpolation to compensate for possible missing beats and/or rejected sections in the HR data of the YLab and the E4 leads to a correlation of .8 or higher.

Method

Procedure

For this validation study participants were welcomed to the experiment room by the researcher. After the participant was informed about the study and their rights, they consented to participate. Then they were given instructions and the sensor placement manual. The manual explained how to attach the sensors of the YLab prototype to themselves, including how to wear the Empatica E4. If the participant preferred they could leave the experiment room to attach the ECG sensors in private as those had to be attached to their torso. Depending on whether the participant wanted to leave the experiment room or not, these preparations took approximately 10 – 20 minutes.

The participant was asked to find a comfortable position with their lower arms resting on the table to prevent movement. The sensors were attached to the YLab by the researcher. Then only the Empatica E4 was started. When the heart rate (HR) signal for the E4 stabilized the experiment was started simultaneously with the recording of the YLab. The experiment instructions were verbally given to the participant by the researcher as well as displayed on a computer screen. For example, a reminder to sit as still as possible until you see "End task" on the introduction slide of the experiment. The entire experiment sequence was presented to the participants in a PowerPoint on a computer screen. The baseline measurement took 5 minutes, during which the participants had to keep their eyes open to prevent them from dozing off. The instructions of/and the current Sing a Song Stress Task (SSST) took approximately 10 minutes. After the task was completed the measuring devices were turned off by the researcher and the participant was asked to remove the sensors. Then, the final questionnaire was filled in by the participant. Finally, the participant was debriefed and thanked for their participation. These closing steps after the SSST took approximately 10 minutes. In total the experiment took 35 to 45 minutes. The procedure was summarized in Figure 2.

Figure 2

YLab Validation Study Procedure Summary
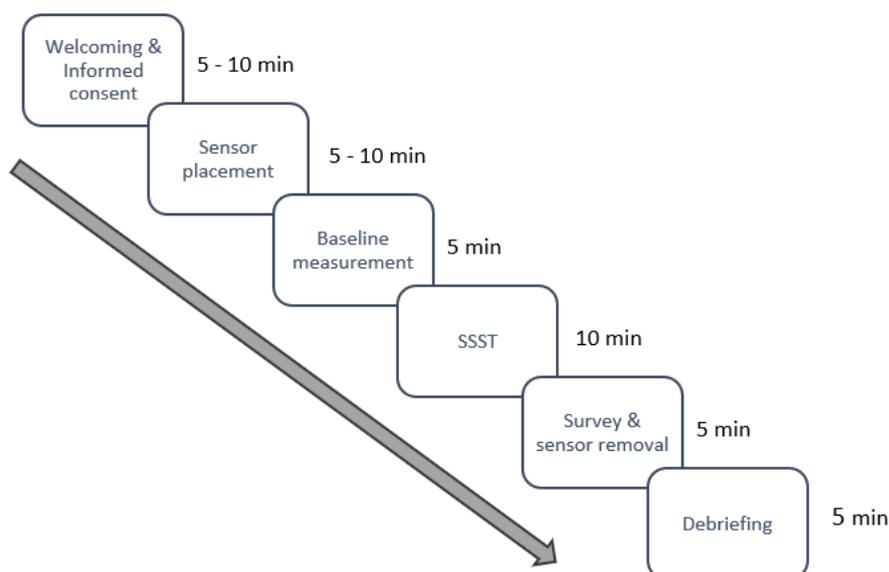
Sensor Placement Manual

The manual included the placement of the Empatica E4, the YLab ECG electrodes, and lastly the YLab electrodermal activity (EDA) sensor (Appendix A). With considerations for the privacy and comfort of the participants, the researcher aimed to stay outside of the participant's personal space as much as possible. The images of the Empatica E4 user manual (Empatica, 2021) were mirrored, to increase the ease of following that part of the manual. Because for the current study the participant had to wear the E4 on their right instead of their left hand. With a copy of the manual to follow, the participant was given the option to place the ECG sensors on their torso in the privacy of a bathroom. Only two of the twelve participants made use of that option. The YLab ECG sensor placement part of the manual included an adapted picture with reference points from Cadogan (2022). Lastly, the YLab EDA sensors part of the manual included an adapted picture from Cionek (2020) on where to wear the finger sleeves. The EDA sensors were attached last as they hindered the placement of the electrodes and E4.

The participants were given the option to let the researcher check their placement, in case a participant felt unsure about their sensor placement. Moreover, when they were comfortable with the researcher placing the sensors they did not have to do it themselves. However, they were encouraged and presented the manual first. Only one of the twelve participants wanted the researcher to place the sensors.

Baseline and Stress Scale

A baseline measurement and stress scale were included in the experiment to have more data to compare the devices with. The baseline measurement was implemented following the advice to do so for EDA measurements (Braithwaite et al., 2015; Li et al., 2022). The stress scale was implemented to correlate the physiological variables with as way to compare the devices. The participant was asked to rate the stress induced by the experiment verbally on a 7-point Likert scale as the following:

1. Not at all stressed
2. Low stress
3. Slightly stressed
4. Average
5. Moderately stressed
6. Very stressed
7. Extremely stressed

The stress scale was presented six times to the participant. The first one was presented immediately after the introduction slide i.e., before the baseline measurement of 5 minutes began. After the baseline measurement was finished the same question was asked again. The other four times the participants were asked to rate their subjective stress were during the SSST. This was done to hopefully capture the increase in perceived stress due to the SSST.

After feedback from one of the first two participants on the meaning of point "4. Average" the scale was changed to:

1. Not at all stressed
2. Low stress
3. Slightly stressed
4. Moderately stressed*
5. Very stressed*
6. Highly stressed*
7. Extremely stressed

Because none of the two participants had chosen the adjusted points their values did not need to be converted. The bipolar scale was intended to capture stress that might arise from participating in an experiment or stress the participants potentially feel in general. In hindsight, it did not make sense to the participants, whereas during the pilot testing nobody questioned the scale.

The Adjusted SSST

A Sing a Song Stress Task (SSST) aims to elicit stress based on social evaluation and singing, while adhering to ethical standards like the original SSST (Brouwer and Hogervorst, 2014). The SSST starts with participants following instructions to read sentences in their mind or out loud alternated with 1-minute intervals. At a certain point, the participant is asked to sing out loud. In experiments, participants can be told their singing will be judged afterwards or utilize confederates to enhance the social stress aspect (Brouwer and Hogervorst, 2014; van der Mee et al., 2020).

The adjusted SSST was inspired by Brouwer and Hogervorst (2014) and van der Mee et al. (2020). The neutral statements in the current SSST were derived via the same method as van der Mee et al. (2020), i.e. English sentences were taken from the Wikipedia page on vacuum cleaners. The length of the adjusted SSST was approximately 10 minutes long, between the length of the SSST (Brouwer and Hogervorst, 2014) and the Short SSST (van der Mee et al., 2020).

The current SSST was displayed in PowerPoint, in the same file as the baseline measurement, immediately after the second subjective stress scale was presented. This SSST is equipped with an example slide, it includes an example sequence of instruction slides and a counter slide. There was no practice set of slides. After the example slide, the first counter begins and the task sequence follows. After the final countdown is presented the end of the task is shown on the final slide. The first counter slide included the 7-point scale alternating with the counter-only slides until the task ended. The adjusted SSST is summarized in Figure 3 and the full version can be found in Appendix B.

Reminder cues were given to the participant in case they did not follow instructions or did not react to the scale, e.g. "Please follow the instructions on screen" and "Please indicate your level of stress." For the singing part of the SSST encouraging cues were prepared, e.g. "Please sing the song you thought of" or if the participant was not responsive after that "Please sing any song that comes to

mind." It was estimated that this adjusted SSST would still be a valid stress inducing task like the ones it is based on, i.e. the original SSST (Brouwer and Hogervorst, 2014) and the Short SSST (van der Mee et al., 2020).

Figure 3

Adjusted Sing a Song Stress Task (SSST) summary.
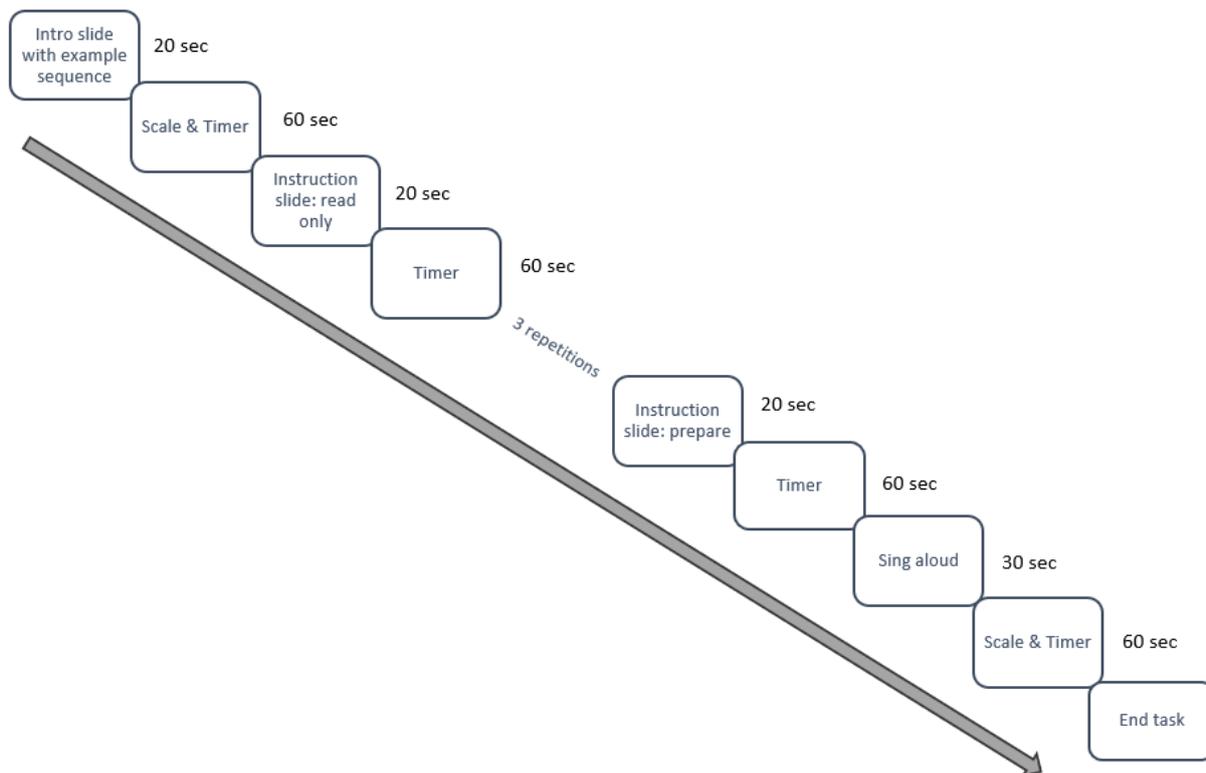
Survey and Informed Consent

The informed consent and survey were presented to the participant with Qualtrics, with a break in between to signal the experiment would begin. This was done to keep all data online and in one file. The survey after the SSST and sensor removal contained questions regarding demographics, the participant's experience, frequency, and comfort regarding singing alone and in front of the researcher. Furthermore, it included a question on previous experience with the SSST. The final questions were to check if they followed the instructions prior to the experiment to stop consuming alcohol and caffeine and cease strenuous exercise. As well as, to check their health status concerning diseases for the exclusion criteria.

The participants were reminded to cease consuming alcohol or caffeine 12 hours before the experiment per email 24 hours prior to the experiment. Strenuous exercise was also asked to be avoided 12 hours before the experiment. Because those behaviors have affected the physiological reaction of the heart (Krejčí et al., 2018; Li et al., 2022). This was only verified in the survey because, despite the possibly influenced data patterns, the YLab prototype and Empatica E4 should still be able to measure the variables with a high degree of accuracy. Furthermore, no inferences will be drawn

based on the physiology of the participants. Therefore, those behaviors did not need to be exclusion criteria. In summary, the survey was collected to potentially explain differences in correlations between the devices and the subjective stress scale.

Materials and Measures

The electrodermal activity (EDA) and heart rate (HR) data from the YLab and the E4 were matched and correlated. For both devices, the variable EDA was measured directly in the same manner. The HR was calculated based on the inter beat interval (IBI) for both devices afterwards. The devices had different sensors to measure the IBI but those were matched and correlated too. The IBI was chosen over the HR as it was closest to the raw data signal from the E4.

Empatica E4

The Empatica E4 contains four sensors of which two will be used to measure stress in the current study: the Photoplethysmography (PPG) and EDA (Empatica, 2020b). In the E4 the EDA is measured with a sampling frequency of 4Hz. The PPG is an optical technique to measure the volumetric changes in blood being pumped through the arteries by the heart. This is measured with a blood volume pulse at 64 Hz, which served as the basis for the IBI and thus the calculated HR. The algorithm which filters the IBI for true R peaks from noise is not publicly available (Empatica, 2020a).

YLab prototype

The YLab prototype (Figure 4) is a low-budget Do-It-Yourself device that consists of various products (Table 1). The basis of the YLab is the Maker Pi Pico (rev 1.2) board on which the Raspberry Pi Pico 2040 (RP2040) is preinstalled. The RP2040 is the microcontroller that is programmed in Python (Appendix C) please refer to the YLab GitHub for all necessary libraries etcetera (Schmettow, 2023). The python code was run in Thonny (v4.0.1) with Python (v3.10.6, 64-bit) and Tk (v8.6.12) to use the YLab prototype. The raw signals are written in CSV files on an SD card. The signals are retrieved via the grove system, to which the grove-compatible GSR sensor (v1.2) to measure EDA and EMG detector (v1.1) to measure ECG were attached.

The EMG detector was fed through the ADS1115 to convert the analog signal to a digital output. The ADS1115 has a 16-bit precision and a programmable gain amplifier which can boost the signal up to 16 times. The EMG detector made use of a three-electrode cord, meaning a 5 or 12 Lead ECG sensor configuration was not possible. Therefore the Lead-II ECG configuration was chosen as it was feasible with the three electrodes and that Lead-II was effective in previous research (Madona et al. 2021). The choice to use ECG measurement instead of a PPG sensor is because ECG is a more accurate and effective measurement of heartbeats (Chandra et al., 2021; Van Lier et al., 2020). The EDA was sampled at 4Hz to match the E4, and the ECG was sampled at 100Hz. The EDA and ECG sensors were called yeda and yema, respectively.
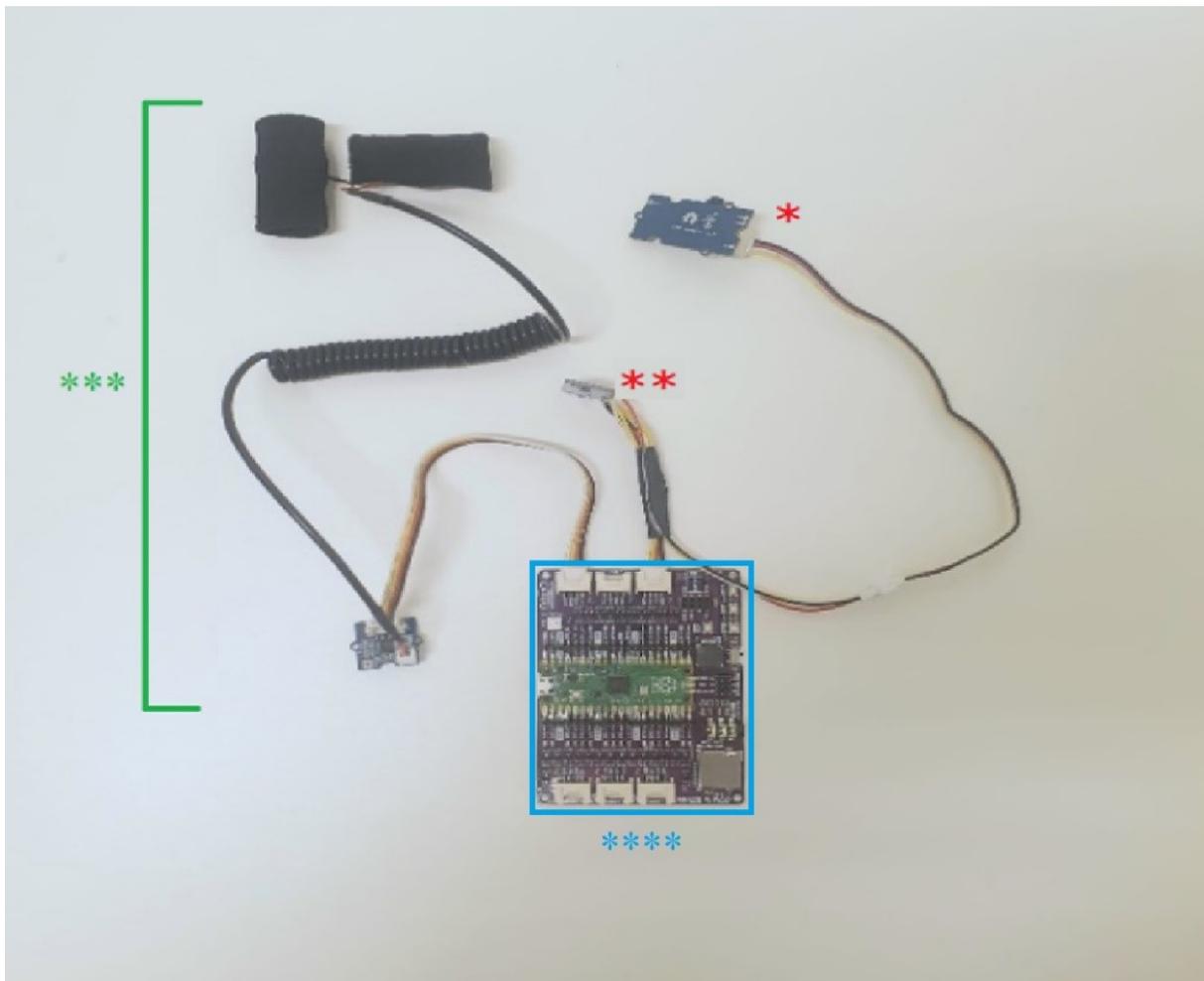
Table 1

Products Used in the YLab Prototype

Product	Producer
Maker Pi Pico rev 1.2	Cytron Technologies Sdn Bhd
ADS1115	Adafruit Industries LLC
Grove – GRS v1.2	Seeed Technology Co., Ltd
Grove – EMG sensor v1.1	Seeed Technology Co., Ltd
Grove – 4 pin cable	Seeed Technology Co., Ltd

Figure 4

YLab Prototype

Note. The parts used to create the YLab are denoted as *Grove – EMG sensor v1.1,** ADS1115, ***

Grove – GRS v1.2 and ****Maker Pi Pico rev 1.2 which are connected by the Grove – 4 pin cables.

Participants

12 participants (9 females) aged 19 to 29 (M = 23.8, SD = 3.8) were recruited via convenience sampling and through the "SONA system" provided by the University of Twente. SONA participants that completed this experiment were rewarded 2 SONA points. The participants recruited in the network of the researcher did not receive a reward for their participation as they were not eligible for SONA points. The exclusion criteria for this study were heart disease or epilepsy.

The sample size for this study was based on the initial power calculations of Van Lier et al. (2020). The study of Van Lier et al. (2020) also concerned validating a device and the Empatica E4 and was thus deemed relevant. In their study, the signal and parameter level sample size were based on a correlation of .8 and effect size of 1.8. While $[.8^2 /(1 - .8^2)]$ and the Bland-Altman analysis. Both suggested the sample size of 6 participants, Van Lier et al. (2020) increased it to 55 participants. However, due to the resource constraints of this study no more than 12 participants could be achieved. This was to an extent justified as this first validation study of the YLab will only use correlations and no other statistical analysis.

Data analysis and Design

This project has been approved by the Ethics committee of the University of Twente in the Netherlands under number 221300. In this study, a within-subjects design was implemented. The individual differences over time in heart rate (HR) and electrodermal activity (EDA) derived from the devices were compared with each other and the subjective stress scores.

During the data analysis correlations within subjects were performed between the YLab and E4, on the signal level for EDA and the parameter level for HR. These variable correlations aimed to answer RQ1 and sub RQ2.1, respectively. Moreover, the YLab and E4 data for both EDA and HR were correlated with the subjective stress scale that aimed to measure the construct stress and answer sub RQ1.1 and sub RQ2.2, respectively. The correlations were then compared between devices.

Time synchronization

The raw data of the YLab had to undergo transformations, most importantly synchronizing the data. Due to the scope and timeframe of this study, no synchronization software could be created to match the start of E4 and YLab recordings. Nor were there resources to create an algorithm to match recordings after the experiment. Other timing issues occurred because the E4 has a lag between pressing start and recording/displaying the measurements of all signals in the E4 realtime (v2.1.1) streaming app. Even when the HR appeared on the streaming app it showed a calibration period where the first estimations of the HR went up from e.g., 52 bpm to stabilize around 82 bpm.

Unlike the E4, the YLab does not have a lag or timing issues in starting the recording of the measurements. Luckily, the ending of the recordings could be done simultaneously for both devices. Therefore, they were stopped at the same time by the researcher. The last observations were matched and the YLab served as a timeline because it started to record with the start of the experiment. Because

the E4 was started in advance of the experiment it had more data than the YLab, given that the E4 algorithm did not reject too much data in the beginning. The excess E4 data from before the experiment was deleted.

The real time for the YLab was calculated based on an arbitrary number of seconds passed which was given to every participant. For example, the first 3 observations of "time" on the YLab were [3422.28, 3422.3, and 3422.32] and so on due to the sampling frequency. When the entire time series was deducted by 3422 the real timeline was created, resulting in the times from the example to become [0.28, 0.3, and 0.32] respectively. This arbitrary number differed per experiment therefore the researcher had to pre-process the data by reducing the start value accordingly.

Similarly to Milstein and Gordon (2020), the synchronization of the output from both devices was done in Excel. This was convenient because the YLab and the E4 both recorded CSV files. After the real time and other pre-processing steps were completed in R the data was exported to Excel. Then the YLab and E4 data were aligned based on the last observation and the subjective stress scale values were inserted. Finally, the newly aligned data was imported back to R for the data analysis.

**Signal and parameter level pre-processing**

The pre-processing was mostly necessary for the data of both the yeda and yema sensors (EDA signal and HR parameter level respectively). However, for the YLab and the E4 the HR had to be calculated with the IBI as mentioned above. Furthermore, both devices underwent the same linear interpolation process. The EDA data from the yeda sensor had to be transformed as the skin resistance in Ohm (Ω) was measured instead of skin conductance in Siemens (S). Per the formula below (reciprocal of Ohms Law: A = Ampere, V = Volt, and Ω = resistance), 1 was divided by the yeda values to get the Siemens under the assumption A is constant. The EDA data could be correlated immediately afterwards as it was sampled at 4Hz for both devices. The Empatica E4 measures the EDA in micro-Siemens (μS) and the yeda was transformed to Siemens.

$$1\,S = \frac{1\,A}{1\,V} = \frac{1}{1\,\Omega}$$

The yema data was first visually inspected per participant for signs of the R peak in the Q-R-S complex of the heartbeat wave. Figure 5 shows a random fragment of 4 seconds for all participants, a visual difference can be seen between the higher peaks resembling R peaks and other lower waves for most participants. Despite that, the peaks are more prominent in some participants, for all participants the R peaks were tried to be detected. The visual inspection was also necessary to answer sub RQ2.3. As mentioned earlier, the HR for both devices were computed based on the inter beat interval (IBI). To calculate the IBI, first, the R peaks and their time stamps had to be detected. Three approaches were tried to detect true R peaks in the yema data.

Figure 5

Excerpt Yema Sensor of the YLab Prototype ECG Values Representing Heart Rate Waves Participants
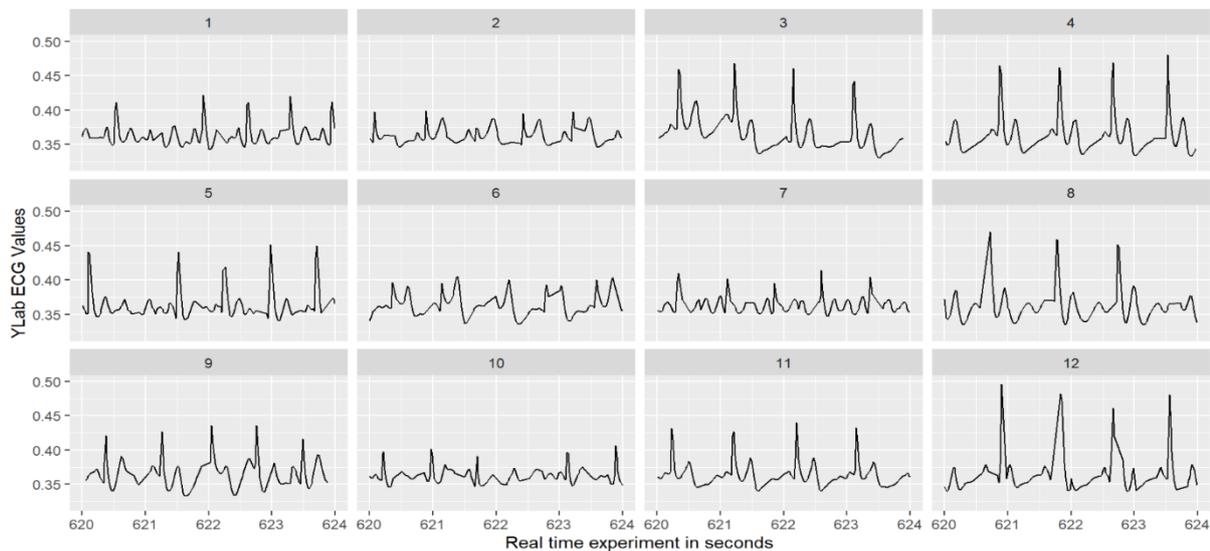
For that purpose, the yema data had to be filtered multiple times. However, the conventional frequency filters such as Butterworth, Fourier transformations, or notch filters were not applied because the yema data looked similar to the result of those conventional filters (Basu & Mamud, 2020; Dotsinsky, 2005; Shankar & Babu, 2020). It is unknown if the E4 applies similar PPG filters to the PPG data e.g., those proposed by Liang et al. (2018). In a way, the filtering process is an attempt to mimic what the E4 algorithm does but after the data is collected. For the yema data, 4 filters based on the work from Francesca et al. (2018) on low-complexity algorithm steps were created. The 4 filters were tried by compounding them one by one. Using the 4 filters together seemed to deliver the best output, therefore they were used together as the first approach to pre-processing the yema data.

Two filters were used to detect potential R peaks, which are distinct and rapid increases in the ECG value. The first filter was based on the height of the peak. For this, the positive outliers of the yema sensor values were used to choose a threshold. The second filter was based on the derivative between the yema sensor values. With this derivative, all peaks in the data could be detected by looking at when a positive value was followed by a negative value. However, for the R peaks, positive outliers from the derivative were used to choose a threshold. High derivative values are essentially big increases in the yema values and thus high peaks, similar to R peaks. Based on the values left over after both filters, the timestamps when potential R peaks and thus heartbeats occurred were extracted.

Then the IBI was calculated between the consecutive potential heartbeats. The yema data thus far contains the potential peaks and some extreme values of IBIs. This is because false peaks, which often have very short IBIs, and missed peaks, which cause very long IBIs, are still included. Two IBI filters were thought of inspired by the E4 algorithm. Besides rejecting false peaks somehow, the algorithm only uses the first 2 accepted peaks after a rejected section to calculate the next recorded IBI

(Empatica, 2020a). In other words, rejected sections are not included in the recorded dataset, to achieve a similar result the extreme IBIs were filtered out.

To this end, two IBI filters were estimated based on outliers from the yema data. The third filter was for the lower end of IBIs. IBIs shorter than .4 seconds were filtered out because, IBIs of 0.01, 0.02, and 0.03 seconds were most likely measured on the rising side of the same peak. Then the fourth filter was for the upper end of the IBIs, filtering out outliers of 1.2 seconds or longer. The pinpointing of the threshold for the filters also took into consideration the HR which would result from the accepted IBIs. Because HR = 60/IBI, the IBI thresholds for the lower and upper end would give a HR of 150 and 50, respectively. This range was deemed broad enough considering the seated participants. This first approach of using the 4 filters served as the fundament for the remaining two approaches to pre-process the yema data.

The second approach was to filter the yema IBIs on the minimum and maximum IBI lengths the E4 algorithm accepted per participant, which were expected to be more conservative. This approach is more influenced by the E4 algorithm as the E4 is already validated. In other words, this approach only differed from approach 1 by the thresholds given to the lower and upper IBI filters.

It must be noted however that filtering the IBI lengths means a section of rejected and/or missed peaks will not show in the data. The exclusion of such sections in the recorded data is stated by Empatica (2020a). Therefore, by mimicking the E4 algorithm neither will such sections show the yema data. Moreover, it is important to note that the E4 and the YLab might reject/filter the data in differing places. In other words, when the yema might have filtered/missed a R peak the E4 might not have, potentially creating a mismatch between the datasets. This potential problem was not controlled for because this method of filtering is part of the validated E4 algorithm. Thus, the yema data of pre-processing approaches 1 and 2 were aligned with the original E4 IBI data to compare.

Interpolation was used in the final third approach, in an attempt to reveal all rejected sections and estimated the misses for both the YLab and the E4. This pre-processing approach could prevent a potential mismatch between the datasets. Various forms of interpolation are available when dealing with missing data (or rejected sections). For example, cubic spline and linear interpolation are popular and relatively easy options (North & Livingstone, 2013). However, when evaluating the cubic spline interpolation on the yema data even more extreme values were estimated. This is most likely due to the number of consecutive misses. Therefore, linear interpolation was chosen to be used on the IBIs. Moreover, Empatica (2020a) also mentions linear interpolating if necessary.

Empatica (2020a) also provides a way to detect the hidden rejected timespans in the data. For the yema, the method is similar but more straightforward as the long IBIs are present in the data. Therefore, for both devices, the very long IBIs were located, and the number of misses was calculated by dividing the long IBI by the median IBI. The median was used because the long outlier IBIs could influence the mean whereas the median is more robust to outliers. Finally, the inserted empty spots were linearly interpolated. This was also necessary to answer sub RQ2.4.

Results

The visual inspection of the data, pre-processing, and analysis were performed in RStudio (v4.2.2.) (RStudio Team, 2020) (Appendix D). First participant level plots were made to look at how the values compare between the YLab and E4 for the electrodermal activity (EDA) and then for the three pre-processing approaches for the heart rate (HR). All visualizations are of excerpts of the data around the last three minutes of the SSST because between seconds 815 – 845 the singing task was performed. Around this timeframe, the stress and thus EDA and HR of the participants were estimated to increase or peak as a result of the singing.

Electrodermal activity

Figure 6

Excerpt EDA Values Yeda Sensor and E4 for last 3 Minutes of the Experiment
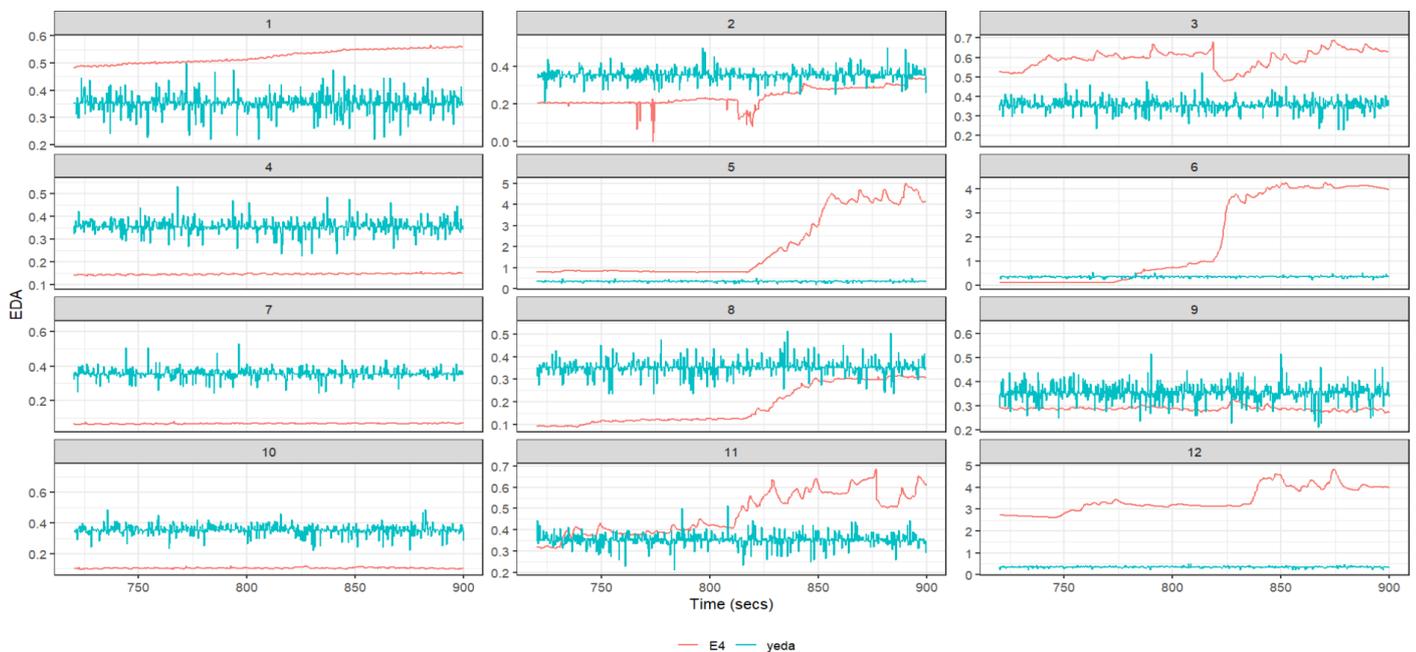
As can be seen above the EDA measurements from the yeda sensor and E4 sensor do not seem similar. Unsurprisingly, the correlations look poor between the devices in Figure 7 and were revealed to be poor after statistical analysis in Table 2. The cluster formations for the E4 in Figure 7 are most likely the peaks and higher plateaus visible in Figure 6. The most prominent plateau occurred in participant 6 around second 820 (Figure 6). In contrast to the yeda values that seem to have peaks but no systematic increase in EDA. The yeda sensor seems to not have measured EDA but possibly static of some sort, as there seems to be an oscillation around a set value that differs per participant (Figure 6). Whereas normally the EDA would contain SCL which is a certain rate of increase or decrease in the baseline (Boucsein et al., 2012). Moreover, the concern for a defective yeda sensor seems to be supported by the E4 data from e.g., participants 5, 8, and 12 where the E4 seemed to have measured SCL (Figure 6).

Figure 7

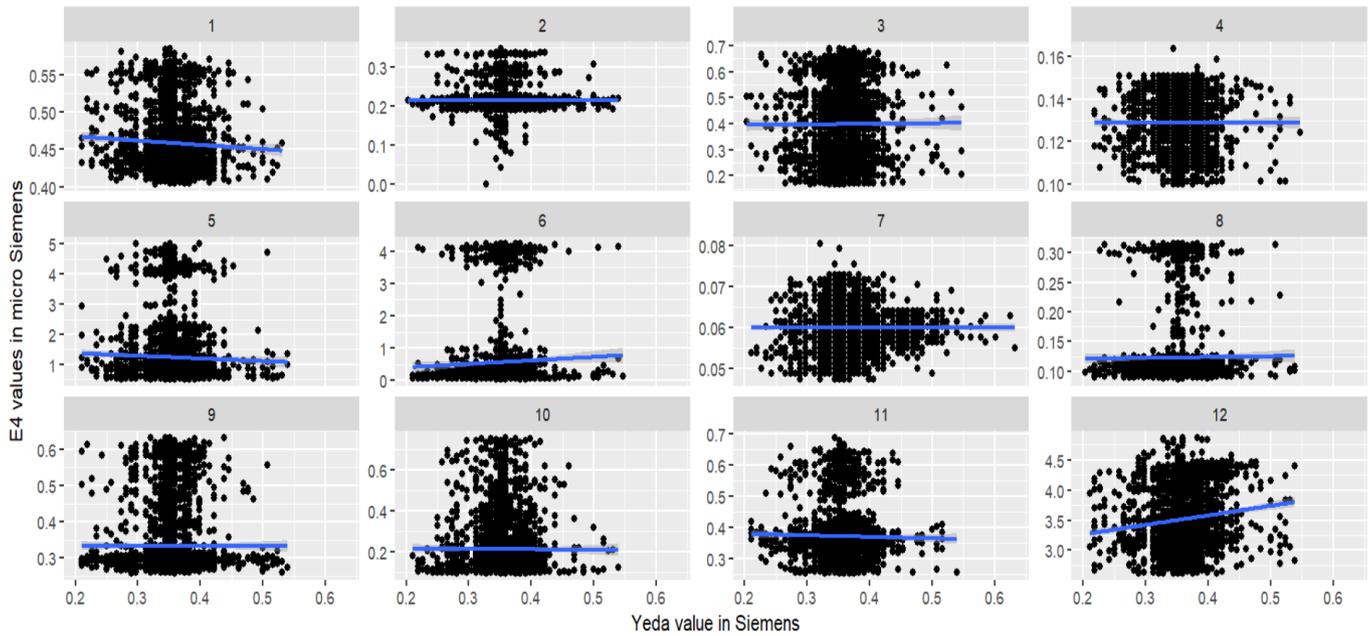*Scatterplot EDA Data between the E4 and the Yeda Sensor of the YLab Prototype*

Table 2

Correlation EDA between Yeda Sensor and the E4

	EDA
Participant	**yeda and E4**
1	-0.043
2	0.002
3	0.005
4	0.000
5	-0.030
6	0.030
7	0.001
8	0.008
9	-0.001
10	-0.004
11	-0.018
12	0.113

Heart Rate

For an exploration of the yema sensor, of the yema values were plotted against the E4. The three approaches were looked at separately. In summary, the approaches were:

1. Use yema IBI outliers to filter the yema IBI (Figures 8 & 9)

2. Use E4 IBI min and max values to filter the yema IBI (Figures 10 & 11)

3. Use linear interpolation on both the yema and E4 IBI data (Figures 12 & 13)

Figure 8

Excerpt HR Values Yema Sensor and E4 for last 3 Minutes Experiment Approach 1
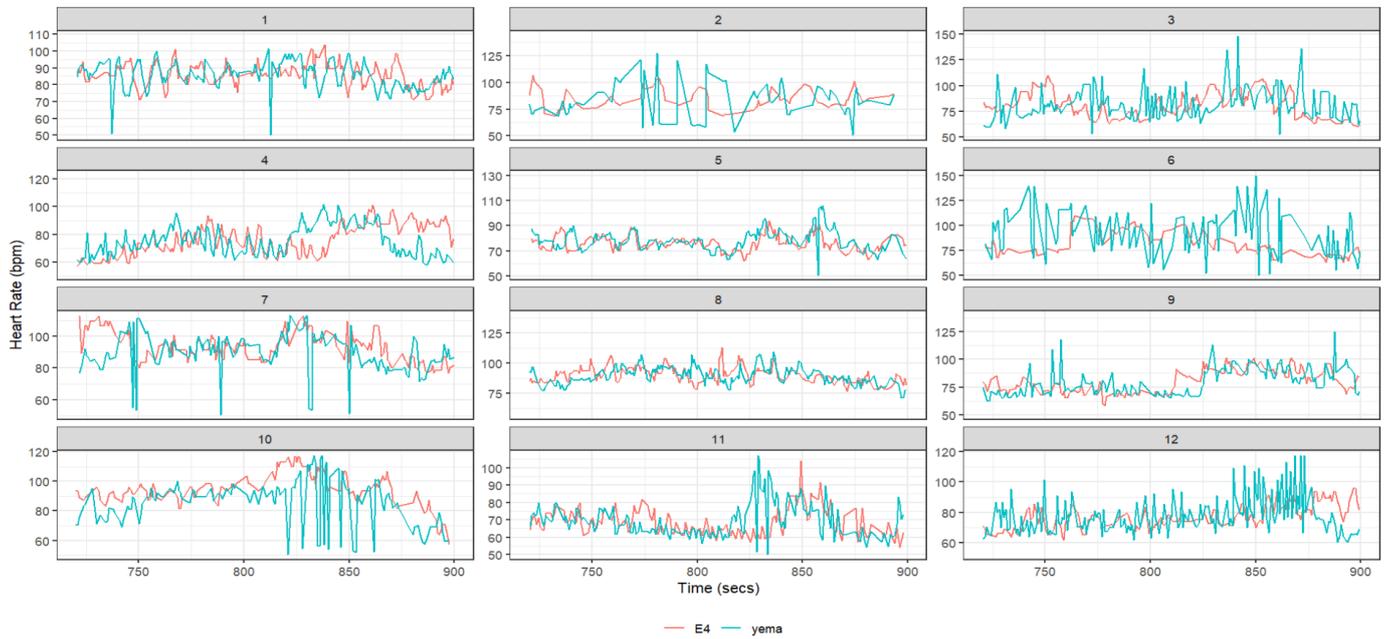

Figure 9

Scatterplot HR Data between the Yema Sensor and the E4 for Approach 1.

Figure 10

Excerpt HR Values Yema Sensor and E4 for last 3 Minutes Experiment Approach 2

Figure 11

Scatterplot HR Data between the Yema Sensor and the E4 for Approach 2.

Figure 12

Excerpt HR Values Yema Sensor and E4 for last 3 Minutes Experiment Approach 3
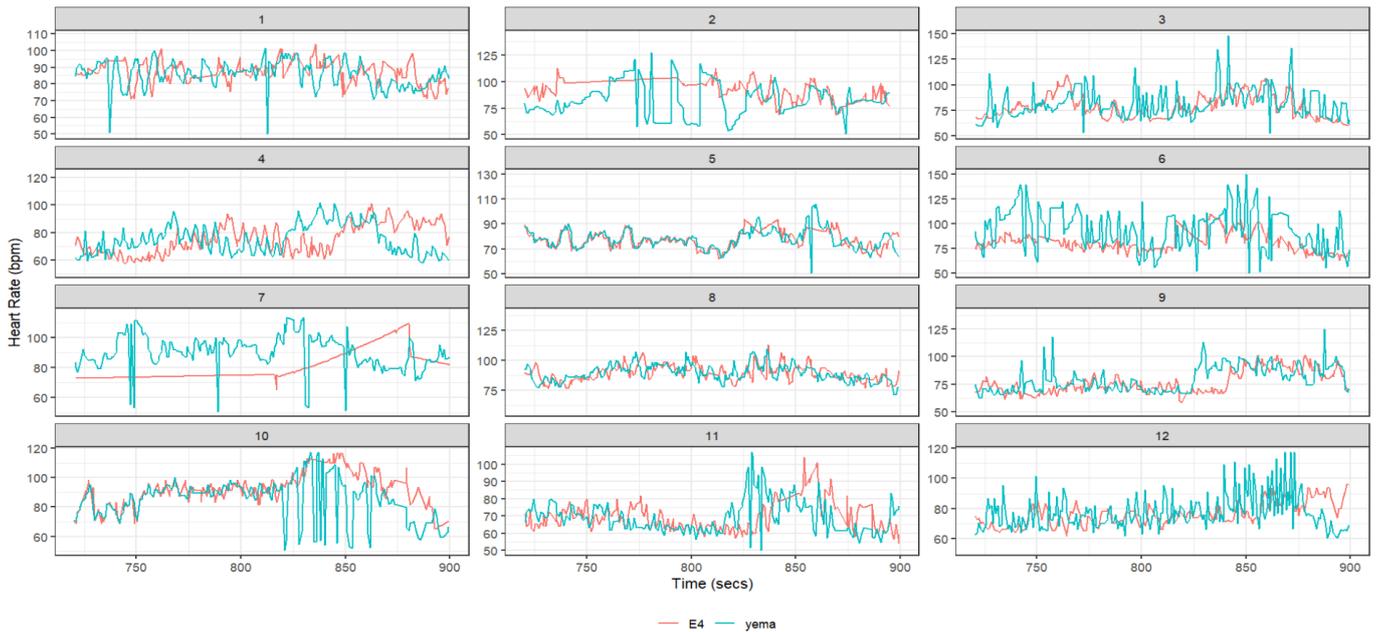
Figure 13

Scatterplot HR Data between the Yema Sensor and the E4 for Approach 3.
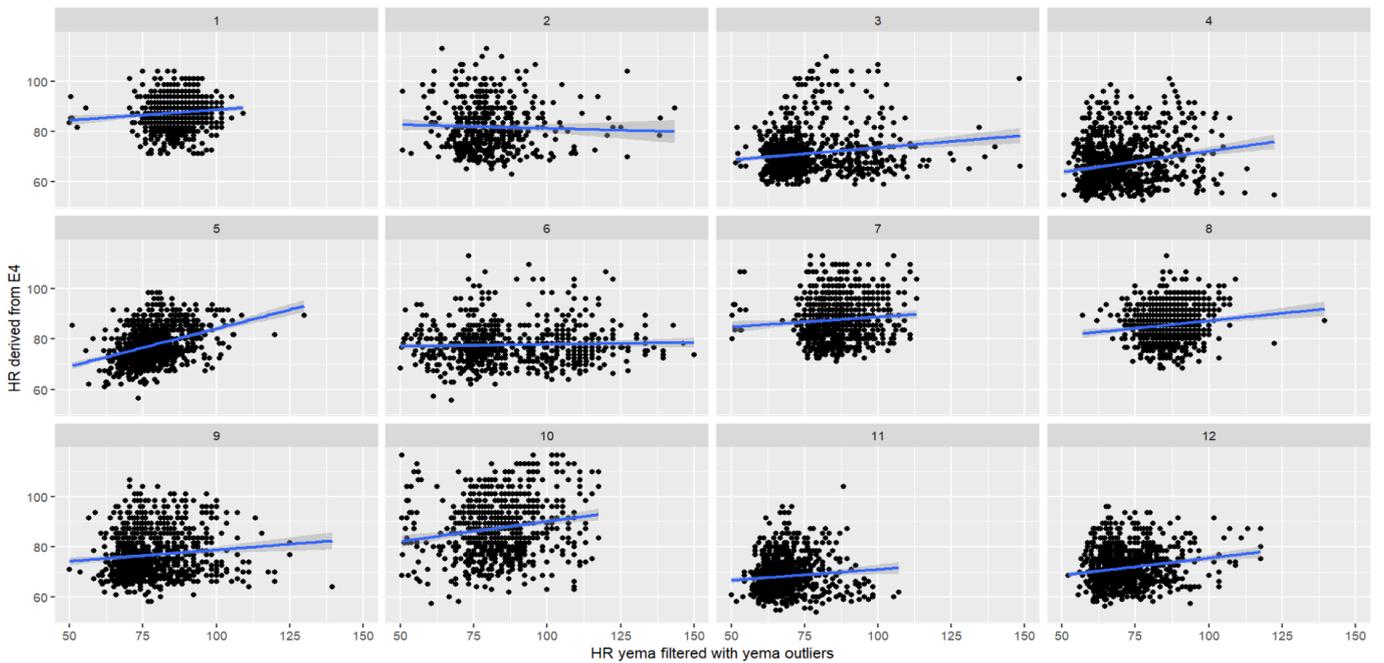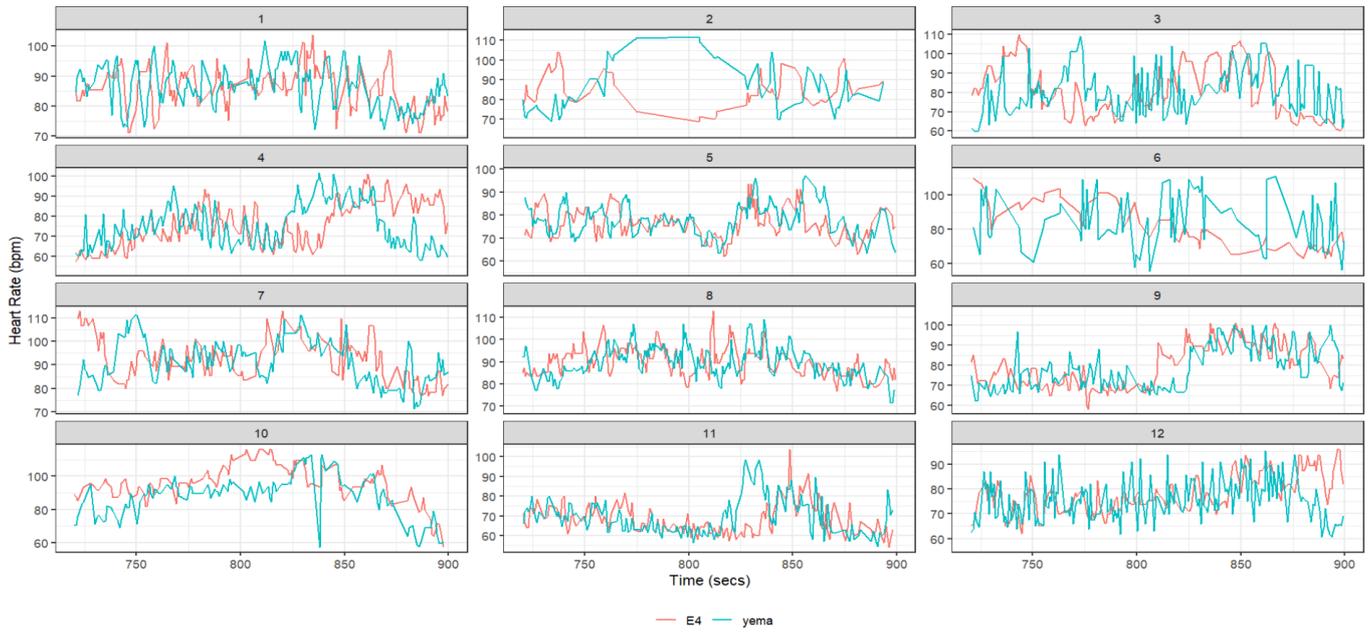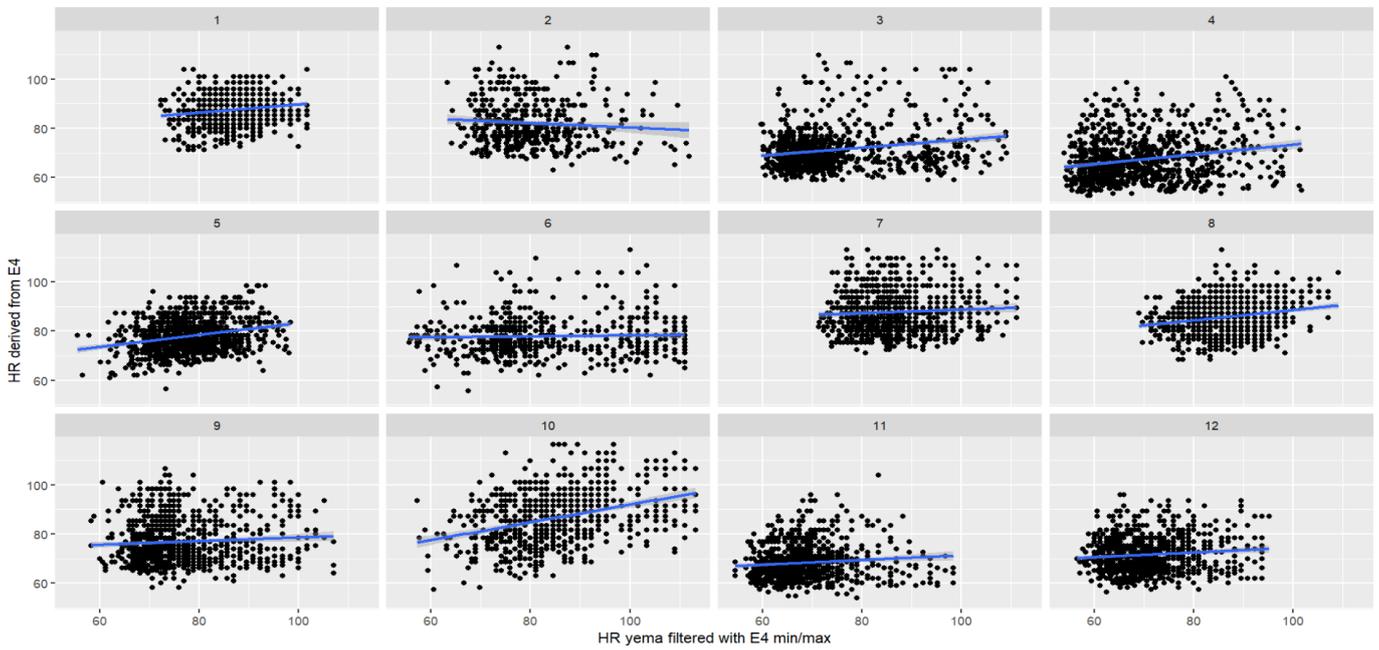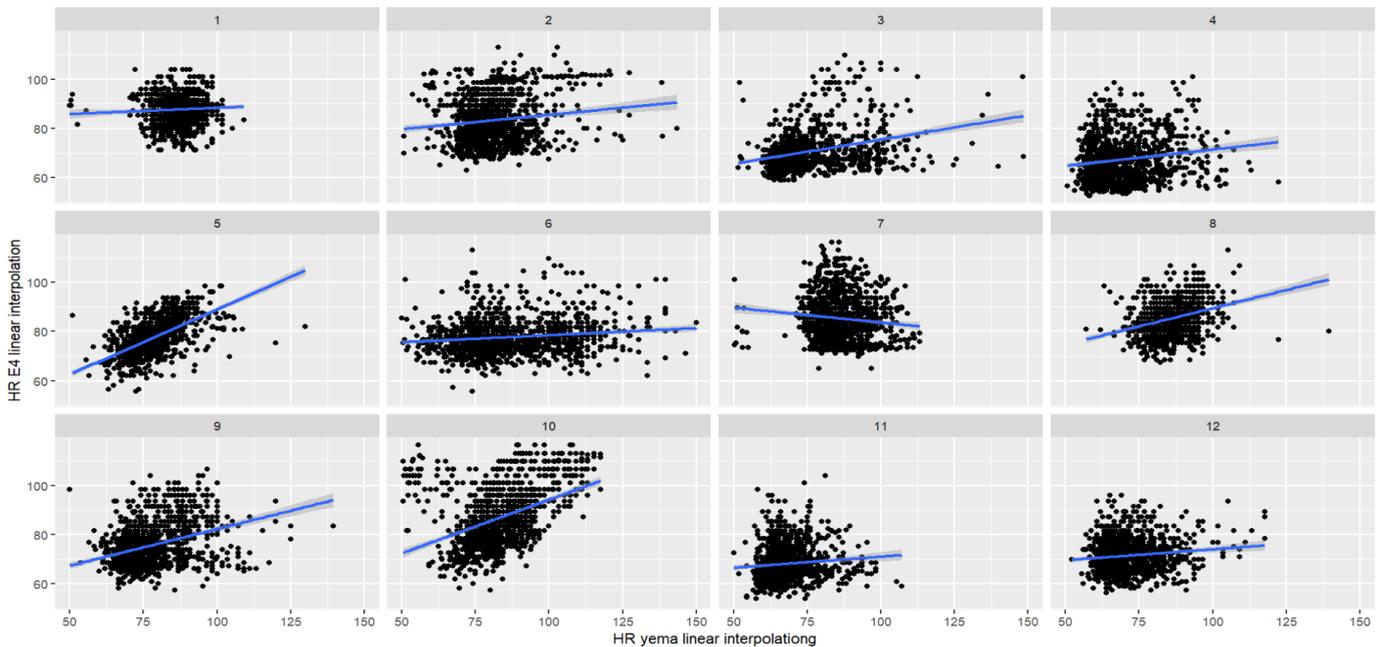

Concerns based on irregular patterns in the data will be described below. The irregular patterns visible in the excerpt of data in the figures above are presumed to be present in the entire recording. For all approaches, it is clear that the yema has radical changes between consecutive HR values. Essentially, the very low to very high HRs are a result of long and short IBIs respectively. This is a concern because it is not normal for the HR to go change this drastically within seconds, such changes could be due to an error in the sensor or pre-processing to detect R peaks. Especially in approach 1, these rapid HR changes are present, but it gets reduced in approach 2 for example in participants 7 and

10. This is due to the E4 min and max IBI thresholds of those participants being more conservative than the min and max IBI from the yema outlier. In other words, hypothetically in approach 1 the lowest IBI accepted for a participant was .4 whereas the minimum E4 IBI value used in approach 2 was .56. Vice versa would be true for the longest accepted IBIs. These changes in threshold values from approach 1 to 2 result in less extreme IBIs being accepted and consequently fewer radical changes in HR being present in participants 7 and 10 (Figure 10).

Furthermore, the third approach raises other concerns with regard to the missed sections in the data. This is due to the interpolation for the rejected sections or "misses", which are most prominently visible in participants 2 and 7 in the E4 data (Figure 12). What has happened in e.g., participant 7 is that the section between timestamps 720 – 815 seconds has been missed. This section has then been interpolated, between similar values, to cause the almost straight line of seemingly repeating values. Afterwards in the section between timestamp 825 – 885 seconds, the same happens but now between vastly differing HR estimations. Therefore, the line connecting the points has a positive slope. These segments are disconcerting as it shows the E4 misses 155 seconds of the 300 seconds displayed in the data visualizations for participant 7.

In regard to the correlations, no systematic visual correlation could be seen for all participants of approaches 1 and 2 (Figures 9 & 11). However, the third approach showed more promising fitted lines (Figure 13). This seems to be partially confirmed by the correlations in Table 3. The small difference in overall correlation between the first and second approach could be due to the minor change in IBI filters. As can be seen clearly with participants 1 and 7, between Figures 9 and 11 only some outliers are removed but the better part of the scatterplots stays the same. In conclusion, no correlations higher than .8 were found for either the EDA or the three HR approaches between the devices. However, the yeda values were correlated to a lower extent than the yema values with the E4.

Table 3

Correlation HR between Yema Sensor and the E4 per Approach

Participant	HR approach 1	HR approach 2	HR approach 3
1	0.098	0.168	0.057
2	-0.037	-0.086	0.136
3	0.152	0.210	0.309
4	0.200	0.218	0.158
5	0.367	0.262	0.597
6	0.039	0.035	0.141
7	0.084	0.068	-0.109
8	0.129	0.210	0.314
9	0.108	0.076	0.355
10	0.172	0.351	0.424
11	0.097	0.105	0.094
12	0.182	0.104	0.117

Subjective Stress Scale

Lastly, the subjective stress scale was correlated per participant per device per measurement. A wide variety of correlations was found between participants as well as between devices (Tables 4 and 5). The E4 EDA seemed to correlate better with the subjective stress scale but the negative correlations for other participants were quite high as well (Table 4). Then, the opposite is the case with the HR where all three approaches of the yema seemed to correlate better to the subjective stress scale (Table 5). This is most noticeable for participants 4, 5, and 9. The exact same correlations for participants 4 and 5 for approaches 1 and 2 are due to the IBI filters not changing the HRs at the time stamp the subjective stress score was appended. Contrarily to the differences between the devices and variables, some similarities are present. Participants 2 and 4 show some consistency in significance for the EDA and HR approach 3. It seems participant 2 did better with the E4 and participant 4 did better with the YLab. Nothing particularly distinctive about these participants or their behavior during the experiment were noted or retrieved from the movement sensor in the E4. The one thing that comes to mind is that the sensor attachment of the YLab might have been exceptionally well done for the fourth participant.

Table 4

Correlation Subjective Stress Scale and Measurements of the EDA for the YLab and the E4

	EDA	
Participant	**yeda**	**E4**
1	0.113	0.430
2	-0.232	0.991**
3	-0.159	0.806*
4	0.829*	0.000
5	0.262	0.620
6	0.223	-0.521
7	-0.333	0.638
8	-0.349	0.971**
9	-0.466	-0.991
10	-0.453	-0.696
11	-0.506	0.658
12	-0.476	-0.105

*high correlation of between .8 and .9 **very high correlation of .9 or above.

Table 5

Correlation Subjective Stress Scale and Measurements of the HR for the YLab and the E4 per Approach

	HR Approach 1		HR Approach 2		HR Approach 3	
Participant	**yema**	**E4**	**yema**	**E4**	**Yema**	**E4**
1	0.049	-0.145	0.049	-0.760	0.049	-0.022
2	0.775	0.284	0.775	0.003	-0.285	0.878*
3	-0.622	0.605	-0.622	-0.224	-0.622	0.466
4	0.897*	0.354	0.897*	0.300	0.921**	0.533
5	0.876*	0.317	0.876*	0.376	0.876*	0.825*
6	-0.224	-0.187	0.439	0.156	-0.340	-0.550
7	-0.392	0.632	-0.392	0.672	-0.252	0.302
8	0.622	-0.570	0.622	-0.486	0.622	0.677
9	0.863*	0.700	0.863*	0.225	0.863*	0.211
10	-0.658	-0.023	-0.658	-0.083	-0.587	0.009
11	-0.389	0.632	-0.389	0.427	-0.413	0.499
12	0.173	0.636	0.173	0.419	0.173	0.190

*high correlation of between .8 and .9 **very high correlation of .9 or above.

In summary, the E4 on average performed better for both the raw EDA measurement and the measurement in correlation to the subjective stress scale. Then the HR measurements seemed to improve with the linear interpolation for both the devices and regarding the correlations with the subjective stress scale. Overall, the results were conclusive that the correlations between the devices for both EDA and HR are not up to the level that the YLab can be fully validated.

Discussion

This study served as a first attempt to validate the YLab prototype when measuring stress in comparison to the Empatica E4. During the experiment that included a baseline and a stress task both devices were used simultaneously to measure the EDA and HR of participants. Because the E4 is validated, the assumption was that if the measurements of the electrodermal activity (EDA) and heart rate (HR) would correlate to a high degree between devices the YLab would be validated. The pre-processed EDA and HR data were correlated but no systematic significant correlations were found.

Therefore, the YLab prototype was not validated as the RQ1 "Explore if the YLab can measure the same EDA as the E4 with a correlation of .8 or higher" and Sub RQ2.1 "Explore if the YLab can measure the same HR as the E4 with a correlation of .8 or higher" were not supported. Essentially, for the HR related research questions no support was found because those depended on RQ2 "find a way to detect the R peaks in the ECG data of the YLab" which was not achieved/supported. In other words, if the R peaks could not be properly detected then the IBI is unreliable, thus the calculated HR of the yema would be faulty. Even the use of linear interpolation on the missing R peaks sections was not fully supported (Sub RQ2.4). Furthermore, the research questions regarding correlations between the devices for both EDA and HR and with the subjective stress scale were not supported either (Sub RQ1.1 and SubRQ2.2 respectively). Several explanations as to why the results did not support any of these research questions will be described in the next sections.

In contrast, only sub RQ2.3 "Explore if placement of the ECG electrodes by participants can foster an accurate signal" seem to have some support. When inspecting the data visually it seems peaks are present, these peaks are considered R peaks (Figure 5). Despite that 11 of the participants placed the sensors themselves with the help of the sensor placement manual. This is an important observation considering the sampling frequency is lower than most other ECG set-ups. It is assumed that when the sampling frequency is increased the signal quality will improve. Thus, for future research, the use of a sensor placement manual can be encouraged. The comfort and privacy of the participant can be maintained with the sensor placement by the participant. Furthermore, the manual is useful because of the current risks associated with close contact between researcher and participant, related to the pandemic, movements, or otherwise.

Yeda

The EDA data from the YLab seems to have picked up a different/erroneous signal because a cable was not functioning properly (Figure 6). Unfortunately, this realization happened after data collection had ended when the results were visualized. When testing if the yeda sensor was defect, it was confirmed something was wrong. The sensor did not produce an appropriate signal change when the two electrodes were placed onto each other. The recorded EDA data was not visualized in the pilot

phase. This is because the ECG data required a visual inspection for R peaks whereas the EDA data did not. This will be a valuable lesson for the researcher and future research with the YLab.

The cables for the GSR sensor, or the entire sensor, need to be replaced before new research with the YLab will be performed. Then, the new data needs to be visualized before data collection and should be similar to the E4 signal (Figure 6). This is because EDA data usually contains slow rises (the SCL) and peak moments (the SCRs) (Critchley, 2002). Another way to potentially improve the sensor once it is working, is to increase the sampling frequency to be higher than the 4Hz of the E4. This is because the YLab has the option to do that and using a higher sampling frequency has resulted in higher quality EDA data in previous research (Caruelle et al., 2019; Li et al., 2022).

Yema

Existing packages

Numerous ways of processing the ECG signal were attempted. Among which, the "heartBeat" R package (Emilpl, 2015). This package was originally developed for a single lead ECG Zephyr signal. Despite that, it was attempted to process the yema data with it. The heartBeat package includes the frequency filters, from wavelet decomposition transformations like Fourier transformations, which were earlier deemed unnecessary for the yema data. The frequency filters were tried because it was a mandatory step in the walkthrough of the package. However, when the package seemingly did not result in any heartbeats detected it was realized problems had occurred. By backtracking each step in the process it became clear where the problems lied.

Starting with, the yema data did not match the required format nor values of the example data from the heartBeat package walkthrough. After the format was adjusted, the threshold value necessary for the R peak detection later on was changed to match the yema data. The selection of an alternative value for the threshold was informed by visual inspection. Now that the heartBeat package functioned, it was still no use. A noticeable contrast was between the example sample frequency of 250Hz and the sampling frequency of the yema. In the current study, the yema sampling frequency was expected to be 100Hz however it came closer to 44Hz. Even when adjusting the sampling frequency in the heartBeat package it did not change the outcome. Whilst unfortunate, this is to be expected with such a discrepancy in sampling frequencies. Further down the script, it became clear how much the yema data was impacted by the sampling frequency.

Wavelet coefficients are influenced by the pronouncement of peaks in the data. The package calculates the wavelet coefficients that will then be filtered based on the earlier established threshold. The wavelet coefficient filtering is done to detect peaks. In Figure 5 the yema peaks are less pronounced and systematic over all participants. Therefore, the package was tried with participant 4 which has one of the most pronounced peaks. However, this still led to the wavelet coefficients being too different from the heartBeat example data. In comparison, the yema data produced inconsistent heights and not as pronounced wavelet coefficient peaks.

From this point on, the use of the heartBeat package was halted. In future research, the sampling frequency of the yema should be increased. Moreover, it could prove useful to adapt the yema output to match the heartBeat package, or other packages like the RHRV package (Martínez et al., 2017), format and values. Matching the yema data to existing formats could lead to smoother pre-processing.

The 3 HR approaches

For the yema sensor and the three approaches to estimate HR, the lack of systematic correlations over all participants seems to point to a need to further define the data pre-processing. Despite that the current pre-processing was derived based on Francesca et al. (2018). And it was similar to the heartBeat package in the use of thresholds for peak detection, more refined pre-processing is necessary. Entirely different approaches might be created for future research. When choosing a method to expand upon in future research the following considerations should be kept in mind.

Between the approaches, the amount of observation differed. There was a 3.1 % decrease in the number of observations between approach 1 (the yema min/max IBI) and approach 2 (the E4 min/max IBI). Whereas the linear interpolation added 17.2% to the observations in from approach 1 for the yema and 13.3% for the E4. Considering that the yema is a prototype and had a reduced sampling frequency these percentages are not shocking. Especially, when comparing the yema increase in observations to that of the E4. In Figure 12 the interpolation method revealed the sections of missing/rejected heartbeats by the E4 algorithm. The interpolation method, while excluded from the E4 algorithm, did cultivate the best correlations between devices in this study. While it would be best for both devices to be able to measure all R peaks accurately, interpolation might be a good method to insert the missing data.

However, linear interpolation might not be the best option. A version of interpolation that takes the best of cubic spline and linear interpolation exists called constrained cubic spline interpolation (Kruger, 2003). It has the flexibility of cubic spline interpolation but stays within the bounds of the two values it must interpolate in between like linear interpolation. This will solve the problem of the under and overshooting from the cubic spline and the rigidness of the linear interpolation. Constrained cubic spline interpolation will hopefully increase the quality of the estimated misses in future research.

Subjective stress scale

Another weakness of the experiment was the stress scale. The need to change the scale does not bode well for the validity of the scale. The participants were informed when making an appointment to partake in the experiment that it involved stress. Therefore, the subjective stress scale should have made sense during the experiment. But one participant mentioned they did not feel the word "stress" was appropriate, because to the participant the experiment did not induce such strong feelings. If this is an issue of the participant having the bias due to knowing it is a stress experiment a double-blind study setup could be used. To accomplish that setup, the lead experimenter should set up the study without revealing the task involves stress. This would include randomly dividing the

participants into the control and intervention groups. The control group might be given an SSST that does not involve the singing instruction and the intervention group would be presented with the adjusted SSST. Then a colleague could collect the data without knowing the division.

Then, the appraisal of the word stress. The appraisal of stress can be negative or positive. This might be why the scores on the scale did not match well with the measurements of the devices. If the participant scored their stress as low because they appraised the feeling to excitement (eustress) their physiological measurements could still behave similarly to a (distress) stress response. This is one of the drawbacks of introspection methods. In future research using subjective measures can be attempted again with a more robust scale.

Hardware and software

Hardware and software difficulties were encountered in this study. As mentioned previously, the sampling frequency of the yema sensor had dropped from the 100Hz to approximately 44 Hz. This seems to be a hardware issue. The YLab has a 16-bit precision which was at first estimated to be enough for the study. However, the current signal takes up approximately 14 kilobits to write all the raw data into the CSV file. This can potentially be solved by upgrading the hardware to increase the processing capacity. Alternatively, it can be tried to write the data with a different method or to another location than an SD card e.g., to the flash storage instead. The storage capacity will have to be checked per length of the experiment as the flash storage might not be big enough for longer experiments.

Another option is to try and solve the data bottleneck. This can be done by applying the algorithms to filter out the yema data before it gets written in the CSV file. It is currently unknown how much the computing will take up on the capacity of the Raspberry Pi Pico 2040 microcontroller. However, because of the low complexity of the filtering steps it should be possible, much like in the work of Francesca et al. (2018). This is similar to what the E4 does and would greatly reduce the amount of data that needs to be recorded. The drawback of this method is the loss of access to the raw signal, ironically this was one of the hinderances with the E4. Therefore, it might still be better to try and solve the lack of processing capacity of the raw data with an upgrade in hardware.

Another thing that could have potentially slowed the YLab down is the use of Python software. Python is one of the favorite programming languages in academia because it is robust and language-oriented, making it easier to learn (Marowka, 2018). However, Python is also regarded as a slow language (Marowka, 2018) because of its line-by-line interpretation e.g., in comparison to C++ which has a compiled interpretation (White, n.d.). Because the YLab initiative is meant for students it might be a leap to convert the YLab prototype to one of the other programming languages. However, it could be done and might help with the speed of processing the data. Furthermore, it could help in overcoming boundaries to existing data processing packages in other languages such as C++ or

MATLAB. It is outside the scope of this study to compare the programming languages available to boost the data recording speed for the YLab.

This data recording speed issue brings up the question of how the E4 manages to process the data. While their algorithm filters the raw data, the E4 also contains a body temperature sensor (4Hz) and MEMS type 3-axis accelerometer sensors (32Hz) (Empatica, 2020b). Therefore, the amount of data that the E4 has to record is estimated to be higher than the YLab. Furthermore, in contrast to the YLab, the E4 also uses Bluetooth to connect with the streaming app. One would assume that should also slow down the processing. Due to a lack of transparency, it is hard to find out how the E4 manages the amount of data. This makes it difficult to draw any specific tips for future research from the E4.

Another concern regarding the Empatica E4, despite the device's validation, is how well it can measure what it describes it does. Similar to Van Lier et al. (2020) who could not fully support the validity of the E4 and stated additional research on the E4 is needed. Moreover, Bradford et al. (2022) state concern about how well a PPG can work on the skin of people of color. In this study, the different sensors and placement might have influenced the measurements to an extent that correlations are insufficient. The PPG on the wrist is one of the lesser placement and sensor combinations. Similarly, the EDA sensor on the wrist has the risk of also picking up thermal regulation signals instead of the true EDA fluctuations (Boucsein et al., 2012). When designing such a (medical) device which also doubles as a research tool, the degrees of freedom for sensor placement are hindered. This loss of degrees of freedom and thus accuracy of measurements might be a valid reason to discontinue the use of the Empatica E4 as a reference device in future research.

Then, the worrisome aforementioned missing data sections of the E4 that were revealed by interpolation. In Figure 12 for participant 7 the rejected sections lasted for 155 seconds in total from the 3 minutes displayed. Empatica warns of data signal loss due to movement (Empatica, 2020a). However, participants were instructed and reminded to sit as still as possible. They were told why that was important and had to lay their lower arms on the table to further prevent the need for movement/adjusting their posture. The exposing of the missed beats by the E4 also reiterates how hard it is to produce a solid algorithm for detecting heartbeats. Unfortunately, whether that is due to the subpar sensor or algorithm is not clear and cannot be investigated without the raw E4 signal. If even in such an experiment an additional 13.3% of observations were revealed by the linear interpolation, it begs the question of how well the E4 functions and how well it has functioned in previous research. Especially in research where more movement was involved.

In a similar vein, how did previous research achieve the synchronization of the data with the E4 and devices without Unix timestamps/clocks like the YLab. The way the E4 is incompatible with other devices makes it exceedingly difficult to use it as a reference device in a validation study. In participants 4, 9, and 11 the time synchronization difficulties are visible most prominently. Whether looking at Figures 8, 10, or 12 the yema signal shows a similar pattern to the E4 data but synchronized

incorrectly. For participants 4 and 11, the yema signal is ca. 25 seconds behind the E4 and if moved to the right would have overlapped to a higher degree. What is interesting is that for participant 9 the yema signal needed to be moved to the right only in Figure 12. In Figures 8 and 10 the yema signal seems to be in front of the E4 signal and would have to move to the left to overlap to a higher degree. This is most likely due to E4 data being interpolated before the 3-minute section shown in Figure 12. Other segments and other participants most likely have similar issues, but for the current study the visualizations were already enough to conclude the time synchronization had failed.

There are various solutions to the encountered time synchronization issues. Possibly by finding a way to remotely control and start the YLab and E4 simultaneously. A variation on that can be a coordinated centralized system for all data streams. Another way is to add a Unix timestamp or clock synchronization of the computer to the YLab, this can then be used as an anchor point to synchronize the devices. The Unix timestamp was tried to be recorded instead of the arbitrary time, but it slowed down the YLab to a degree it did not function properly anymore. If none of that is achievable with the E4, it might be more convenient to find a different validated reference device that can be synchronized e.g., the Polar V800 HR monitor. A change in reference device also enables comparison to devices that incorporate ECG signals which are deemed more reliable than that of a PPG (Kim et al., 2018; Laborde et al., 2017). One could even go as far as to use multiple renowned devices to correlate each of the YLab sensors with. However, it is outside the scope of this study to compare all other available devices for future YLab research.

The E4 was chosen as a reference device because it is often used in research, had the sensors of interest, and was available at the time this study took place. However, in hindsight, it might not have been the right choice as a reference device due to a lack of transparency. Although it makes sense that a company like Empatica cannot share its algorithm or specific hardware set-up from a business perspective. Empatica remains marketing the E4 towards use in research, but the research setting comes with the need for transparency. Transparency is the cornerstone of science. Therefore, from a scientific perspective, it might not be appropriate to use such a device with an undisclosed algorithm that can not be reviewed as a reference device. The E4 might not be suitable for use in validation studies for that reason alone, not to mention the time synchronization issues and missed data in non-movement conditions.

Conclusion

The strength of this study lies in the extensive exploration on how to process the ECG signal and the recommendations for future research. Another commendable thing is that the YLab has the degrees of freedom to prioritize which and where to apply the sensors based on the literature. The current findings might not have been what was hoped for the YLab prototype, but they are important nonetheless. The findings of this study highlight how difficult it is to produce a suitable processing procedure for physiological measurements.

In conclusion, the substantial first step was made in investigating the YLab, a device that uses open-source software and low-budget hardware for measuring physiology. Both the YLab and the Empatica E4 have their issues. Therefore, it might as well have been impossible to correlate the data to .8 or higher. The research questions were pragmatic and ambitious, maybe a bit too ambitious for a first validation study. However, by implementing the mentioned recommendations for future research the YLab will most likely live up to the expectations eventually. On the condition of at least increasing the sampling frequencies and finding a way to synchronize the time with the reference device, whichever may be used next. This should be prioritized in future validation studies. More work will need to be done to achieve the goal of the YLab initiative and get the YLab prototype validated.

References

Basu, S., & Mamud, S. (2020, September). Comparative Study on the Effect of Order and Cut off Frequency of Butterworth Low Pass Filter for Removal of Noise in ECG Signal. *IEEE 1st International Conference for Convergence in Engineering (ICCE) (pp. 156-160).* IEEE. DOI:10.1109/ICCE50343.2020.9290646

Bradford, D. E., DeFalco, A., Perkins, E., Carbajal, I., Kwasa, J., Goodman, F., ... & Joyner, K. (2022). Whose Signals Are We Amplifying? Towards a More Equitable Clinical Psychophysiology. Preprint doi:10.31234/osf.io/c2naf

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments. Psychophysiology, 49(1), 1017-1034.

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., ... & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. Psychophysiology, 49(8), 1017-1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Brouwer, A. M., & Hogervorst, M. A. (2014). A new paradigm to induce mental stress: the Sing-a-Song Stress Test (SSST). Frontiers in neuroscience, 8, 224. https://doi.org/10.3389/fnins.2014.00224

Cadogan, M. (2022, Jan 30). *ECG Lead positioning*. Life in the Fastlane. Retrieved from https://litfl.com/ecg-lead-positioning/ with permission under Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).

Caruelle, D., Gustafsson, A., Shams, P., & Lervik-Olsen, L. (2019). The use of electrodermal activity (EDA) measurement to understand consumer emotions–A literature review and a call for action. Journal of Business Research, 104, 146-160. https://doi.org/10.1016/j.jbusres.2019.06.041

Chandra, V., Priyarup, A., & Sethia, D. (2021). Comparative Study of Physiological Signals from Empatica E4 Wristband for Stress Classification. In International Conference on Advances in Computing and Data Sciences (pp. 218-229). Springer, Cham.

Cilhoroz, B., Giles, D., Zaleski, A., Taylor, B., Fernhall, B., & Pescatello, L. (2020). Validation of the Polar V800 heart rate monitor and comparison of artifact correction methods among adults with hypertension. PloS one, 15(10), e0240220.https://doi.org/10.1371/journal.pone.0240220

Cionek, J. (2020, Sep 22). *Galvanic Skin Response module (GSR module) EEG BCI.* Brain Support. Retrieved from https://www.brainlatam.com/blog/galvanic-skin-response-module-gsr-module-eeg-bci-2089

Cosoli, G., Spinsante, S., & Scalise, L. (2021). Wearable Devices and Diagnostic Apps: Beyond the Borders of Traditional Medicine, But What About Their Accuracy and Reliability?. IEEE Instrumentation & Measurement Magazine, 24(6), 89-94. doi:10.1109/MIM.2021.9513636

Critchley, H. D. (2002). Electrodermal responses: what happens in the brain. The Neuroscientist, 8(2), 132-142. https://doi.org/10.1177/107385840200800209

De Looff, P. C., Cornet, L. J. M., Embregts, P. J. C. M., Nijman, H. L. I., & Didden, H. C. M. (2018). Associations of sympathetic and parasympathetic activity in job stress and burnout: A systematic review. PLoS One, 13(10), e0205741. https://doi.org/10.1371/journal.pone.0205741

Dotsinsky, I. (2005). Suppression of AC railway power-line interference in ECG signals recorded by public access defibrillators. *BioMedical Engineering OnLine, 4*(1), 1-8. DOI:10.1186/1475-925X-4-65

Emilpl. (2015). Walkthrough for the heartBeat package [R script]. Retrieved on 23rd of December 2022 from https://github.com/ehrscape/R-project/blob/master/HeartBeat-walkthrough.md

Empatica. (2020a, Jan 24). E4 data – IBI expected signal. Retrieved on 06 Dec 2022 from https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal

Empatica. (2020b, Jan 24). E4 wristband technical specifications. Retrieved on 29 Sept 2022 from: https://support.empatica.com/hc/en-us/articles/202581999-E4-wristband-technical-specifications

Empatica. (2021, Oct 21). Empatica E4 wristband user's manual [v1.2]. Retrieved from https://www.empatica.com/manuals/

Epel, E. S., Crosswell, A. D., Mayer, S. E., Prather, A. A., Slavich, G. M., Puterman, E., & Mendes, W. B. (2018). More than a feeling: A unified view of stress measurement for population science. Frontiers in neuroendocrinology, 49, 146-169. doi:10.1016/j.yfrne.2018.03.001

Francesca, S., Carlo, C. G., Luca, F. D. N., & Marco, R. (2018). Comparison of low-complexity algorithms for real-time QRS detection using standard ECG database. *International Journal on Advanced Science, Engineering and Information Technology, 8*(2), 307. ISSN: 2088-5334

Hinde, K., White, G., & Armstrong, N. (2021). Wearable devices suitable for monitoring twenty four hour heart rate variability in military populations. Sensors, 21(4), 1061. https://doi.org/10.3390/s21041061

Indikawati, F. I., & Winiarti, S. (2020). Stress detection from multimodal wearable sensor data. In IOP Conference Series: Materials Science and Engineering (Vol. 771, No. 1, p. 012028). IOP Publishing. doi:10.1088/1757-899X/771/1/012028

Jauniaux, J., Tessier, M. H., Regueiro, S., Chouchou, F., Fortin-Côté, A., & Jackson, P. L. (2020). Emotion regulation of others' positive and negative emotions is related to distinct patterns of heart rate variability and situational empathy. PloS one, 15(12), e0244427. https://doi.org/10.1371/journal.pone.0244427

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. Psychiatry investigation, 15(3), 235. doi: 10.30773/pi.2017.08.17

Kodithuwakku Arachchige, S. N., Burch V, R. F., Chander, H., Turner, A. J., & Knight, A. C. (2022). The use of wearable devices in cognitive fatigue: current trends and future intentions. Theoretical Issues in Ergonomics Science, 23(3) 374-386. doi:10.1080/1463922X.2021.1965670

Krejčí, J., Botek, M., & McKune, A. J. (2018). Stabilization period before capturing an ultra-short vagal index can be shortened to 60 s in endurance athletes and to 90 s in university students. PloS one, 13(10), e0205115. https://doi.org/10.1371/journal.pone.0205115

Kruger, C. J. C. (2003). Constrained cubic spline interpolation. *Chemical Engineering Applications, 1*(1).

Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting. Frontiers in psychology, 8, 213. https://doi.org/10.3389/fpsyg.2017.00213

Li, S., Sung, B., Lin, Y., & Mitas, O. (2022). Electrodermal activity measure: A methodological review. Annals of Tourism Research, 96, 103460. https://doi.org/10.1016/j.annals.2022.103460

Liang, Y., Elgendi, M., Chen, Z., & Ward, R. (2018). An optimal filter for short photoplethysmogram signals. *Scientific data, 5*(1), 1-12. doi: 10.1038/sdata.2018.76

Madona, P., Basti, R. I., & Zain, M. M. (2021). PQRST wave detection on ECG signals. Gaceta Sanitaria, 35, S364-S369. https://doi.org/10.1016/j.gaceta.2021.10.052

Marowka, A. (2018). Python accelerators for high-performance computing. *The Journal of Supercomputing, 74*(4), 1449-1460. https://doi.org/10.1007/s11227-017-2213-5

Martínez, C. A. G., Quintana, A. O., Vila, X. A., Touriño, M. J. L., Rodríguez-Liñares, L., Presedo, J. M. R., & Penín, A. J. M. (2017). Heart rate variability analysis with the R package RHRV. Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-65355-6

Marques, A. H., Silverman, M. N., & Sternberg, E. M. (2010). Evaluation of stress systems by applying noninvasive methodologies: measurements of neuroimmune biomarkers in the sweat, heart rate variability and salivary cortisol. *Neuroimmunomodulation*, *17*(3), 205-208. https://doi.org/10.1159/000258725

Muse. (2022, Sept 6). Muse Products Technical Specifications. Retrieved on 28 Sept 2022 from: https://choosemuse.force.com/s/article/Muse-Products-Technical-Specifications?language=en_US

North, R. P., & Livingstone, D. M. (2013). Comparison of linear and cubic spline methods of interpolating lake water column profiles. *Limnology and Oceanography: Methods, 11*(4), 213-224. https://doi.org/10.4319/lom.2013.11.213

Page, A., Hassanalieragh, M., Soyata, T., Aktas, M. K., Kantarci, B., & Andreescu, S. (2017). Conceptualizing a real-time remote cardiac health monitoring system. In Medical Imaging:

Concepts, Methodologies, Tools, and Applications (pp. 160-193). IGI Global. doi: 10.4018/978-1-4666-8662-5.ch001

Paloniemi, S., Penttonen, M., Eteläpelto, A., Hökkä, P., & Vähäsantanen, K. (2022). Integrating self-reports and electrodermal activity (EDA) measurement in studying emotions in professional learning. In Methods for Researching Professional Learning and Development (pp. 87-109). Springer, Cham. https://doi.org/10.1007/978-3-031-08518-5_5

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

Schmettow, M. (2023). YLab [Python software]. Retrieved on 24th of October from https://github.com/schmettow/YLab.git

Segerstrom, S.C., & Miller, G.E. (2004). Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry. Psychological bulletin, 130(4), 601-630. https://dx.doi.org/10.1037/0033-2909.130.4.601

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Frontiers in public health, 258. https://doi.org/10.3389/fpubh.2017.00258

Shankar, M. G., & Babu, C. G. (2020). An exploration of ECG signal feature selection and classification using machine learning techniques. *Int. J. Innovative Technol. Exploring Eng. Regul, 9*(3), 797-804. DOI:10.35940/ijitee.C8728.019320

Shei, R. J., Holder, I. G., Oumsang, A. S., Paris, B. A., & Paris, H. L. (2022). Wearable activity trackers–advanced technology or advanced marketing?. European Journal of Applied Physiology, 1-16. https://doi.org/10.1007/s00421-022-04951-1

Torkamani-Azar, M., Lee, A., & Bednarik, R. (2022). Methods and measures for mental stress assessment in surgery: A systematic review of 20 years of literature. IEEE Journal of Biomedical and Health Informatics, , 1-14. doi:10.1109/JBHI.2022.3182869

Tsao, L., Li, L., & Ma, L. (2019). Human work and status evaluation based on wearable sensors in human factors and ergonomics: A review. IEEE Transactions on Human-Machine Systems, 49(1), 72-84. doi:10.1109/THMS.2018.2878824

Van der Mee, D. J., Duivestein, Q., Gevonden, M. J., Westerink, J. H. D. M., & de Geus, E. J. C. (2020). The short Sing-a-Song Stress Test: A practical and valid test of autonomic responses induced by social-evaluative stress. Autonomic Neuroscience, 224, 102612. https://doi.org/10.1016/j.autneu.2019.102612

White, M. T. (n.d.) The Need for Speed: Julia Vs. Python. Retrieved from https://www.researchgate.net/profile/Mason-White-5/publication/357825090_The_Need_for_Speed_Julia_Vs_Python/links/61e109f4c5e3103375918089/The-Need-for-Speed-Julia-Vs-Python.pdf

Wójcikowski, M., & Pankiewicz, B. (2020). Photoplethysmographic time-domain heart rate measurement algorithm for resource-constrained wearable devices and its implementation. Sensors, 20(6), 1783.https://doi.org/10.3390/s20061783

Appendix A

Sensor Placement Manual

### How to attach the Empatica E4 on your right wrist*

Wear the E4 with the case on top of your wrist. Wear it snugly, so that it does not move around, but not so tight that it is uncomfortable. The EDA electrodes (under the snaps) should be on the inside of the wrist. You may optionally line them up with a finger, e.g. the third (ring) finger, but this is not required.
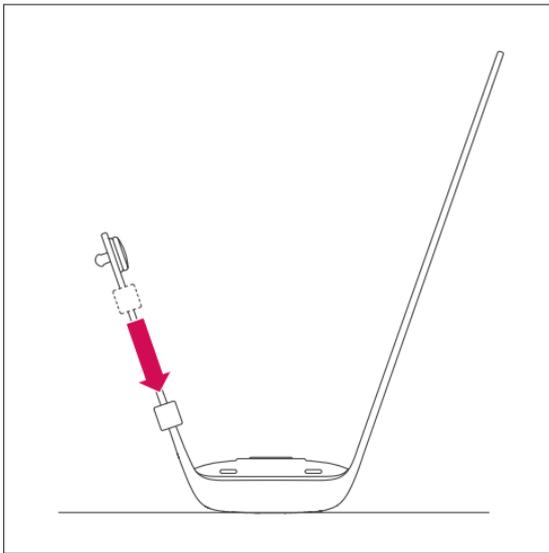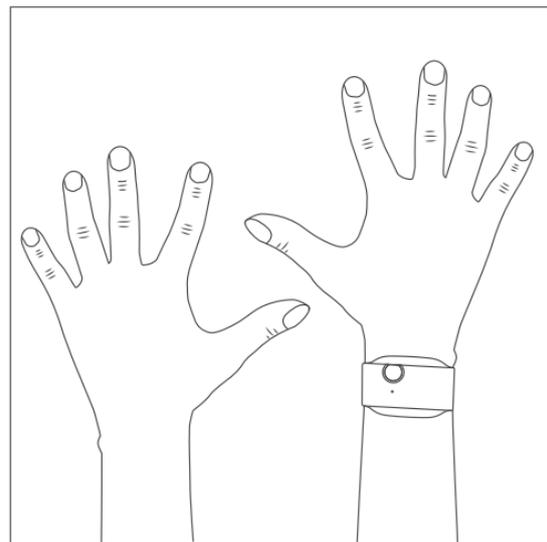

1 - Slide loop towards the case.
2 - Place the E4 top-down on a surface.
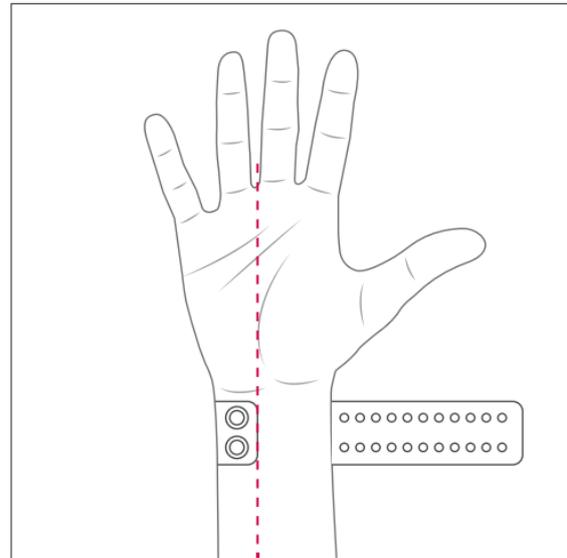
3 - Wrap the band around the wrist.
4 - Line up the snaps under the middle and ring fingers.
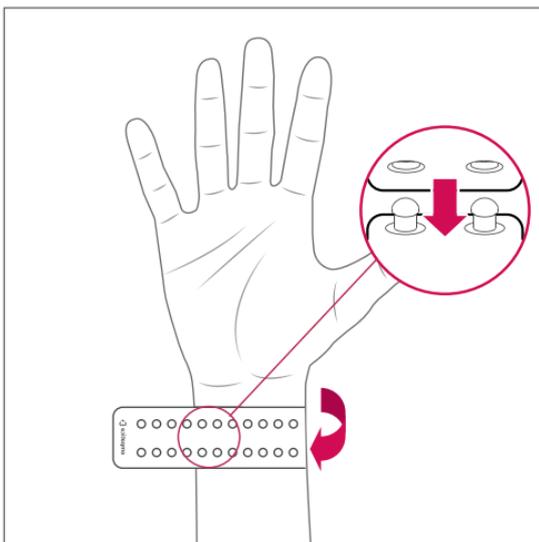
5 - Wrap the band over snaps and tighten.
6 - To secure, connect one snap at a time.
7 - If too tight, loosen by one snap.

* This Manual is derived from the Empatica user manual.

How to attach the 3 ECG electrodes to your torso

Please attach the electrodes matching the colours White, Black, and Red on your torso as shown in the picture. If you went to the bathroom to place the electrode, please return to the experiment room when you are done.

When placing the stickers with the White and Black cord attached, try to place them on the same height/on a horizontal line approximately 2 cm below your collarbones on each side respectively.

- Place the white electrode under your collarbone on your right side.
- Place the black electrode under your collarbone on the left side.
- Place the red electrode at the lower part of your left ribcage.
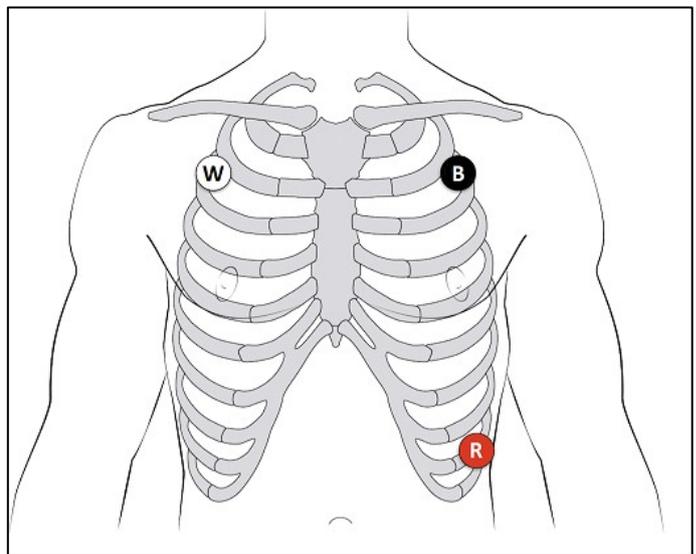
How to attach the finger sleeves to your fingers

Put on the finger sleeves with the metal touching the middle part between the two joints on your index and middle finger on the left palm.

The sleeve with the red cable should be on your index finger and the sleeve with the black cable should be on your middle finger.

Where the metal should touch your fingers →

*
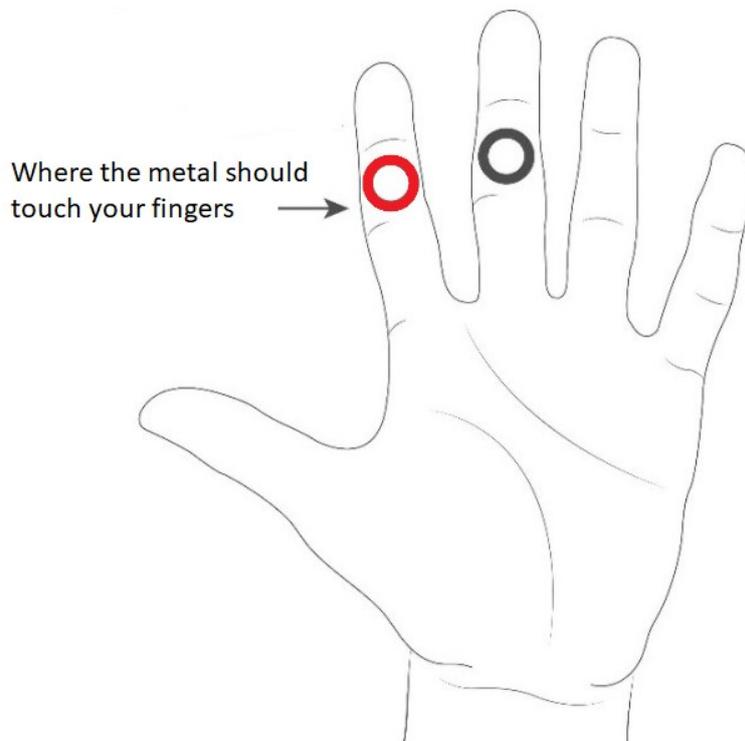
Appendix B

Experiment Sequence: Baseline and Adjusted Sing-a-Song Stress Task (SSST)

Slide 1:

Please sit as still as possible until you see "End task"

All slides will move on automatically please be patient

Slide 2:

How stressed do you feel due to the experiment?

Likert scale:
1. Not at all stressed
2. Low stress
3. Slightly stressed
4. Moderately stressed
5. Very stressed
6. Highly stressed
7. Extremely stressed

Slide 3:

Baseline measurement

Slide 4:

How stressed do you feel due to the experiment?

Likert scale:
1. Not at all stressed
2. Low stress
3. Slightly stressed
4. Moderately stressed
5. Very stressed
6. Highly stressed
7. Extremely stressed

Slide 5:

Please follow the instructions on the screen

Example sequence:

Instruction: read only
"example sentence"

01:00

Instruction: read aloud
"example sentence"

Read this in your head   Read this out loud

Slide 6:

Please indicate how stressed you feel due to the experiment?

1 – Not at all stressed
2 – Low stress
3 – Slightly stressed
4 – Moderately stressed
5 – Very stressed
6 – Highly stressed
7 – Extremely stressed

01:00

Slide 7:

Instruction: read only

"A **vacuum cleaner** is also known simply as a **vacuum** or a **hoover**"
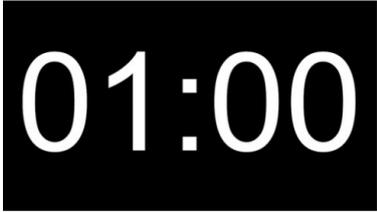
Slide 8:

01:00

Instruction: read only

"A **vacuum cleaner** is generally electrically driven, often via a cable."

9

Please indicate how stressed you feel due to the experiment?
1 – Not at all stressed
2 – Low stress
3 – Slightly stressed
4 – Moderately stressed
5 – Very stressed
6 – Highly stressed
7 – Extremely stressed

01:00

10

Instruction: read aloud

"A **vacuum cleaner** is generally electrically driven, often via a cable."

11

01:00

12

Instruction: read only

"The vacuum cleaner evolved from the carpet sweeper via manual vacuum cleaners."

13

Please indicate how stressed you feel due to the experiment?
1 – Not at all stressed
2 – Low stress
3 – Slightly stressed
4 – Moderately stressed
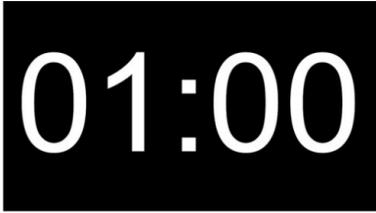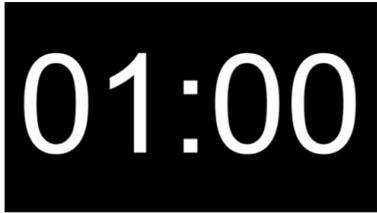5 – Very stressed
6 – Highly stressed
7 – Extremely stressed

01:00

14

Instruction: prepare

Think of a song you could sing for 30 seconds

15

01:00

16

Instruction: sing aloud

Sing the song you thought of aloud until the next counter begins

17

Please indicate how stressed you
feel due to the experiment?

1 – Not at all stressed
2 – Low stress
3 – Slightly stressed
4 – Moderately stressed
5 – Very stressed
6 – Highly stressed
7 – Extremely stressed

01:00

18

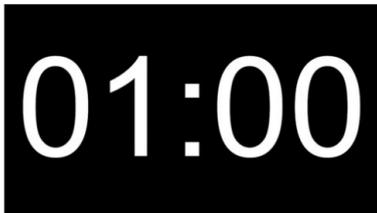This was an adjusted Sing-a-Song stress task (SSST)

END TASK

19

Appendix C

Python Code YLab Prototype

```python
"""
YLab1(Ylab0): YEDA and MOI
"""

import board
import time

from sensory import Yeda, Yema_ads, MOI, Sensory from yui import Button, RGB
from ydata import SDcard

def main():
    STATE = "Init"
    print(STATE)

    sensory = Sensory([MOI(pins = board.GP21),
            MOI(pins = board.GP22),
            Yeda(),
            Yema_ads()])
    sensory.connect()

    SDcard.init()
    drive = SDcard(sensory, filename = "ylab1_" + str(time.time()) + ".csv")
    drive.connect()

    btn = Button()
    btn.connect()

    rgb = RGB()
    rgb.connect()
    rgb.white()

    ## Fast while loop
    while True:
        ################ Interactive transitionals #############
        if btn.update():
```

```
if btn.update_event():
    if btn.event == "short":
        ## Init --> Pause
        if STATE == "Init":
            STATE = "Pause"
        ## Record --> Pause
        elif STATE == "Record":
            STATE = "Pause"
        ## Pause --> Record
        elif STATE == "Pause":
            STATE = "Record"
        ## Stop --> Record
        elif STATE == "Stop":
            sensory.reset_data()
            drive.filename = "ylab1_" + str(time.time()) + ".csv"
            STATE = "Record"
        ## --> STOP
    elif btn.event == "long":
        if STATE == "Stop":
            drive.disconnect()
            rgb.off()
            STATE = "End"
            print("YLab0 says bye.")
            break
        else:
            STATE = "Stop"


    ## Updating the static displays ##
    if STATE == "Record":
        rgb.red()
    elif STATE == "Pause":
        rgb.green()
    elif STATE == "Stop":
        rgb.white()


    print(STATE)
```

```
############### Continuous processing #############


if STATE == "Record":
    if sensory.sample():
        sensory.print()
        sensory.record()
    drive.update()
elif STATE == "Pause":
    if sensory.sample():
        sensory.print()
elif STATE == "Stop":
    pass
else:
    pass


main()
```

Appendix D

R Script - YLab Prototype Yeda and Yema Sensors (pre-)processing

YLab validation study data transformations

I.G.M. Jongmans

2023-31-01

YLab validation study '22-'23

The following script documents the data (pre-)processing, exploration and analysis necessary for the YLab validation study. It was also written with future research in mind wherein the current time synchronization issue is no longer present. The time synchronization and inserting of the subjective stress scale values for the current study was done in Excel.

```r
#Load packages
library(tidyverse)
library(readr)
library(ggplot2)
library(dplyr)
library(lme4)
library(zoo)


library(devtools)
library(heartBeat)


setwd("~/2022-2023/Masterthese/Analysis")
```

1. Data pre-processing

Data importation

The raw data of the YLab is imported. Simultaneously the arbitrary time is reduced by the corresponding value per participant to create the Real Time (RT). Furthermore, the minimum and maximum inter beat interval (IBI) values from the E4 are added.

```r
# Import all raw data for the YLab
# Add participant number and calculate real time (RT). Manually look
 up the time that should be deducted from the time column per partic
ipant.
# It would be best to figure out how to have R auto select the right
 arbitrary time from the first observation in the time column per pa
rticipant.
P1 <- read.csv(file = "ylab1_1668879901.csv", sep = ",", header = TR
UE)
P2 <- read.csv(file = "ylab1_1669484532.csv", sep = ",", header = TR
UE)
P3 <- read.csv(file = "ylab1_1669557636.csv", sep = ",", header = TR
UE)
P4 <- read.csv(file = "ylab1_1669644271.csv", sep = ",", header = TR
UE)
P5 <- read.csv(file = "ylab1_1669651688.csv", sep = ",", header = TR
UE)
P6 <- read.csv(file = "ylab1_1669721446.csv", sep = ",", header = TR
UE)
P7 <- read.csv(file = "ylab1_1669738294.csv", sep = ",", header = TR
UE)
P8 <- read.csv(file = "ylab1_1669811947.csv", sep = ",", header = TR
UE)
P9 <- read.csv(file = "ylab1_1669815456.csv", sep = ",", header = TR
UE)
P10 <- read.csv(file = "ylab1_1670319435.csv", sep = ",", header = T
RUE)
P11 <- read.csv(file = "ylab1_1670330430.csv", sep = ",", header = T
RUE)
P12 <- read.csv(file = "ylab1_1670333844.csv", sep = ",", header = T
RUE)

#add in values including the min and max values E4 necessary later o
```

```r
n
P1 <- P1 %>% mutate(part = 1, RT = time - 3422, E4min = .58, E4max =
 .84)
P2 <- P2 %>% mutate(part = 2, RT = time - 45, E4min = .53, E4max = .
95)
P3 <- P3 %>% mutate(part = 3, RT = time - 153, E4min = .55, E4max =
1.02)
P4 <- P4 %>% mutate(part = 4, RT = time - 48, E4min = .59, E4max = 1
.14)
P5 <- P5 %>% mutate(part = 5, RT = time - 39, E4min = .61, E4max = 1
.08)
P6 <- P6 %>% mutate(part = 6, RT = time - 919, E4min = .53, E4max =
1.08)
P7 <- P7 %>% mutate(part = 7, RT = time - 646, E4min = .53, E4max =
.86)
P8 <- P8 %>% mutate(part = 8, RT = time - 1006, E4min = .53, E4max =
 .89)
P9 <- P9 %>% mutate(part = 9, RT = time - 4231, E4min = .56, E4max =
 1.05)
P10 <- P10 %>% mutate(part = 10, RT = time - 2188, E4min = .52, E4ma
x = 1.05)
P11 <- P11 %>% mutate(part = 11, RT = time - 3469, E4min = .58, E4ma
x = 1.11)
P12 <- P12 %>% mutate(part = 12, RT = time - 710, E4min = .63, E4max
 = 1.06)


# Create long data frame for all participants
YLAB <- rbind(P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11, P12)
```

HR visualizations

To check if the yema signal shows signs of QRS complexes the data is plotted.

```r
#create a YLab data object with only the Yema_ads0 input
YEMA <- filter(YLAB, ID == "Yema_ads0")
```

```r
#plot a segment of yema values for each participant
#for the xlim values between 0 and approximately 900 should work how
ever as all experiments lasted at least that long
YEMA %>%
    ggplot(aes(x = RT,
            y = value)) +
  facet_grid(ID ~ 1, scales = "free_y") +
  geom_line() + xlim(620, 624) + ylim(0.32, 0.5) + xlab("Real time e
xperiment in seconds") + ylab("YLab ECG Values") + facet_wrap(~YEMA$
part)


#per participant check
YEMA %>%  filter(part == 1) %>%
            ggplot(aes(x = RT, y = value)) +
  facet_grid(scales = "free_y") +
  geom_line() + xlim(620, 630) + ylim(0.32, 0.5) + xlab("Real time e
xperiment in seconds") + ylab("YLab ECG Values")



#The signal looks like it contains QRS complexes with distinguishabl
e R peaks. Therefore, it was deemed good enough to continue towards
data (pre-) processing
```

HR and EDA transformations

The YLab data needed to be transformed and pre-processed before it would be useful to analyze. For the Hr transformations the YLab ECG sensor is referred to as yema, for the EDA sensor it will be denoted as yeda.

HR transformations

Four ways of pre-processing the yema data will be compared based on correlation with the E4 data on acquiring the heart rate (HR) from and pre-processing the yema data.

Pre-processing the yema data with an existing package 0) use heartBeats from (https://github.com/ehrscape/R-project/blob/master/HeartBeat-walkthrough.md) to detect heartbeats and then HR

Pre-processing the yema data step by step and then 1) filtering with yema IBI outliers to filter the yema IBI 2) filtering based on IBI values and algorithm from the E4 (https://support.empatica.com/hc/en-us/articles/360030058011-E4-data-IBI-expected-signal) 3) applying linear interpolation to estimate the missing inter beat intervals (IBIs)

The HeartBeat package

The yema data is similar to filtered data from the literature so it should probably not be filtered with the usual frequency filters. However, for this package the instructions were followed to pass the data through the wavelet decomposition transformations.

```r
# transform df to necessary df for package functions
# Try per participant
HR0.1 <- YEMA %>% filter(part == 1) %>% select(RT, value)
# convert time to unix timestamp by adding a random unix timestamp to the RT
# the unix from Monday the 2nd of January at 12:00 = 1672657200
HR0.1 <- HR0.1 %>% mutate(time = format(round(RT + 1672657200, 2), nsmall = 2))
HR0.1 <- HR0.1 %>% transmute(time = time, ecg = value)
options(digits=12)
HR0.1 <-transform(HR0.1, time = as.numeric(time),
ecg = as.numeric(ecg))


#adjust to YLab settings
SampleFreq = 44 # 100Hz adjusted to the recorded amount of observations (around 42-46 obs per second)
thr = 0.009 # initial threshold of 9 / 1000 to match the yema values



# 4-level decomposition is used with the Daubechie d4 wavelet. (as in the package example)
```

```r
wavelet <- "d4"
level <- 4L


X <- as.numeric(HR1.1$ecg)


library(wavelets)
ecg_wav <- dwt(X, filter = wavelet, n.levels = level, boundary = "periodic",
    fast = TRUE)
str(ecg_wav)


# Coefficients of the second level of decomposition are used for R peak detection.
x <- ecg_wav@W$W2


#plot the coefficients
plot(x, type = "l")
plot(x)


# after the visualizations a more appropriate threshold seems to be .02
thr = .02


#plot x again, but zoom to get a grasp on what the yema data produces thus far
#to do so the wavelet coefficients of x are made into a df and prepared for ggplot
xx <- as.data.frame(x) #make df from the object x
xx <- add_column(xx, indx = NA) #add column for row names
xx <- xx %>%  mutate(indx = row_number())


#adjust xlim to scan the wavelet coefficients wherever you want
ggplot(data = xx, aes(x = indx, y = xx$V1)) + geom_line() + xlim(140
```

```
0, 1500) + xlab("observation count") + ylab("wavelet coefficients fr
om x")


# Unfortunately, the wavelet coefficients seem to be irregular and t
his could be due to the sampling rate the yema currently has. Due to
 the low sampling rate the signal clarity for various participants i
s not enough for the data processing of the hearBeat package.
# From here on, someone could continue with the process if the sampl
ing rate is increased to match the example data to a higher degree.
```

The heartBeat package does not seem to work on the yema data yet, so it will be skipped for now. This leaves the 3 other options of further pre-processing the data, the filtering based on yema outliers, E4 data and applying linear interpolation. The rest of the heartBeat package code would be to continue at the "Peak detection" chapter, except that [x <- ecg_wav@W$W2] is already performed in the R chunk above.

Pre-processing the yema data step by step

The following steps were chosen to extract the IBI and HR from the raw yema sensor data.

Basic Filter round 1

The threshold of .39 and the derivative of .01 were chosen based on outliers and visual examination. Subsequently the data is filtered with those values as cut off points.

```
#find filter based on the height of the peaks in the yema data
#after the visual inspection a threshold of around .4 should work fo
r most peaks
#however a check was done based on outliers in the yema data
YEMA_out <- c(boxplot.stats(YEMA$value, do.out = T))


summary(YEMA_out$out)
#the min value of .0947 seems like an error based on the visual insp
ection
#the outliers will have to be above .33 to find a suitable min value
summary(YEMA_out$out[YEMA_out$out >= .33])
```

```r
#select threshold for value filter
value_t <- .39


#find filter based on the calculated derivative
YEMA <- YEMA %>%  group_by(part) %>% mutate(deriv = value-lag(value,
 default = first(value)))


#find the positive (> 0) outliers in the derivatives that are part o
f the rising R peak
deriv_out <- c(boxplot.stats(YEMA$deriv, do.out = T))


summary(deriv_out$out[deriv_out$out >= 0])


#select threshold for derivative filter
deriv_t <- .01


#value filter
YEMA <- YEMA %>%
  mutate(value_filter = if_else(condition = value > value_t,
                        true = 1,
                        false = 0), .after = value)


#derivative filter
YEMA <- YEMA %>%
  mutate(deriv_filter = if_else(condition = deriv > deriv_t,
                        true = 1,
                        false = 0), .after = deriv)


YEMA %>% count()
```

Basic Filter round 2

The next filter is based on the length of the Inter Beat Interval (IBI). Based on an examination of the outliers the cut off point is determined at being longer than .4 seconds. Then the Heart Rate (HR) is calculated. These are the final steps of approach 1.

```r
#take the values that met the criteria from filter round 1 and put them in the inter beat interval (IBI) df
IBI <- dplyr::filter(YEMA, value_filter == 1 & deriv_filter == 1)


#calculate the IBI based on the differences between the time stamps of the selected peaks
IBI <- IBI %>%  group_by(part) %>% mutate(IBI = RT-lag(RT, default = first(RT)))


#find the positive (> 0) outliers in the derivatives for the derivative filter
IBI_out <- c(boxplot.stats(IBI$IBI, do.out = T))


summary(IBI_out$out)
#hist(IBI_out$out)
#from the lower end outliers anything under .4 will be filtered out
#from the upper end a more narrowed look might be better
summary(IBI_out$out[IBI_out$out <= .5716]) #lower end
summary(IBI_out$out[IBI_out$out >= .5716]) #upper end


#select threshold for IBI filters
IBI_t1 <- .4 #lower end also due to respective HR of 150 (based on upper spectrum max HR)
IBI_t2 <- 1.2 #upper end also due to respective HR of 50 (based on lower spectrum resting HR)



#filter IBI
IBI <- IBI %>% mutate(IBI_filter = if_else(condition = IBI > IBI_t1 & IBI < IBI_t2,
```

```
                              true = 1,

                              false = 0), .after = IBI)


#take the filtered IBI values to calculate the HR
HR <- dplyr::filter(IBI, IBI_filter == 1)


HR <- HR %>% mutate(HR = 60/IBI)


HR %>% group_by(part) %>% count()
#the HR for the yema sensor of the YLab prototype is now computed an
d can be used in statistical tests
#However, some HR are lower because no R peaks were present based on
 the current filters
#after checking the data some segments seem to miss R peaks, which w
as also confirmed by visualizations.
#it appeared the data needed some further processing hence the next
approaches, filtering based on E4 values and linear interpolation,
```

Further filtering yema data based on E4 values

In approach 2 the goal is to filter out both the short and long IBIs based on the min/max E4
IBIs, respectively. The Empatica uses an algorithm to determine the difference between true
heartbeats and false beats. From the accepted beats the IBI and the HR are calculated. When
using the IBI to synchronize the YLab and E4 data the false beats are left out of the E4 data
thus to create a similar effect in the YLab data it is filtered further based on the min and max
IBI lengths of the E4.

```
#to clean up the df keep columns relevant to synchronization later o
n
#filter out yema IBI outside boundaries min and max E4 IBI
HR2 <- IBI %>% group_by(part) %>%  transmute(IBI = ifelse(IBI < E4mi
n | IBI > E4max, NA, IBI), RT, E4min, E4max)


#calculate HR based on accepted IBIs
HR2 <- HR2 %>%  group_by(part) %>%  mutate(HR = 60/IBI)
```

```
#filter out NAs to obtain new HR list to be synchronized to the E4 l
ist
HR2 <- na.omit(HR2)


HR2 %>%  group_by(part) %>% count()
#the accepted data is reduced greatly by this procedure and might no
t be the best way to go
```

Linear interpolation for IBI

In approach 3 the goal is to identify where beats were missed based on too long IBIs. Calculate the amount of missed beats to be inserted. Prepare the rows associated with missed beat to have NA, then interpolate.

```
#to clean up the df keep columns relevant to synchronization later o
n
HR3 <- IBI %>% select(part, RT, IBI, E4min, E4max)


#filter out erroneous beats with the lower end IBI threshold
HR3 <- HR3 %>% group_by(part) %>%  transmute(IBI = ifelse(IBI < IBI_
t1, NA, IBI), RT, E4min, E4max)
HR3 <- na.omit(HR3)


#find IBIs that are too long based on the upper end IBI threshold
#use median per participant to determine amount of missed beats
#the row containing the too long IBI will be interpolated later thus
 misses -1
HR3 <- HR3 %>% group_by(part) %>% mutate(misses = if_else(condition
= IBI > IBI_t2, round(IBI / median(IBI) -1), 0), .after = IBI)


#change long IBIs to NA
HR3 <- HR3 %>% group_by(part) %>% mutate(IBI = ifelse(IBI > IBI_t2,
NA, IBI))
```

```r
#add rows based on amount of misses and copy all other columns relev
ant to synchronization
HR3 <- HR3 %>%
  rowwise() %>%
  do(if(.$misses > 0) {
    data.frame(part = c(rep(.$part, .$misses), .$part),
               RT = c(rep(.$RT, .$misses), .$RT),
               IBI = c(rep(NA, .$misses), .$IBI),
               misses = c(rep(.$misses, .$misses), .$misses),
               E4min = c(rep(.$E4min, .$misses), .$E4min),
               E4max = c(rep(.$E4max, . $misses), .$E4max))
  } else if (.$misses == 0) {
    data.frame(part = .$part,
               RT = .$RT,
               IBI = .$IBI,
               misses = .$misses,
               E4min = .$E4min,
               E4max = .$E4max)
  } else {
  }) %>%
  ungroup()
```

Interpolate the inserted rows

```r
#linear interpolation for the IBI values
HR3 <- na.approx(HR3)
HR3 <- data.frame(HR3)


#make the affected rows easier to read
HR3 <- HR3 %>%  mutate(IBI = round(IBI, digits = 3))


#calculate HR based with interpolated IBI
HR3 <- HR3 %>% mutate(HR = 60/IBI)
HR3 %>%  group_by(part) %>% count()
```

E4 linear interpolation

After interpolating the yema data the interpolation steps will also be performed on the E4 data.

Import E4 data

```
#Import and change data accordingly
PE1 <- read.csv(file = "IBI_P1.csv", sep = ",", header = TRUE)
PE2 <- read.csv(file = "IBI_P2.csv", sep = ",", header = TRUE)
PE3 <- read.csv(file = "IBI_P3.csv", sep = ",", header = TRUE)
PE4 <- read.csv(file = "IBI_P4.csv", sep = ",", header = TRUE)
PE5 <- read.csv(file = "IBI_P5.csv", sep = ",", header = TRUE)
PE6 <- read.csv(file = "IBI_P6.csv", sep = ",", header = TRUE)
PE7 <- read.csv(file = "IBI_P7.csv", sep = ",", header = TRUE)
PE8 <- read.csv(file = "IBI_P8.csv", sep = ",", header = TRUE)
PE9 <- read.csv(file = "IBI_P9.csv", sep = ",", header = TRUE)
PE10 <- read.csv(file = "IBI_P10.csv", sep = ",", header = TRUE)
PE11 <- read.csv(file = "IBI_P11.csv", sep = ",", header = TRUE)
PE12 <- read.csv(file = "IBI_P12.csv", sep = ",", header = TRUE)


PE1 <- PE1 %>% mutate(part = 1, E4min = .58, E4max = .84)
PE2 <- PE2 %>% mutate(part = 2, E4min = .53, E4max = .95)
PE3 <- PE3 %>% mutate(part = 3, E4min = .55, E4max = 1.02)
PE4 <- PE4 %>% mutate(part = 4, E4min = .59, E4max = 1.14)
PE5 <- PE5 %>% mutate(part = 5, E4min = .61, E4max = 1.08)
PE6 <- PE6 %>% mutate(part = 6, E4min = .53, E4max = 1.08)
PE7 <- PE7 %>% mutate(part = 7, E4min = .53, E4max = .86)
PE8 <- PE8 %>% mutate(part = 8, E4min = .53, E4max = .89)
PE9 <- PE9 %>% mutate(part = 9, E4min = .56, E4max = 1.05)
PE10 <- PE10 %>% mutate(part = 10, E4min = .52, E4max = 1.05)
PE11 <- PE11 %>% mutate(part = 11, E4min = .58, E4max = 1.11)
PE12 <- PE12 %>% mutate(part = 12, E4min = .63, E4max = 1.06)
```

```r
colnames(PE1)[1] <- "RT"
colnames(PE2)[1] <- "RT"
colnames(PE3)[1] <- "RT"
colnames(PE4)[1] <- "RT"
colnames(PE5)[1] <- "RT"
colnames(PE6)[1] <- "RT"
colnames(PE7)[1] <- "RT"
colnames(PE8)[1] <- "RT"
colnames(PE9)[1] <- "RT"
colnames(PE10)[1] <- "RT"
colnames(PE11)[1] <- "RT"
colnames(PE12)[1] <- "RT"


E4_ALL <- rbind(PE1, PE2, PE3, PE4, PE5, PE6, PE7, PE8, PE9, PE10, P
E11, PE12)
```

Calculate misses E4 based on derivative

```r
#calculate derivative RT
E4_ALL <- E4_ALL %>%  group_by(part) %>% mutate(misses = RT-lag(RT,
default = first(RT)), .before = RT)


#calculate the amount of misses based on median IBI
#The row the misses were located in remains the same in contrast to
the yema data
#thus -1 is no longer necessary
E4_ALL <- E4_ALL %>% group_by(part) %>% mutate(misses = if_else(cond
ition = misses > IBI, round(misses/median(IBI)), 0))


#sum amount of misses per participant
E4_ALL %>%  group_by(part) %>% summarize(Freq = sum(misses)) #amount
 of misses
```

Insert rows with NA into E4 data

```r
#because the IBI in the row the misses were located is an accepted I
BI by the E4 algorithm it does not need to be corrected
#only the missing rows need to be inserted
E4_ALL <- E4_ALL %>%
  rowwise() %>%
  do(if(.$misses > 0) {
    data.frame(part = c(rep(.$part, .$misses), .$part),
               RT = c(rep(.$RT, .$misses), .$RT),
               IBI = c(rep(NA, .$misses), .$IBI),
               misses = c(rep(.$misses, .$misses), .$misses),
               E4min = c(rep(.$E4min, .$misses), .$E4min),
               E4max = c(rep(.$E4max, . $misses), .$E4max))
  } else if (.$misses == 0) {
    data.frame(part = .$part,
               RT = .$RT,
               IBI = .$IBI,
               misses = .$misses,
               E4min = .$E4min,
               E4max = .$E4max)
  } else {
  }) %>%
  ungroup()


#Use linear interpolation to insert the IBI
E4_ALL <- na.approx(E4_ALL)
E4_ALL <- data.frame(E4_ALL)

#calculate HR based with interpolated IBI
E4_ALL <- E4_ALL %>% mutate(HR = 60/IBI, .after = IBI)
E4_ALL %>%  group_by(part) %>% count()
```

EDA transformation

To match the E4 and the use of Siemens (S) the same unit is calculated for the YLab from Ohm (Ω) with S = 1/Ω. Based on Ohms law, when A remains constant the reciprocal from ohm will be Siemens.

```
#create a YLab Data object with only the Yeda input
YEDA <- filter(YLAB, ID == "Yeda0")


#compute skin conductance in Siemens (S) from the yeda values in Ohm
 (Ω)
YEDA <- YEDA %>% mutate(yeda = 1/value, .after = value)


YEDA
```

Export transformed data to by synchronized in Excel

This is done by performing the following steps per participant: 1. reversing the order of the observations by sorting the RT from high to low 2. copying the E4 data next to it 3. deleting excess E4 data at the bottom 4. sorting the RT from low to high to get the correct order of observations for both the YLab and E4 5. searching for the time stamps in RT to append the subjective stress scale values 6. creating a CSV with all participants below each other (as a long data format) 7. import the synchronized data back into R

```
#convert objects to CSV files to be synchronized in Excel
write_excel_csv(HR, "~/2022-2023/Masterthese/Analysis/HRoutliermin-m
ax.csv", col_names = T)
write_excel_csv(HR2, "~/2022-2023/Masterthese/Analysis/HRE4min-max.c
sv", col_names = T)
write_excel_csv(HR3, "~/2022-2023/Masterthese/Analysis/HRlinear(E4).
csv", col_names = T)
write_excel_csv(E4_ALL, "~/2022-2023/Masterthese/Analysis/E4linear.c
sv", col_names = T)
write_excel_csv(HR3, "~/2022-2023/Masterthese/Analysis/HRlinear(outl
iers).csv", col_names = T)
```

2. Data visualization

plot EDA and HR data from the YLab and E4 per participant

```r
#import synchronized data
# Approach 1 (4 yema base filters)
HR1_ALL <- read.csv(file = "HRlong.csv", sep = ",", header = TRUE)
# Approach 2 (IBI filtered further with E4 min/max)
HR2_ALL <- read.csv(file = "HR2long.csv", sep = ",", header = TRUE)
# Approach 3 (Linear interpolated yema and E4)
HR3_ALL <- read.csv(file = "HR3long.csv", sep = ",", header = TRUE)
# EDA data
EDA_ALL <- read.csv(file = "EDA_long.csv", sep = ",", header = TRUE)
```

plot individual participants per device

```r
#plot excerpts from participants in the last 3 minutes (c.a. second
720 - 900) per device
#this is done for a detailed look at the data
#EDA yema and E4
EDA_ALL %>%
    ggplot(aes(x = RT,
               y = yeda)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("EDA yeda") + facet_wrap(~EDA_ALL$part, scales = "free_y
")


EDA_ALL %>%
    ggplot(aes(x = RT,
               y = E4)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("EDA E4") + facet_wrap(~EDA_ALL$part, scales = "free_y")


#HR approaches yema and E4
#Approach 1
```

```r
HR1_ALL %>%
    ggplot(aes(x = RT,
               y = yema)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR yema") + facet_wrap(~HR1_ALL$part, scales = "free_y"
)


HR1_ALL %>%
    ggplot(aes(x = RT,
               y = E4)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR E4") + facet_wrap(~HR1_ALL$part, scales = "free_y")

#Approach 2
HR2_ALL %>%
    ggplot(aes(x = RT,
               y = yema)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR yema") + facet_wrap(~HR2_ALL$part, scales = "free_y"
)


HR2_ALL %>%
    ggplot(aes(x = RT,
               y = E4)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR E4") + facet_wrap(~HR2_ALL$part, scales = "free_y")

#Approach 3
HR3_ALL %>%
    ggplot(aes(x = RT,
               y = yema)) +
  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR yema") + facet_wrap(~HR2_ALL$part, scales = "free_y"
```

```
)


HR3_ALL %>%

    ggplot(aes(x = RT,

            y = E4)) +

  geom_line() + xlim(720, 900) + xlab("Real time experiment in secon
ds") + ylab("HR E4") + facet_wrap(~HR2_ALL$part, scales = "free_y")
```

plot the data for both devices per participant

```
#Change format to use plotting function later on
#EDA
#Select relevant columns
EDA_ALLx <- EDA_ALL %>% select(part, RT, E4, sub)
EDA_ALLy <- EDA_ALL %>% select(part, RT, yeda, sub)


#transform data
EDA_ALLx <- EDA_ALLx %>% transmute(device = "E4", value = E4, part,
RT, sub)
EDA_ALLy <- EDA_ALLy %>% transmute(device = "yeda", value = yeda, pa
rt, RT, sub)


#complete necessary format
EDA_ALLz <- rbind(EDA_ALLx, EDA_ALLy)


#repeat steps for HR per approach
#Aproach 1
HR1_ALLx <- HR1_ALL %>% select(part, RT, E4, sub)
HR1_ALLy <- HR1_ALL %>% select(part, RT, yema, sub)


HR1_ALLx <- HR1_ALLx %>% transmute(device = "E4", value = E4, part,
RT, sub)
HR1_ALLy <- HR1_ALLy %>% transmute(device = "yema", value = yema, pa
rt, RT, sub)
```

```r
HR1_ALLz <- rbind(HR1_ALLx, HR1_ALLy)


#Aproach 2
HR2_ALLx <- HR2_ALL %>% select(part, RT, E4, sub)
HR2_ALLy <- HR2_ALL %>% select(part, RT, yema, sub)


HR2_ALLx <- HR2_ALLx %>% transmute(device = "E4", value = E4, part,
RT, sub)
HR2_ALLy <- HR2_ALLy %>% transmute(device = "yema", value = yema, pa
rt, RT, sub)


HR2_ALLz <- rbind(HR2_ALLx, HR2_ALLy)


#Approch 3
HR3_ALLx <- HR3_ALL %>% select(part, RT, E4, sub)
HR3_ALLy <- HR3_ALL %>% select(part, RT, yema, sub)


HR3_ALLx <- HR3_ALLx %>% transmute(device = "E4", value = E4, part,
RT, sub)
HR3_ALLy <- HR3_ALLy %>% transmute(device = "yema", value = yema, pa
rt, RT, sub)


HR3_ALLz <- rbind(HR3_ALLx, HR3_ALLy)


# plot data with both devices for EDA and HR approaches
#EDA
ggplot(data = EDA_ALLz,aes(x = RT, y = value)) +
    facet_wrap(~part, scales = "free_y", nrow = 4, ncol = 3) +
    geom_line(aes(colour = device, group = device)) +
    labs(x = NULL, y = NULL, colour = "") +
    theme_bw() + xlim(720, 900) + xlab("Time (secs)") + ylab("EDA")
+   theme(legend.position="bottom")
```

```
#HR
#Approach 1
ggplot(data = HR1_ALLz,aes(x = RT, y = value)) +
    facet_wrap(~part, scales = "free_y", nrow = 4, ncol = 3) +
    geom_line(aes(colour = device, group = device)) +
    labs(x = NULL, y = NULL, colour = "") +
    theme_bw() + xlim(720, 900) + xlab("Time (secs)") + ylab("Heart
Rate (bpm)") + theme(legend.position="bottom")


#Approach 2
ggplot(data = HR2_ALLz,aes(x = RT, y = value)) +
    facet_wrap(~part, scales = "free_y", nrow = 4, ncol = 3) +
    geom_line(aes(colour = device, group = device)) +
    labs(x = NULL, y = NULL, colour = "") +
    theme_bw() + xlim(720, 900) + xlab("Time (secs)") + ylab("Heart
Rate (bpm)") + theme(legend.position="bottom")


#Approach 3
ggplot(data = HR3_ALLz,aes(x = RT, y = value)) +
    facet_wrap(~part, scales = "free_y", nrow = 4, ncol = 3) +
    geom_line(aes(colour = device, group = device)) +
    labs(x = NULL, y = NULL, colour = "") +
    theme_bw() + xlim(720, 900) + xlab("Time (secs)") + ylab("Heart
Rate (bpm)") + theme(legend.position="bottom")
```

visualize correlations between devices

```
#plot EDA for all participants
EDA_ALL %>%
  ggplot(aes(x= yeda, y= E4)) + facet_wrap(~EDA_ALL$part, scales = "
free_y") + geom_point() + geom_smooth(method = "lm") + labs(x= "Yeda
 value in Siemens", y= "E4 values in micro Siemens")


#plot HR approaches
HR1_ALL %>%
```

```
  ggplot(aes(x= yema, y= E4)) + facet_wrap(~HR1_ALL$part) + geom_poi
nt() + geom_smooth(method = "lm") + labs(x= "HR yema filtered with y
ema outliers", y= "HR derived from E4")


HR2_ALL %>%

  ggplot(aes(x= yema, y= E4)) + facet_wrap(~HR2_ALL$part) + geom_poi
nt() + geom_smooth(method = "lm") + labs(x= "HR yema filtered with E
4 min/max", y= "HR derived from E4")


HR3_ALL %>%

  ggplot(aes(x= yema, y= E4)) + facet_wrap(~HR3_ALL$part) + geom_poi
nt() + geom_smooth(method = "lm") + labs(x= "HR yema linear interpol
ation", y= "HR E4 linear interpolation")
```

3. Results (correlations)

Correlate the YLab and E4 data with each other for the EDA and HR. For the HR correlate the 3 approaches: 1. The yema data based on the process of the E4 algorithm 2. The yema data based on the results of the E4 algorithm (E4 min/max IBI values) 3. The linear interpolated data of the yema and E4

```
##See if the EDA values of the E4 and YLab correlate for participant
EDA_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yeda, E4))


#See if the HR values of the E4 and YLab correlate for participant
HR1_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, E4))


HR2_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, E4))
```

```
HR3_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, E4))
```

Correlate the subjective stress scale with each device and sensor per data set

```
##EDA
#correlate subjective stress with Ylab
EDA_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yeda, sub, use = "complete.obs"))


#correlate subjective stress with E4
EDA_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(E4, sub, use = "complete.obs"))


#HR
#Approach 1
#correlate subjective stress with Ylab
HR1_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, sub, use = "complete.obs"))


#correlate subjective stress with E4
HR1_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(E4, sub, use = "complete.obs"))


#Approach 2
#correlate subjective stress with Ylab
HR2_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, sub, use = "complete.obs"))
```

```r
#correlate subjective stress with E4
HR2_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(E4, sub, use = "complete.obs"))



#Approach 3
#correlate subjective stress with Ylab
HR3_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(yema, sub, use = "complete.obs"))


#correlate subjective stress with E4
HR3_ALL %>%
  group_by(part) %>%
  summarize(cor=cor(E4, sub, use = "complete.obs"))
```

Cable test EDA (post hoc)

```r
#test if the cable is broken post hoc
#because the yeda signal seemed to not be working an just picking up
 surrounding frequencies the cable was tested. This was done by maki
ng a recording where 30 sec a normal recording was simulated, 15 sec
 without touching the sensor, 15 sec sensors touching each other, ~1
5 sec again without touching the sensor. This should lead to a chang
e in signal around the timestamps of 30 sec, 45 sec and again at 60
sec. If that does not happen it means the cable is not broken.

T1 <- read.csv(file = "ylab1_cabletest.csv", sep = ",", header = TRU
E)
T1 <- T1 %>% mutate(RT = time - 136)
T1 <- T1 %>% filter(ID == "Yeda0")


T1 %>% ggplot(aes(x= RT, y=value)) + geom_line() + labs(x= "Real tim
```

```
e broken cable test", y= "Yeda signal")


#The cable does not produce an appropriate signal when the electrode
s touch, therefore the cable is deemed broken.
```