Detecting Stress Using Wearables

Matty Pahren, Scott Heng On, Ethan Shen

1 Introduction

Affect recognition seeks to detect a person's affective state. One of these states is stress, and long-term stress is known to have severe implications on human health and well-being. This study uses the Wearable Stress and Affect Detection (WESAD) dataset, which was introduced in Schmidt et al. [1], and achieves classification accuracies of 86.2% on wrist data and 90.8% on all data for the binary classification problem (stress vs. amusement). In conducting this analysis, we aim to determine whether sensor data are useful in predicting stress, and if so, which predictors are most significant when discriminating between stress and amusement. Furthermore, we seek to understand which types of sensor data are most useful in predicting stress, alone or in combination, and to determine if stress can be detected only using the wrist-worn wearable, which is more convenient to wear than a chest-worn device. Finally, we quantify the heterogeneity across individuals in their responses to stress.

Section 2 focuses on providing descriptions of the WESAD data, how the raw data was processed and the engineering of various features. Exploratory data analysis plots are included to provide initial insights to our study based on the data. In section 3, we provide an extensive literature review on previous research, put our research in context of them. Section 4 describes our data and domain-driven approaches to feature engineering, model selection and type of evaluation metrics based on the nature of our study and the dataset. Section 5 discusses the results of our models, predictive accuracies of models being trained only on wrist data and on wrist and chest data combined, along with multiple diagnostic techniques to address model assumptions and check model fit. Section 6 consists of interpretations from the study's models and addressing issues on heterogeneity, collinearity and the predictive significance of various features.

2 Data

The WESAD dataset contains 63 million samples of raw physiological and motion signal data collected from 15 test subjects [1]. There were originally 17 participants in the study, but data for two subjects had to be thrown out. The data was collected using 2 devices: a chest-worn device (RespiBAN) and a wrist-worn device (Empatica E4). The RespiBAN chest data includes electrocardiography (ECG) data in mV, electrodermal activity (EDA) in microseconds, electromyography (EMG) data in mV, skin temperature (TEMP) in °C, displacement of the thorax for males and the abdomen for females induced by inhaling or exhaling (RESP) as a percentage, and 3-axis accelerometer (ACC) data in g - the acceleration of gravity. All RespiBAN data were recorded at 700 Hz. The E4 wrist data collected includes blood volume pulse (BVP) data (64Hz), ACC data in units of 1/64g (32Hz), EDA in microseconds (4Hz), and TEMP in °C (4Hz). Each sample also references a label corresponding to the study protocol condition, with each label referring to a different condition (1-baseline, 2-stress, 3-amusement, 4-meditation). These conditions were recorded by having participants exposed to sources of stimulation that would elicit the desired condition. The labels 0, 5, 6, 7 indicate that the data should be ignored. The dataset also included information about each subject, such as height, weight and age, as well as personal responses to PANAS and SSSQ questionnaires.

Considering the nature of the WESAD dataset and the scope of our study, we decided to filter out samples that did not represent stress or amusement, and standardized sampling rates for all RespiBAN and E4 variables to 4Hz. We did this by downsampling data collected at a more frequent rate, and we chose to use the recorded signals from these time periods instead of a mean since we planned on calculating mean windows of our data later. This downsampling allowed for more feasible computation with a dataset of reasonable size and standardized windows for multiple variables. Even though we lost some information by downsampling, we thought this was a good trade-off so we could test more models with a smaller dataset.

For example, Figure 1.1 and 1.2 show plots of the EDA signal data by participant and by condition showing different values when the participant experienced different conditions, and this difference is captured both in the chest and wrist data. Since we only took samples for 2 conditions, we could desirably observe a relative equal distribution of samples in both groups. However, raw signal data only provides limited information and difficult interpretability, and therefore from these raw data signals, we performed data and domain-driven feature engineering to generate potential predictor variables for our model that were easily interpretable and well represented the information in the data. We generated a total of 50 features across all the raw signal variables.

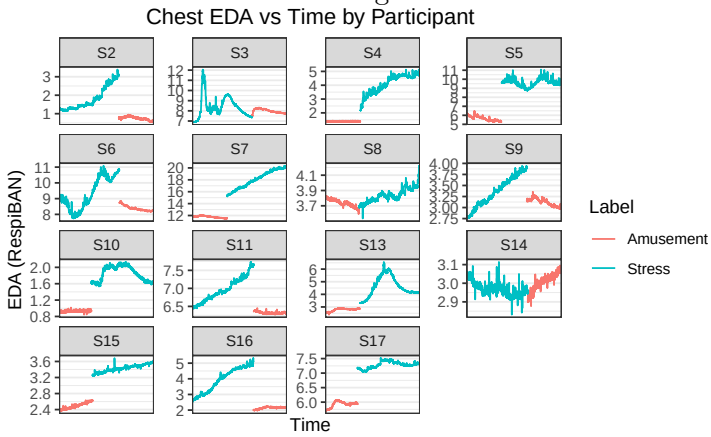

Fig 1.1 Plot of RespiBAN (Chest) EDA samples collected over time by participant and label (2−stress, 3−amusement)

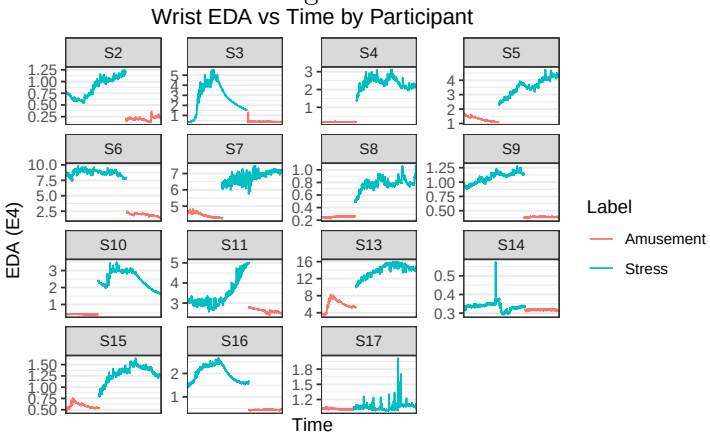

Fig 1.2 Plot of E4 (Wrist) EDA samples collected over time by participant and label (2−stress, 3−amusement)

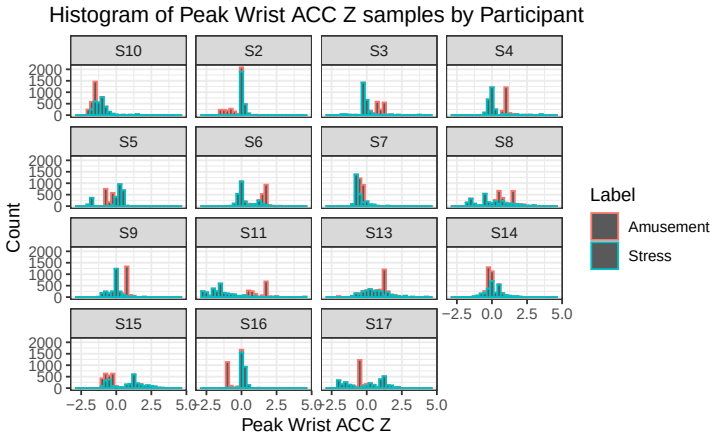

Fig 1.3 Histogram of peak acc wrist samples by participant and label (2−stress, 3−amusement)

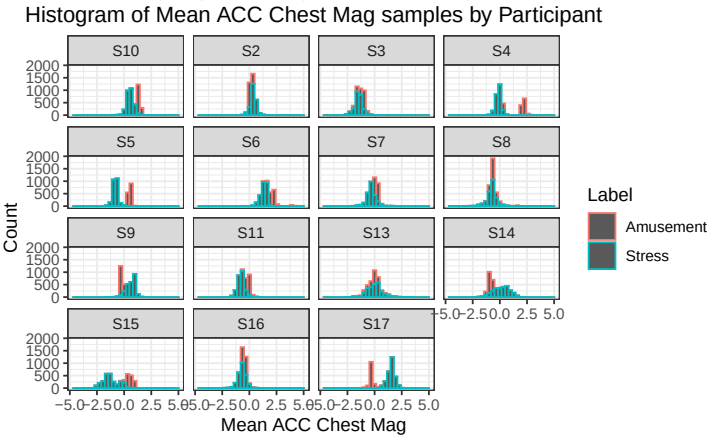

Fig 1.4 Histogram of mean acc chest magnitude samples by participant and label (2−stress, 3−amusement)

Performing exploratory data analysis on the engineered features provided more insightful information related to our study. For example, Figure 1.3 and 1.4 are histograms of Peak Wrist ACC Z and Mean ACC Chest Magnitude samples engineering from the raw signal ACC data. From these plots we can observe that these ACC feature values have different distributions based on the different conditions and could potentially be significant in predicting stress or amusement.

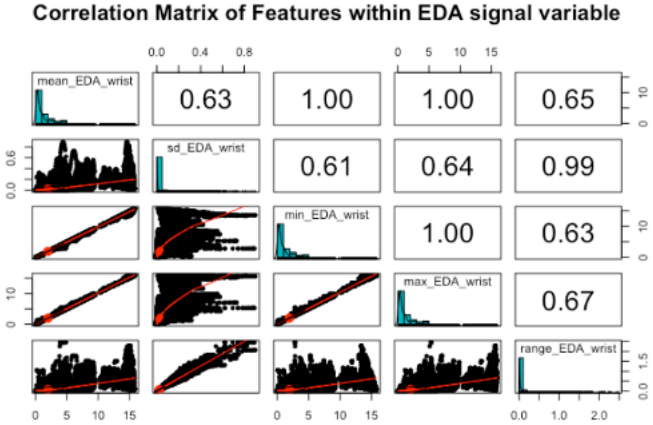

Fig 5.3 Correlation Matrix of Features within EDA signal variable

Other interesting insights like potential issues of multicollinearity were observed when plotting correlation matrices of the features. Figure 2 shows features engineered only from EDA and we can see that there exists a high level of collinearity. On the other hand, there was low correlation between features of different signal variables. High collinearity in predictor variables will influence the strategies and thresholds chosen for model selection and interaction effects, which we elaborate further in the paper in relation to our proposed model.

3 Literature Review

Stress can impact heart rate, blood pressure, and skin temperature [3]. When a threat is presented, it only takes the brain .9 milliseconds, or 1111 Hz, to shift a body into a full stress response [4]. In that time, the brain filters information relevant to our personal safety, redirects energy to only areas that are biologically relevant for survival, and releases stress hormones. While stress can be useful in some situations in that it activates fight-or-flight response and can help us during a dangerous situation, chronic stress can have many negative long-term impacts, even reducing brain mass in some cases.

However, stress is not the only emotion that impacts us physically. In fact, many of our physiological measures change based on our emotional state. L. Shu et. al. [5] examine how anger, anxiety, embarrassment, fear, amusement, happiness, and joy have different effects on our cardiovascular, electrodermal, respiratory, and electroencephalographic measures. Specifically, they saw that when patients are amused, heart rate becomes more variable and electrodermal activity, respiration rate and brain activity all increase.

With more advanced technology these days, studies have shown that changes in emotion can be effectively measured by sensors of physiological data [6]. This study found that emotional valence and arousal could all be reliably estimated by using low-cost sensors used to collect data on electroencephalography (EEG), galvanic skin response (GSR), and electromyographic signal (EMG) data.

There has also been research done on the appropriate time intervals to use when measuring different body measures such as motion, respiration, electrodermal activity, and heart rate. Different studies use various sampling rates of the raw signals and window sizes for feature engineering, and it appears that there is no consensus on the best-suited way to choose these values. Arif and Kattan [7] used a five second rolling window to analyze ACC data, and Uy et. al. [8] also used a five second rolling window for measuring Galvanic Skin Response, Blood Volume Pulse, and Respiratory Variability. Since we found multiple sources using a five-second window for predictors similar to the ones in our dataset, we decided to use this approach as well.

4 Methodology

4.1 Feature Engineering

As mentioned earlier, the brain takes .9 milliseconds to shift the body into a full stress response[4], which suggests that these data should be sampled at approximately 1111 Hz to properly identify a state of stress. However, the highest-resolution data is sampled at 700 Hz, which suggests that these data are not granular enough to fully capture a shift to stress. Thus, with computational cost in mind, we downsampled our data to 4Hz. We assigned a unique ID to each of the 15 subjects, and also added their personal characteristics. Since we are primarily interested in differentiating between stress and amusement, we only kept samples that were labeled 2 (stress) or 3 (amusement). We then relabeled amusement as 0 and stress as 1.

Next, we calculated features for the different modalities. Segmentation of the physiological signals was done using a sliding window, with a window shift of 0.25 seconds. The ACC and physiological features were all extracted using a 5 second window size [7][8]. To calculate some BVP and ECG features, we utilized peak detection algorithms [9] to identify the time between each successive peaks and used the peak intervals to calculate the mean, standard deviation and root-mean squared of heart rate variability within each assigned windows.

To calculate the respiration variables, we had to separate out inhales and exhales. Inhale data corresponded to respiration values greater than zero and exhale data corresponded to respiration values less than zero. For five-second intervals where someone either did not inhale or exhale, we coded the corresponding observations to equal zero instead of NA. Additionally, we found total respiration range which took both inhale and exhale values into account when calculating the maximums and minimums. Next, we calculated the inhale to exhale ratio by dividing our mean inhale column by our mean exhale column. Since this would have produced some infinite values

due to the fact that periods where someone was not exhaling were set to zero, we set these inhale/exhale ratios to equal 10. Additionally, we had a couple of intervals where the inhale/exhale ratio was very large, sometimes even greater than 1500. This was happening because some five-second windows in our data could have consisted of the participant inhaling for most of the time, thus we would be dividing by an extremely low exhale value. So, we re-coded any ratio greater than 10 to equal an upper bound of 10. This is a fair assumption, since it would not be physically possible for someone to inhale incredibly more than they exhale. We next calculated a breath rate by first summing the number of rows in one breath cycle, which is defined as one full inhale and one full exhale. Then, we took this sum and divided it by four, since the observations in our dataset were taken every quarter of a second, to get the time in seconds of one full breath cycle.

For ACC, we also calculated the 3D magnitude, $ACC = \sqrt{ACC_X^2 + ACC_Y^2 + ACC_Z^2}$. We then calculated the mean, standard deviation, maximum, minimum and range of these modalities; a summary of all predictor variables is given in Table 1.

Table 1: Summary of Sensor Data Categories, with the Extracted Predictors and Respective Symbols

Metric	**Features**
Chest Sensor Data	
Chest ACC	Mean (μ^C_{ACC}), SD (σ^C_{ACC})
Chest ECG	Mean (μ^C_{ECG}), SD (σ^C_{ECG}), Mean HRV (μ^C_{HRV}), SD HRV (σ^C_{HRV})
Chest EDA	Mean (μ^C_{EDA}), SD (σ^C_{EDA}), Min (Min^C_{EDA}) Max (Max^C_{EDA}), Range ($Range^C_{EDA}$)
Chest EMG	Mean (μ^C_{EMG}), SD (σ^C_{EMG}), Min (Min^C_{EMG}) Max (Max^C_{EMG}), Range ($Range^C_{EMG}$)
Chest Resp	Mean (μ^C_{inhale}), SD (σ^C_{inhale}) Mean (μ^C_{exhale}), SD (σ^C_{exhale}) Min (Min^C_{resp}), Max (Max^C_{resp}), Range ($Range^C_{resp}$) Mean (μ^C_{breath}), SD (σ^C_{breath}), Min (Min^C_{breath}) Max (Max^C_{breath}), Range ($Range^C_{breath}$), I/E Ratio
Chest Temp	Mean (μ^C_T), SD (σ^C_T), Min (Min^C_T), Max (Max^C_T), Range ($Range^C_T$)
Wrist Sensor Data	
Wrist ACC	Mean (μ^W_{ACC}), SD (σ^W_{ACC})
Wrist BVP	Mean (μ^W_{ECG}), SD (σ^W_{ECG}), Mean HRV (μ^W_{HRV}), SD HRV (σ^W_{HRV})
Wrist EDA	Mean (μ^W_{EDA}), SD (σ^W_{EDA}), Min (Min^W_{EDA}), Max (Max^W_{EDA}), Range ($Range^C_{EDA}$)
Wrist Temp	Mean (μ^W_T), SD (σ^W_T), Min (Min^W_T), Max (Max^W_T), Range ($Range^W_T$)

4.2 Model Selection

We first split our data into a 80/20 train/test split, perform 5-fold cross validation on the training data and lastly assess model performance by predicting on the testing data.

The extracted features serve as predictors for our classification models. We propose two different models, both using different types of sensor data: features from the wrist-worn device, and features from both chest-worn and wrist-worn devices. Since our research goals are concerned with the predictive power and interpretability of our models, logistic regression was chosen over other classification algorithms. To determine our final model, we performed variable selection by first including all relevant features and *participant* and then used backwards selection with AIC as the selection criterion. Additionally, we iteratively removed variables that had p-values greater than 0.1. Interaction effects between the remaining predictors and with *participant* were fit to determine if they would improve the predictive power of the models. We also noted that the models with the subjects' personal

characteristics did not increase accuracy, F_1-score or AUC values, and also led to increased computational cost. Thus, we did not include any personal characteristics in the final models.

We also tried fitted multilevel logistic regression models, where *participant* was the random effect. However, these models were computationally expensive, and did not improve the classification metrics. We also utilized the random forest algorithm; these models resulted in better predictions, but lack sufficient interpretation. Both of these algorithms are discussed in the Appendix.

The final models are below (Model 1 was fitted using only the wrist-worn data and Model 2 was fitted using all the sensor data):

$$
\begin{aligned}
Y_i &\sim Bernoulli(\pi_i) \\
log(\frac{\pi_i}{1-\pi_i}) &= \beta_0 + \beta_1\ \sigma^W_{iEDA} + \beta_2\ \sigma^W_{iACC} + \beta_3\ \sigma^W_{iHRV}\ + \beta_4\ \sigma^W_{iEDA} * \sigma^W_{iHRV}\ + \\
&\quad \beta_5\ \sigma^W_{iEDA} * \sigma^W_{iACC} + \beta_6\ \sigma^W_{iHRV} * \sigma^W_{iACC}\ + \\
&\quad \sum_{j=2}^{15} \beta_{7j}\ I(subject_i = j)\ + \sum_{j=2}^{15} \beta_{8j}\ \sigma^W_{iACC} * I(subject_i = j), \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
Y_i &\sim Bernoulli(\pi_i) \\
log(\frac{\pi_i}{1-\pi_i}) &= \beta_0 + \beta_1\ \mu^C_{iACC} + \beta_2\ \sigma^C_{iACC} + \beta_3\ \sigma^W_{iACC} + \beta_4\ \sigma^C_{iHR} + \beta_5\ \sigma^W_{iHR}\ + \beta_6\ \sigma^W_{iHRV}\ + \\
&\quad \beta_7\ rms^C_{iHRV} + \beta_8\ rms^W_{iHRV} + \beta_9\ \sigma^C_{iEDA}\ + \beta_{10}\ \sigma^W_{iEDA} + \beta_{11}\ min^C_{iEMG} + \beta_{12}\ \sigma_{i\ inhale}\ + \\
&\quad \beta_{13}\ \sigma_{i\ exhale} + \beta_{i14}\ max_{i\ breath} + \beta_{15}\ ie_{i\ ratio} + \beta_{16}\ \sigma^C_{iT} + \beta_{17}\ min^W_{iT} + \beta_{18}\ \sigma^C_{iEDA} * min^W_{iT}\ + \\
&\quad \sum_{j=2}^{15} \beta_{19j}\ I(subject_i = j) + \sum_{j=2}^{15} \beta_{20j}\ rms^W_{iHRV} * I(subject_i = j)\ + \sum_{j=2}^{15} \beta_{21j}\ \sigma^C_{iHR} * I(subject_i = j), \quad (2)
\end{aligned}
$$

i corresponds to each sample, and Y_i indicates the corresponding state (1 = stress or 0 = amusement). $subject_i$ represents the unique subject ID that was assigned to each of the 15 participants. The reference level for *subject* is Subject 10 (S10), who had the highest number of samples in the filtered dataset. The remaining levels $j \in (2, 3, \ldots, 14, 15)$ correspond to S2, S3, S4, S5, S6, S7, S8, S9, S11, S13, S14, S15, S16 and S17 respectively. The other predictors correspond to those displayed in Table 1.

4.3 Evaluation Metrics

We used accuracy, F_1-score, and AUC as evaluation metrics. Accuracy is the number of correctly classified observations over the total number of samples. We then define the F_1-score and AUC. The True Positive Rate (TPR/Recall) is the proportion of positive (stress) samples that are correctly identified. The False Positive Rate (FPR) is the proportion of negative (amusement) samples that are predicted to be positive over the total number of negative samples. Precision is the proportion of samples that are positive and classified as one over the total number of samples that are predicted to be positive. Thus, the F_1-score is the harmonic average of the precision and recall, and is a good choice for imbalanced classification tasks. The Receiver Operating Characteristics (ROC) curve plots the TPR against the FPR. The AUC is thus the area under the ROC curve. Both the F_1-score and AUC have a range of [0,1], where a higher value indicates better classification, and a value of 1 indicates perfect classification.

5 Results

5.1 Wrist Sensor Data

Using only data from the wrist-worn wearable, our model reports an accuracy of 86.2%, an F_1-score of 80.45 and an AUC of 84.7 when predicting on the testing data. This suggests that we can detect stress vs. amusement reasonably well using only the wrist-worn wearable, which may be preferable since a wrist-worn device is minimally intrusive and more convenient than a chest-worn device. A naive classifier using only wrist sensor data that predicts

Table 2: Model Results with 5-fold Cross Validation, with Mean and Standard Deviation reported

	Cross Validation Data			Test Data		
Model	Accuracy	F1-Score	AUC	Accuracy	F1-Score	AUC
Wrist Only	85.83% ± 0.45%	79.81 ± 0.77	84.17 ± 0.64	86.2%	80.45	84.70
Chest & Wrist	90.88% ± 0.05%	87.23 ± 0.08	90.05 ± 0.13	90.83%	87.27	90.19

only the majority class from each test fold reports an accuracy of 64.26 ± 0.00; this indicates our model exceeds that of a naive classifier.

5.2 All Sensor Data

Using data from both the chest-worn and wrist-worn wearables, our model reports an accuracy of 90.83%, an F_1-score of 87.27 and an AUC of 90.19 when predicting on the testing data. These values suggest that sensor data are useful in discriminating between stress and amusement conditions. These metrics are slightly better than the previous model, which indicates that having data from both the chest-worn and wrist-worn wearables is better than having just data from the wrist-worn device. A naive classifier using all sensor data that predicts only the majority class from each test fold reports an accuracy of 64.26 ± 0.00; similarly, this indicates our model exceeds that of a naive classifier.

5.3 Model Diagnostics

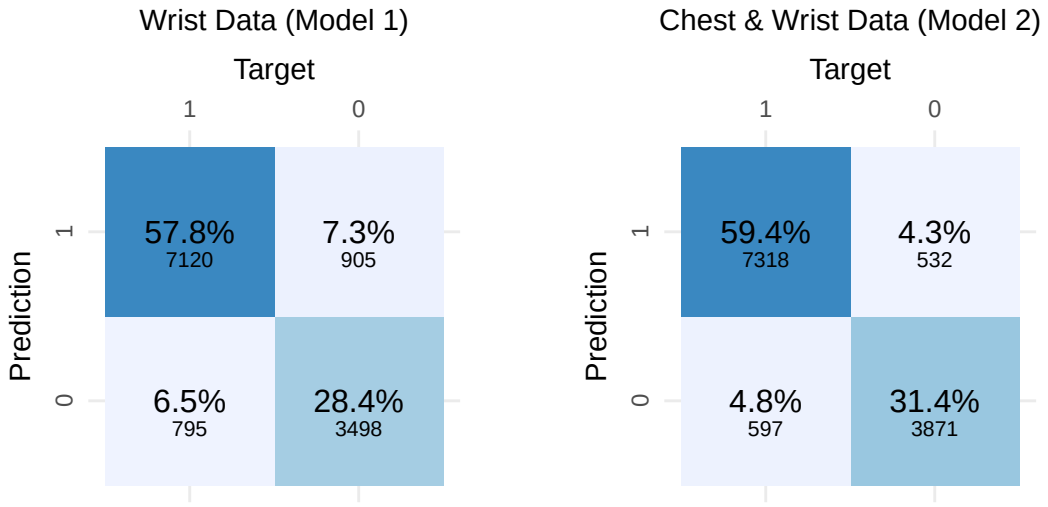

Figure 3: Confusion Matrices on Model Predictive Accuracy for model trained on only Wrist Data and Wrist and Chest data combined

From the metrics displayed in Table 2 and the confusion matrices in Figure 3, it seems that both the wrist model and chest & wrist models have relatively high classification accuracies when detecting stress vs. amusement.

It is also worth discussing independence when evaluating our logistic regression model. Adding *participant* addresses independence at the subject level, but each individual sample is not independent; presumably, a person's heart rate from one second ago affects their current heart rate. In addition, we display variance inflation factor (VIF) values for both our models in the Appendix in Tables 9-11. VIF quantifies the severity of multicollinearity in a logistic regression model, and a smaller VIF value indicates less multicollinearity. In general, most VIF values are below 10; however, *participant* and predictors that are used in interaction effects may have higher VIF values. However, since the standard errors of our estimates and interaction effects are relatively small, the larger VIF values are acceptable.

5.3.1 Sensitivity Analysis

Since there is no standard sampling frequency in the literature, we first examine the sensitivity of our sampling rate of 4 Hz. Table 3 displays the results of the wrist data and all data models, with the same predictors as Models 1 and 2, when using data that was sampled at 1 Hz. In general, the models using data sampled at 4 Hz perform better on the training data, while the models using data sampled at 1 Hz perform better on the testing data. There is also higher variability in the metrics of the models using data sampled at 1 Hz - this is because we are using a much smaller sample size, which leads to higher sampling variability. However, the difference in metrics is very minimal; thus we can verify that our results are not highly sensitive to different sampling frequencies.

Table 3: Model Results with Data Sampled at 1 Hz

	Cross Validation Data			Test Data		
Model	Accuracy	F1-Score	AUC	Accuracy	F1-Score	AUC
Wrist Only	85.8% ± 0.56%	79.73 ± 0.88	84.12 ± 0.69	86.43%	80.65	84.81
Chest & Wrist	90.78% ± 0.5%	87.08 ± 0.66	89.95 ± 0.51	91.1%	87.51	90.24

6 Discussion

6.1 Interpretation of Predictors

6.1.1 Wrist-Only Model

When standard deviation of wrist EDA increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 39.997 and 52.625 with mean 45.879, holding all else constant. When standard deviation of wrist ACC increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 15.812 and 99.409 with mean 1.56, holding all else constant. When standard deviation of wrist HRV increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.429 and 1.704 with mean 56.092, holding all else constant.

Holding standard deviation of wrist HRV constant, when standard deviation of wrist EDA and wrist ACC both increase by 1, we are 95% confident that the odds of being stressed versus amused multiply by an additional factor between 0.304 and 0.344 with mean 0.323. Holding standard deviation of wrist ACC constant, when standard deviation of wrist EDA and wrist HRV both increase by 1, we are 95% confident that the odds of being stressed versus amused multiply by an additional factor between 1.401 and 1.94 with mean 1.649. Holding standard deviation of wrist EDA constant, when standard deviation of wrist HRV and wrist ACC both increase by 1, we are 95% confident that the odds of being stressed versus amused multiply by an additional factor between 1.118 and 1.483 with mean 1.288.

6.1.2 Chest and Wrist Model

When standard deviation of chest EDA increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 2.957 and 3.712 with mean 3.313, holding all else constant. When standard deviation of wrist EDA increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 4860.614 and 8682.632 with mean 6496.378, holding all else constant. When standard deviation of chest temperature increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.143 and 1.232 with mean 1.186, holding all else constant. When minimum wrist temperature increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 0.15 and 0.182 with mean 0.165, holding all else constant. When minimum chest EMG increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 0.832 and 0.903 with mean 0.867, holding all else constant. When standard deviation of inhale increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.685 and 1.866 with mean 1.774, holding all else constant. When standard deviation of exhale increases by 1, we are 95%

confident that the odds of being stressed versus amused multiply by a factor between 0.736 and 0.799 with mean 0.766, holding all else constant. When maximum breath cycle time increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.889 and 1.827 with mean 1.953, holding all else constant. When inhale/exhale ratio increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 0.722 and 0.777 with mean 0.749, holding all else constant. When mean chest ACC increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 0.99 and 1.092 with mean 1.04, holding all else constant. When standard deviation of chest ACC increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 2.111 and 2.356 with mean 2.23, holding all else constant. When standard deviation of wrist ACC increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.374 and 1.582 with mean 1.474, holding all else constant. When standard deviation of wrist HR increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.176 and 1.292 with mean 1.232, holding all else constant. When standard deviation of wrist HRV increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.143 and 1.236 with mean 1.189, holding all else constant. When rms of wrist HRV increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.186 and 1.482 with mean 1.326, holding all else constant. When standard deviation of chest HR increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 0.396 and 0.575 with mean 0.477, holding all else constant. When rms of chest HR increases by 1, we are 95% confident that the odds of being stressed versus amused multiply by a factor between 1.022 and 1.105 with mean 1.063, holding all else constant.

Additionally, holding all else constant, when standard deviation of wrist HRV and minimum wrist temperature both increase by 1, we are 95% confident that the odds of being stressed versus amused multiply by an additional factor between 0.008 and 0.011 with mean 0.009.

It is also interesting to note that in the model built with chest and wrist data, most of our predictors appear to be standard deviation measures, and in the wrist-only model, participant is the only non-standard deviation predictor. This observation suggests that the variability in physiological signals is a more important distinguisher of stress and amusement as opposed to the magnitude of these signals in many cases.

6.2 Quantifying Heterogeneity

In the wrist only model, when compared to the baseline S10 and holding all else constant, we expect the odds of being stressed for S17 to multiply by $e^{0.256}$, or 1.29, while we expect the odds of being stressed for S7 to multiply by $e^{-5.765}$, or 0.003, and the odds of being stressed for S3 to multiply by $e^{-4.34}$, or 0.013.

For the standard deviation of the magnitude of wrist ACC, when compared to the baseline S10 and holding all else constant, as the standard deviation of the magnitude of wrist ACC increases by one, we expect the odds of being stressed for S7 to multiply by $e^{10.803}$, or 49168.08 more and the odds of being stressed for S5 to multiply by $e^{6.297}$, or 542.9 more. The values for the wrist only model are detailed in Table 5.

In the all data model, when compared to the baseline S10 and holding all else constant, we expect the odds of being stressed for S8 to multiply by $e^{0.413}$, or 1.51, and the odds of being stressed for S15 to multiply by $e^{0.178}$, or 1.19, while we expect the odds of being stressed for the other 12 subjects to multiply by a value less than 1.

For the standard deviation of wrist HR, when compared to the baseline S10 and holding all else constant, as the standard deviation of wrist HR increases by one, we expect the odds of being stressed for S11 to multiply by $e^{13.179}$, or 529135.6 more and the odds of being stressed for S17 to multiply by $e^{8.979}$, or 7934.7 more; conversely, as the standard deviation of wrist HR increases by one, we expect the odds of being stressed for S8 to multiply by $e^{-4.834}$, or 0.008. The values for the wrist only model are detailed in Table 8.

Although some variation between subjects is expected, within both models, we see large variability among the odds of being stressed. These results indicate that there is heterogeneity across individuals' responses to stress vs. amusement. Overall, although we only see the differences between S10 vs. other subjects (but not S2 vs. S3 for example), the disparity in the odds of being stressed across the 15 participants is pretty apparent. This could also be due to model instability or a small number of participants - a larger pool of participants could reduce heterogeneity and variability between the estimates.

7 Conclusion

While we were able to distinguish between stress and amusement conditions fairly well with our model, it is worth pointing out some limitations. First, we only had data for the 15 participants in this study, and participant ID is an important predictor in our models, so we are unable to generalize the likelihood of being stressed or amused to the general population. Also, since all of the data was collected in one sitting during an experiment, our data is very highly correlated and suggests there is heterogeneity across individuals' responses to stress vs. amusement. We can distinguish between stress and amusement in this controlled laboratory setting, but we are unable to generalize the differences in these emotions in a person's day-to-day life, or what stress and amusement might even look like if the measurements were taken on a different day.

Since it appears that the wrist wearable is sufficient for detecting stress versus amusement in the laboratory setting, further research could study wrist data for individuals over a long time period, perhaps a day or even a week, and see if different emotions are still distinguishable. Also, studies could be done with a larger random sample of individuals to try and see if there are any findings about stressed versus amused physiological responses generalizable to all individuals instead of just at the participant level.

8 Appendix

8.1 Bibliography

1. P. Schmidt et al. (2018). Introducing WESAD, a Multimodal Dataset for Wearable Stress and Affect Detection. Proceedings of the 2018 on International Conference on Multimodal Interaction - ICMI 18.
2. D. Watson, L. A. Clark and A. Tellegen. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of Personality and Social Psychology, 54, 6, 1063.
3. H. Yaribeygi et al. (2017). The impact of stress on body function: A review. EXCLI Journal, 2017, 16, pp. 1057–1072.
4. D. Strickland. (2019). .9 Milliseconds. Transitions.
5. L. Shu et al. (2018). A Review of Emotion Recognition Using Physiological Signals. Sensors (Basel), 18, 7.
6. D. Girardi et al. (2017). Emotion Detection Using Noninvasive Low Cost Sensors
7. M. Arif and A. Kattan. (2015). Physical Activities Monitoring Using Wearable Acceleration Sensors Attached to the Body. PLoS ONE 10, 7.
8. G. Uy et al. (2013). Correlation of Stress Inducers and Physiological Signals with Continuous Stress Annotations. WCTP 2012, PICT 7, pp. 178–183.
9. P.Westermark (2015). Package 'peakPick'. https://cran.r-project.org/web/packages/peakPick/peakPick.pdf

8.2 Random Forest Model

In addition to the logistic regression, we also fit several other types of models. The first of those was a random forest model. We built a random forest model using all of our data and just the wrist data, and had great overall performance for both. Our model built with all the data consisted of only three predictors: participant ID, max wrist electrodermal activity, and magnitude of standard deviation of chest movement, and it had a misclassification rate of 0.74%. The model built using only wrist data used five predictors, which included participant ID, minimum electrodermal activity, maximum electrodermal activity, the range of electrodermal activity, and standard deviation of Z-plane movement, and it had an even lower misclassification rate of 0.002%.

Despite the great performance of these models, we opted to go with the logistic regression as our final model since interpretability was important to the goals of the case study. Random forest is great for prediction, however it is difficult to interpret how exactly the predictors affect the response. Therefore, we were willing to sacrifice a little bit of accuracy in order to make more concrete conclusions about how our predictor variables affected stress and amusement.

8.3 Multilevel Model

Another method we tried was using a multilevel model (MLM) with the features engineered from the raw signal data being the random effects of the model, while survey and participant-specific data as the fixed effects. One large issue with running MLM was the computational intensity it took after including more than 2 fixed effects and 4 random effects. Considering the number of predictor variables we wished to include and analyze, we decided that it was not feasible to run a multi-level model due to the nature of our study and dataset having a large number of samples and predictor variables.

8.4 Model Output

8.5 VIF

Table 4: Coefficients obtained with Model 1 (Wrist Model)

Coefficient	Estimate	Std. Error	p-value	Lower CI	Upper CI
(Intercept)	4.027	0.236	< 2e-16	3.56444	4.48956
σ^W_{EDA}	3.826	0.07	< 2e-16	3.6888	3.9632
σ^W_{ACC}	3.68	0.469	< 2e-16	2.76076	4.59924
σ^W_{HRV}	0.445	0.045	< 2e-16	0.3568	0.5332
σ^W_{EDA} * σ^W_{ACC}	-1.129	0.031	< 2e-16	-1.18976	-1.06824
σ^W_{EDA} * σ^W_{HRV}	0.5	0.083	< 2e-16	0.33732	0.66268
σ^W_{ACC} * σ^W_{HRV}‘	0.253	0.072	< 2e-16	0.11188	0.39412
Participant					
S10	(Reference)				
S2	-1.479	0.242	< 2e-16	-1.95332	-1.00468
S3	-2.149	0.241	< 2e-16	-2.62136	-1.67664
S4	-2.279	0.244	< 2e-16	-2.75724	-1.80076
S5	12.174	0.736	< 2e-16	10.73144	13.61656
S6	-4.238	0.245	< 2e-16	-4.7182	-3.7578
S7	20.392	1.27	< 2e-16	17.9028	22.8812
S8	-1.56	0.241	< 2e-16	-2.03236	-1.08764
S9	0.656	0.312	0.035	0.0444800000000001	1.26752
S11	-2.505	0.243	< 2e-16	-2.98128	-2.02872
S13	-5.57	0.25	< 2e-16	-6.06	-5.08
S14	-0.034	0.27	0.899	-0.5632	0.4952
S15	-1.548	0.242	< 2e-16	-2.02232	-1.07368
S16	-1.537	0.249	< 2e-16	-2.02504	-1.04896
S17	-1.362	0.24	< 2e-16	-1.8324	-0.8916
σ^W_{ACC} *** Participant**					
σ^W_{ACC} * S10	(Reference)				
σ^W_{ACC} * S2	-2.561	0.489	< 2e-16	-3.51944	-1.60256
σ^W_{ACC} * S3	-2.686	0.477	< 2e-16	-3.62092	-1.75108
σ^W_{ACC} * S4	-2.457	0.491	< 2e-16	-3.41936	-1.49464
σ^W_{ACC} * S5	26.264	1.482	< 2e-16	23.35928	29.16872
σ^W_{ACC} * S6	-4.417	0.484	< 2e-16	-5.36564	-3.46836
σ^W_{ACC} * S7	45.062	2.59	< 2e-16	39.9856	50.1384
σ^W_{ACC} * S8	-2.671	0.478	< 2e-16	-3.60788	-1.73412
σ^W_{ACC} * S9	1.091	0.622	0.079	-0.12812	2.31012
σ^W_{ACC} * S11	-2.302	0.483	< 2e-16	-3.24868	-1.35532
σ^W_{ACC} * S13	-2.119	0.475	< 2e-16	-3.05	-1.188
σ^W_{ACC} * S14	0.902	0.559	0.107	-0.19364	1.99764
σ^W_{ACC} * S15	-2.553	0.48	< 2e-16	-3.4938	-1.6122
σ^W_{ACC} * S16	-2.554	0.508	< 2e-16	-3.54968	-1.55832
σ^W_{ACC} * S17	-2.787	0.479	< 2e-16	-3.72584	-1.84816

Table 5: Coefficients obtained with Model 2 (Chest and Wrist Model)

Coefficient	Estimate	Std. Error	p-value	Lower CI	Upper CI
(Intercept)	5.596	0.115	< 2e-16	5.3706	5.8214
σ^{C}_{EDA}	1.198	0.058	< 2e-16	1.08432	1.31168
σ^{W}_{EDA}	8.779	0.148	< 2e-16	8.48892	9.06908
σ^{C}_{T}	0.171	0.019	< 2e-16	0.13376	0.20824
min^{W}_{T}	-1.801	0.05	< 2e-16	-1.899	-1.703
min^{C}_{EMG}	-0.143	0.021	< 2e-16	-0.18416	-0.10184
σ_{inhale}	0.573	0.026	< 2e-16	0.52204	0.62396
σ_{exhale}	-0.266	0.021	< 2e-16	-0.30716	-0.22484
max_{breath}	0.636	0.017	< 2e-16	0.60268	0.66932
ie_{ratio}	-0.289	0.019	< 2e-16	-0.32624	-0.25176
μ^{C}_{ACC}	0.039	0.025	0.127	-0.01	0.088
σ^{C}_{ACC}	0.802	0.028	< 2e-16	0.74712	0.85688
σ^{W}_{ACC}	0.388	0.036	< 2e-16	0.31744	0.45856
σ^{W}_{HR}	0.209	0.024	< 2e-16	0.16196	0.25604
σ^{W}_{HRV}	0.173	0.02	< 2e-16	0.1338	0.2122
rms^{W}_{HRV}	0.282	0.057	< 2e-16	0.17028	0.39372
σ^{C}_{HR}	-0.74	0.095	< 2e-16	-0.9262	-0.5538
rms^{C}_{HRV}	0.061	0.02	0.002	0.0218	0.1002
σ^{W}_{EDA} * min^{W}_{T}	-4.665	0.08	< 2e-16	-4.8218	-4.5082
Participant					
S10	(Reference)				
S2	-0.171	0.116	0.14	-0.39836	0.05636
S3	-2.475	0.165	< 2e-16	-2.7984	-2.1516
S4	-1.63	0.213	< 2e-16	-2.04748	-1.21252
S5	-2.028	0.156	< 2e-16	-2.33376	-1.72224
S6	-4.503	0.152	< 2e-16	-4.80092	-4.20508
S7	-3.272	0.143	< 2e-16	-3.55228	-2.99172
S8	-1.01	0.12	< 2e-16	-1.2452	-0.7748
S9	0.308	0.115	0.007	0.0826	0.5334
S11	-3.188	0.142	< 2e-16	-3.46632	-2.90968
S13	-4.831	0.178	< 2e-16	-5.17988	-4.48212
S14	0.007	0.115	0.951	-0.2184	0.2324
S15	-0.115	0.149	0.443	-0.40704	0.17704
S16	-0.64	0.158	< 2e-16	-0.94968	-0.33032
S17	0.397	0.116	0.001	0.16964	0.62436
rms^{W}_{HRV} *** Participant**					
rms^{W}_{HRV} * S10	(Reference)				
rms^{W}_{HRV} * S2	-0.284	0.077	< 2e-16	-0.43492	-0.13308
rms^{W}_{HRV} * S3	0.542	0.095	< 2e-16	0.3558	0.7282
rms^{W}_{HRV} * S4	-0.424	0.254	0.095	-0.92184	0.07384
rms^{W}_{HRV} * S5	-0.561	0.16	< 2e-16	-0.8746	-0.2474
rms^{W}_{HRV} * S6	-0.3	0.094	0.001	-0.48424	-0.11576
rms^{W}_{HRV} * S7	-1.794	0.11	< 2e-16	-2.0096	-1.5784
rms^{W}_{HRV} * S8	-0.192	0.09	0.032	-0.3684	-0.0156
rms^{W}_{HRV} * S9	-0.358	0.078	< 2e-16	-0.51088	-0.20512
rms^{W}_{HRV} * S11	1.116	0.092	< 2e-16	0.93568	1.29632
rms^{W}_{HRV} * S13	-1.357	0.181	< 2e-16	-1.71176	-1.00224
rms^{W}_{HRV} * S14	-0.309	0.071	< 2e-16	-0.44816	-0.16984
rms^{W}_{HRV} * S15	-1.252	0.08	< 2e-16	-1.4088	-1.0952
rms^{W}_{HRV} * S16	0.089	0.092	0.331	-0.09132	0.26932
rms^{W}_{HRV} * S17	-1.226	0.075	< 2e-16	-1.373	-1.079

Table 6: Coefficients obtained with Model 2 (Chest and Wrist Model)

Coefficient	Estimate	Std. Error	p-value	Lower CI	Upper CI
σ_{HR}^{C} * **Participant**					
σ_{HR}^{C} * S10	(Reference)			NA	NA
σ_{HR}^{C} * S2	0.213	0.117	0.07	-0.01632	0.44232
σ_{HR}^{C} * S3	0.276	0.111	0.013	0.05844	0.49356
σ_{HR}^{C} * S4	0.725	0.129	$<$ 2e-16	0.47216	0.97784
σ_{HR}^{C} * S5	-0.04	0.13	0.757	-0.2948	0.2148
σ_{HR}^{C} * S6	0.902	0.131	$<$ 2e-16	0.64524	1.15876
σ_{HR}^{C} * S7	0.561	0.126	$<$ 2e-16	0.31404	0.80796
σ_{HR}^{C} * S8	-0.629	0.112	$<$ 2e-16	-0.84852	-0.40948
σ_{HR}^{C} * S9	0.309	0.11	0.005	0.0934	0.5246
σ_{HR}^{C} * S11	1.714	0.125	$<$ 2e-16	1.469	1.959
σ_{HR}^{C} * S13	0.655	0.137	$<$ 2e-16	0.38648	0.92352
σ_{HR}^{C} * S14	0.377	0.105	$<$ 2e-16	0.1712	0.5828
σ_{HR}^{C} * S15	0.29	0.113	0.01	0.06852	0.51148
σ_{HR}^{C} * S16	0.548	0.11	$<$ 2e-16	0.3324	0.7636
σ_{HR}^{C} * S17	1.168	0.106	$<$ 2e-16	0.96024	1.37576

Table 7: Coefficients obtained with Model 1 - Quantifying Hetereogeneity

Term	Log Odds	Std. Error	Odds	Lower CI	Upper CI
Participant					
S2	-1.479	0.242	2.280000e-01	1.420000e-01	3.660000e-01
S3	-2.149	0.241	1.170000e-01	7.300000e-02	1.870000e-01
S4	-2.279	0.244	1.020000e-01	6.300000e-02	1.650000e-01
S5	12.174	0.736	1.936412e+05	4.580491e+04	8.186222e+05
S6	-4.238	0.245	1.400000e-02	9.000000e-03	2.300000e-02
S7	20.392	1.270	7.178041e+08	5.955920e+07	8.650935e+09
S8	-1.560	0.241	2.100000e-01	1.310000e-01	3.370000e-01
S9	0.656	0.312	1.926000e+00	1.046000e+00	3.548000e+00
S11	-2.505	0.243	8.200000e-02	5.100000e-02	1.320000e-01
S13	-5.570	0.250	4.000000e-03	2.000000e-03	6.000000e-03
S14	-0.034	0.270	9.660000e-01	5.690000e-01	1.640000e+00
S15	-1.548	0.242	2.130000e-01	1.320000e-01	3.420000e-01
S16	-1.537	0.249	2.150000e-01	1.320000e-01	3.500000e-01
S17	-1.362	0.240	2.560000e-01	1.600000e-01	4.100000e-01
σ^W_{ACC} * **Participant**					
σ^W_{ACC} * S2	-2.561	0.489	7.700000e-02	3.000000e-02	2.020000e-01
σ^W_{ACC} * S3	-2.686	0.477	6.800000e-02	2.700000e-02	1.740000e-01
σ^W_{ACC} * S4	-2.457	0.491	8.600000e-02	3.300000e-02	2.250000e-01
σ^W_{ACC} * S5	26.264	1.482	2.548917e+11	1.395468e+10	4.655768e+12
σ^W_{ACC} * S6	-4.417	0.484	1.200000e-02	5.000000e-03	3.100000e-02
σ^W_{ACC} * S7	45.062	2.590	3.716331e+19	2.318608e+17	5.956642e+21
σ^W_{ACC} * S8	-2.671	0.478	6.900000e-02	2.700000e-02	1.760000e-01
σ^W_{ACC} * S9	1.091	0.622	2.979000e+00	8.810000e-01	1.007400e+01
σ^W_{ACC} * S11	-2.302	0.483	1.000000e-01	3.900000e-02	2.580000e-01
σ^W_{ACC} * S13	-2.119	0.475	1.200000e-01	4.700000e-02	3.050000e-01
σ^W_{ACC} * S14	0.902	0.559	2.465000e+00	8.240000e-01	7.377000e+00
σ^W_{ACC} * S15	-2.553	0.480	7.800000e-02	3.000000e-02	1.990000e-01
σ^W_{ACC} * S16	-2.554	0.508	7.800000e-02	2.900000e-02	2.100000e-01
σ^W_{ACC} * S17	-2.787	0.479	6.200000e-02	2.400000e-02	1.580000e-01

Table 8: Coefficients obtained with Model 2 - Quantifying Hetereogeneity

Term	Log Odds	Std. Error	Odds	Lower CI	Upper CI
Participant					
S2	-0.171	0.116	0.843	0.671	1.058
S3	-2.475	0.165	0.084	0.061	0.116
S4	-1.630	0.213	0.196	0.129	0.297
S5	-2.028	0.156	0.132	0.097	0.179
S6	-4.503	0.152	0.011	0.008	0.015
S7	-3.272	0.143	0.038	0.029	0.050
S8	-1.010	0.120	0.364	0.288	0.461
S9	0.308	0.115	1.360	1.086	1.703
S11	-3.188	0.142	0.041	0.031	0.054
S13	-4.831	0.178	0.008	0.006	0.011
S14	0.007	0.115	1.007	0.803	1.263
S15	-0.115	0.149	0.892	0.665	1.195
S16	-0.640	0.158	0.527	0.387	0.718
S17	0.397	0.116	1.487	1.185	1.867
rms^{W}_{HRV} *** Participant**					
rms^{W}_{HRV} * S2	-0.284	0.077	0.753	0.647	0.876
rms^{W}_{HRV} * S3	0.542	0.095	1.719	1.427	2.072
rms^{W}_{HRV} * S4	-0.424	0.254	0.654	0.398	1.076
rms^{W}_{HRV} * S5	-0.561	0.160	0.571	0.417	0.781
rms^{W}_{HRV} * S6	-0.300	0.094	0.740	0.616	0.891
rms^{W}_{HRV} * S7	-1.794	0.110	0.166	0.134	0.206
rms^{W}_{HRV} * S8	-0.192	0.090	0.826	0.693	0.984
rms^{W}_{HRV} * S9	-0.358	0.078	0.699	0.600	0.815
rms^{W}_{HRV} * S11	1.116	0.092	3.054	2.551	3.655
rms^{W}_{HRV} * S13	-1.357	0.181	0.258	0.181	0.367
rms^{W}_{HRV} * S14	-0.309	0.071	0.734	0.639	0.843
rms^{W}_{HRV} * S15	-1.252	0.080	0.286	0.244	0.334
rms^{W}_{HRV} * S16	0.089	0.092	1.094	0.913	1.310
rms^{W}_{HRV} * S17	-1.226	0.075	0.293	0.253	0.340
σ^{C}_{HR} *** Participant**					
σ^{C}_{HR} * S2	0.213	0.117	1.238	0.983	1.558
σ^{C}_{HR} * S3	0.276	0.111	1.318	1.060	1.638
σ^{C}_{HR} * S4	0.725	0.129	2.065	1.603	2.659
σ^{C}_{HR} * S5	-0.040	0.130	0.961	0.745	1.239
σ^{C}_{HR} * S6	0.902	0.131	2.465	1.908	3.183
σ^{C}_{HR} * S7	0.561	0.126	1.752	1.369	2.244
σ^{C}_{HR} * S8	-0.629	0.112	0.533	0.428	0.664
σ^{C}_{HR} * S9	0.309	0.110	1.363	1.098	1.690
σ^{C}_{HR} * S11	1.714	0.125	5.552	4.349	7.088
σ^{C}_{HR} * S13	0.655	0.137	1.924	1.471	2.517
σ^{C}_{HR} * S14	0.377	0.105	1.458	1.187	1.792
σ^{C}_{HR} * S15	0.290	0.113	1.337	1.070	1.669
σ^{C}_{HR} * S16	0.548	0.110	1.729	1.395	2.144
σ^{C}_{HR} * S17	1.168	0.106	3.215	2.610	3.959

Table 9: VIF Values - Model 1

Term	VIF
σ^W_{EDA}	3.399
σ^W_{ACC}	256.719
σ^W_{HRV}	15.832
$\sigma^W_{EDA} * \sigma^W_{ACC}$	2.657
$\sigma^W_{EDA} * \sigma^W_{HRV}$	10.747
$\sigma^W_{ACC} * \sigma^W_{HRV}$‘	9.249
S2	34.238
S3	26.465
S4	24.095
S5	105.757
S6	21.026
S7	247.069
S8	30.031
S9	56.252
S11	21.161
S13	14.416
S14	34.539
S15	32.267
S16	35.668
S17	34.546
$\sigma^W_{ACC} *$ S2	26.664
$\sigma^W_{ACC} *$ S3	42.962
$\sigma^W_{ACC} *$ S4	20.439
$\sigma^W_{ACC} *$ S5	105.100
$\sigma^W_{ACC} *$ S6	23.205
$\sigma^W_{ACC} *$ S7	246.714
$\sigma^W_{ACC} *$ S8	38.854
$\sigma^W_{ACC} *$ S9	56.521
$\sigma^W_{ACC} *$ S11	25.507
$\sigma^W_{ACC} *$ S13	73.648
$\sigma^W_{ACC} *$ S14	28.493
$\sigma^W_{ACC} *$ S15	41.220
$\sigma^W_{ACC} *$ S16	26.571
$\sigma^W_{ACC} *$ S17	36.072

Table 10: VIF Values - Model 2

Term	VIF
σ^{C}_{EDA}	1.727
σ^{W}_{EDA}	8.985
σ^{C}_{T}	1.699
min^{W}_{T}	13.052
min^{C}_{EMG}	1.336
σ_{inhale}	1.562
σ_{exhale}	1.680
max_{breath}	1.238
ie_{ratio}	1.530
μ^{C}_{ACC}	2.551
σ^{C}_{ACC}	1.713
σ^{W}_{ACC}	1.546
σ^{W}_{HR}	2.780
σ^{W}_{HRV}	1.791
rms^{W}_{HRV}	15.271
σ^{C}_{HR}	37.769
rms^{C}_{HRV}	1.859
σ^{W}_{EDA} * min^{W}_{T}	11.358
S2	6.115
S3	6.461
S4	7.441
S5	6.633
S6	3.540
S7	2.442
S8	4.281
S9	5.542
S11	3.831
S13	5.346
S14	5.590
S15	6.011
S16	8.148
S17	5.366
rms^{W}_{HRV} * S2	2.299
rms^{W}_{HRV} * S3	3.608
rms^{W}_{HRV} * S4	5.803
rms^{W}_{HRV} * S5	4.010
rms^{W}_{HRV} * S6	1.908
rms^{W}_{HRV} * S7	1.555
rms^{W}_{HRV} * S8	1.974
rms^{W}_{HRV} * S9	2.253
rms^{W}_{HRV} * S11	2.447
rms^{W}_{HRV} * S13	3.145
rms^{W}_{HRV} * S14	2.970
rms^{W}_{HRV} * S15	2.156
rms^{W}_{HRV} * S16	3.948
rms^{W}_{HRV} * S17	3.040

Table 11: VIF Values - Model 2

Term	VIF
σ_{HR}^{C} * S2	3.718
σ_{HR}^{C} * S3	4.708
σ_{HR}^{C} * S4	2.297
σ_{HR}^{C} * S5	2.953
σ_{HR}^{C} * S6	2.460
σ_{HR}^{C} * S7	2.827
σ_{HR}^{C} * S8	3.962
σ_{HR}^{C} * S9	4.557
σ_{HR}^{C} * S11	2.883
σ_{HR}^{C} * S13	2.512
σ_{HR}^{C} * S14	5.135
σ_{HR}^{C} * S15	3.782
σ_{HR}^{C} * S16	3.906
σ_{HR}^{C} * S17	4.722