Personality differences in experienced stress during a blood donation: a DISTRESS study.

Ruben E.M. Lohle

Student number: 1261953

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN COMMUNICATION AND INFORMATION SCIENCES,
MASTER TRACK DATA SCIENCE: BUSINESS AND GOVERNANCE,
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

Thesis committee:

dr. E.M.J. Huis in 't Veld

dr. E.J.M. van Wingerden

Tilburg University

School of Humanities and Digital Sciences

Department of Cognitive Science & Artificial Intelligence

Tilburg, the Netherlands

May 2019

Preface

Finally, here we are, ready to finalize my student period!

After completing my Masters in Fiscal Economics at Tilburg University, I wanted to challenge myself in the field of data science. This was a challenging step for me because I was ignorant in the world of big data. However, I have learned that you can learn anything if you are willing to put enough effort and time into it. By taking on this challenge, I am now ready to combine my knowledge in fiscal economics and data science in my future career.

I could not have achieved this without the support of some people. I would like to thank the following people explicitly.

First, I particularly want to express my thanks to Elisabeth Huis in ‘t Veld and Judita Rudokaite for the feedback and valuable conversations. You gave me new insights into conducting research, making analyses, and academic writing. As a result, I have learned a lot during the whole process of writing a Master thesis in the field of data science.

In addition, I want to express my gratitude to my family and friends. In particular to Jorien, Berend, Casper and Thomas who have spent many hours with me in the library. It was nice to have such people around for relaxation between the sometimes long and stressful study hours.

Ruben Lohle

Tilburg, May 2019

Abstract

The experience of stress among donors fluctuates during the whole course of a blood donation. Especially the moment of needle insertion is regarded as a highly salient event where there is a peak in the experience of stress. In order to find the cause for the experience of stress, it seems not unlikely to state that personality could influence the experience of stress during a blood donation.

In this study, a unique dataset will be used to examine whether it is possible to predict which donors experience less or more stress during a blood donation on the basis of personality. The analysis is conducted by performing several machine learning algorithms like decision trees, random forests and support vector machines. The hypothesis of this study predicts that the experience of stress during a blood donation can indeed be predicted on the basis of personality.

This study contributes to the existing literature by empirically identifying the predictive value of personality on the experience of stress during a blood donation. No other studies have performed such research in the setting of a blood donation.

Given the findings of this study, the experience of stress during a blood donation cannot be predicted on the basis of personality. Future research should strive for applying different methods that provide trustworthy findings on the relationship between personality and the experience of stress during a blood donation.

Contents

1 INTRODUCTION 5

1.1 CONTEXT 5

1.2 RESEARCH QUESTION 6

1.3 FRAMEWORK 6

2 LITERATURE REVIEW 7

2.1 INTRODUCTION 7

2.2 EXPERIENCED STRESS RESPONSES DURING A BLOOD DONATION 7

2.3 PERSONALITY TRAITS RELATED TO EXPERIENCED STRESS 8

2.4 JUSTIFICATION OF THE RESEARCH QUESTION 8

3 METHODOLOGY 10

3.1 ORIGINAL DATASET DESCRIPTION 10

3.2 FEATURE EXTRACTION 10

3.3 OUTLIERS 12

3.4 MISSING DATA 13

3.5 MACHINE LEARNING ALGORITHMS 15

3.6 EVALUATION METHOD 17

4 RESULTS 19

4.1 PREDICTION RESULTS 19

4.2 POST-HOC ANALYSIS 20

5 DISCUSSION 25

5.1 STUDY LIMITATIONS 26

5.2 FUTURE RESEARCH 27

6 CONCLUSION 28

7 REFERENCES 29

APPENDIX A 32

APPENDIX B 33

APPENDIX C 35

1 Introduction

1.1 Context

In the Netherlands, almost 330,000 voluntary blood donors donate every year to save 25,000 human lives (Sanquin, 2019). Sanquin is the blood supply organization that is responsible for the blood supply in the Netherlands, on a not-for-profit basis. Since the donated blood can only be stored for a particular time, it is crucial that healthy donors provide their blood on a regular basis to ensure a sufficient and safe blood supply. To achieve a healthy and safe blood supply on a regular basis, recruitment and retention strategies of donors are of high importance for Sanquin.

One way to ensure that blood supply organizations such as Sanquin maintain their donor pool is to conduct research on the emotional experiences of donors during donation, such as stress and fear, that can be harmful for (the retention of) donors. Sanquin aims to ensure that donors can donate their blood in a safe and pleasant way and wants to minimize the experience of fear or stress that relate to adverse donation experiences such as fainting.

Many previous studies already looked at the experience of stress during blood donation, however, the DISTRESS study by Hoogerwerf et al. (2017; 2018) was the first to assess self-reported psychological, physiological and hormonal stress patterns during multiple stages of blood donation. They report that self-reported psychological stress (arousal, stress), but also objectively measured physiological stress (blood pressure, heart rate) and hormonal stress levels (cortisol) fluctuate during the whole blood donation process from start to end, where especially the needle-insertion is a highly salient event.

Nonetheless, there is still little knowledge about the effect of personality on the experienced stress during a blood donation. Since the participants in the DISTRESS study have also completed a personality inventory questionnaire, the Five-Factor Personality Inventory ("FFPI"), these data could be analyzed to gain new insights into who is most likely to experience stress during a blood donation. These results could provide valuable insights into who is more sensitive to experiencing stress and can help Sanquin provide more targeted, personalized or improved retention strategies.

1.2 Research question

This study aims to build on the results by Hoogerwerf et al. (2017; 2018) by modelling perceived psychological and physiological stress on different personality traits from the DISTRESS study. This thesis will use machine learning techniques to examine the following research question:

> *"Can we predict which donors will experience less or more stress during a blood donation on the basis of personality?"*

The following two sub-questions can help in formulating an answer to the research question above:

1. What technique will be used to label donors into different categories of experienced stress?
2. Can experienced stress during a blood donation be predicted based on personality?

The first sub-question is relevant because no labels have so far been assigned to which donors perceive "high" levels of stress versus those who do not. This is also a critical question to answer since the analysis of the data consists of different classification algorithms. As it is not possible to perform classification algorithms on non-labeled data, the labelling process is a requirement for a successful experimental procedure.

After the first sub-question has been answered, the second sub-question is relevant for the question whether experienced stress during a blood donation can be predicted based on personality. To find an appropriate answer to this question, different classification techniques will be used like decision trees, random forests, and support vector machines.

1.3 Framework

To give an overview of how this thesis will be structured, section 2 will start with a literature study on related work. Subsequently, section 3 describes the dataset accompanied by the experimental procedure and the algorithms used. Section 4 will present and interpret the results of the analysis, followed by a post-hoc descriptive analysis. Finally, section 5 and 6 will respectively provide a discussion and present the conclusions of this study accompanied by recommendations for future research.

2 Literature review

2.1 Introduction

This section gives an overview of previous research on experienced stress responses during a blood donation and on previous research regarding the relationship between different personality traits and experienced stress. Section 2.2 indicates that previous research has shown that experienced psychological and physiological stress responses fluctuate during the whole blood donation. Section 2.3 provides an overview of conducted research that demonstrates that in particular the personality trait neuroticism is highly related to experienced stress. Finally, section 2.4 illustrates the relevance of this thesis by describing that no previous research has been conducted to assess whether different personality traits are related to experienced psychological stress and physiological stress during a blood donation

2.2 Experienced stress responses during a blood donation

Since 2000, six studies were found that investigated experienced stress responses during a blood donation. Two of these studies assessed both physiological and psychological stress responses (Ulrich et al. 2003; Byrne & Ditto 2005). Two studies solely assessed psychological stress responses (Hanson & France 2009; Ditto & France 2006). Finally, the studies of Hoogerwerf et al. (2017; 2018) have measured psychological, hormonal and physiological stress responses throughout multiple stages of the blood donation.

2.2.1 Psychological stress responses

All studies found that blood donation incur some kind of psychological stress, such as anxiety (Hanson & France 2009; Ditto & France 2006; Byrne & Ditto 2005), arousal (Hoogerwerf et al. 2017) and fear (Ulruch et al. 2013), which were significantly higher during the blood donation than when leaving the donation center. Additionally, Hanson & France (2009) investigated levels of anxiety at three moments around the donation, and found that anxiety among donors was higher during the actual donation as compared to pre-donation and post-donation. This is corroborated by the results reported by Hoogerwerf et al. (2017), who found a steady increase with a peak in self-reported stress towards needle insertion.

2.2.2 Physiological stress responses

Three recent studies have been conducted in the field of physiological stress responses during a blood donation and found that heart rate increased (Ulrich et al., 2003) and systolic blood pressure decreased (Byrne & Ditto, 2005) from the start to the end. However, Hoogerwerf et al. (2018) measured stress during multiple stages of the blood donation, as opposed to just a pre and post measurement, and found that systolic blood pressure and heart rate variability increased towards needle insertion and then decreased after uncoupling the needle. Diastolic blood pressure increased during the whole process of the blood donation, and heart rate showed a U-shape curve during the whole procedure. They concluded that physiological stress increased as a consequence of needle insertion, followed by a decrease when leaving the donation center.

2.3 Personality traits related to experienced stress

Most psychologists recognize that a comprehensive description of personality must at least include information of the following five factors: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (Kumar, 2016). Neuroticism is described as a personality trait that is associated with tension, irritability, and vulnerability to stress and consists of sub-traits such as anxiety, sadness, self-consciousness, and hostility. It may therefore not be surprising that people who experience higher levels of neuroticism (and conscientiousness, characterized by goal-oriented behavior) are more likely to experience stress (Ebstrup et al. 2011; Garbarino et al. 2014, Roohafza et al. 2016; Saklofske et al. 2011). On the other hand, these studies found that people high in extraversion (positive emotionality, sociability) and agreeableness (pro-social attitude, altruism, compliance, trust) are less likely to experience stress, whereas openness to experience was not related to stress levels at all. On a critical note, Garbarino et al. (2014) found fairly weakly associations and reported that less than 10 % of the variance of stress was predicted by personality traits.

2.4 Justification of the research question

Given the previous research on different stress responses during a blood donation and the relationship between personality traits and the experience of stress, to my knowledge, no studies have been conducted that have investigated the association between personality traits and the experience of stress during a blood donation. Besides, the studies (outside a blood donation setting) that investigate the relationship between personality traits and stress use a methodology where stress is measured by conducting a questionnaire. However, measuring stress as self-reported stress could lead to a response bias, which is not the case when measuring stress as physiological stress. Therefore, it would be interesting to conduct a study where the relationship between personality traits and physiological stress during a blood donation is investigated. Hence, this study will investigate whether personality traits are associated with the experience of stress during a blood donation using the DISTRESS study by Hoogerwerf (2017; 2018) who collected data on the Five Factor Personality Inventory ("FFPI") personality traits (rather than the Big Five personality traits). The FFPI consists of the following five personality traits (Hendriks et al. 1999): extraversion, agreeableness, conscientiousness, emotional stability, and autonomy. These FFPI traits show many similarities with the Big Five personality traits. By using these FFPI scores and different stress responses (psychological and physiological), the following research question will be answered:

> *"Can we predict which donors will experience less or more stress during a blood donation on the basis of personality?"*

The hypothesis towards this research question - based on previous research - will be as follows:

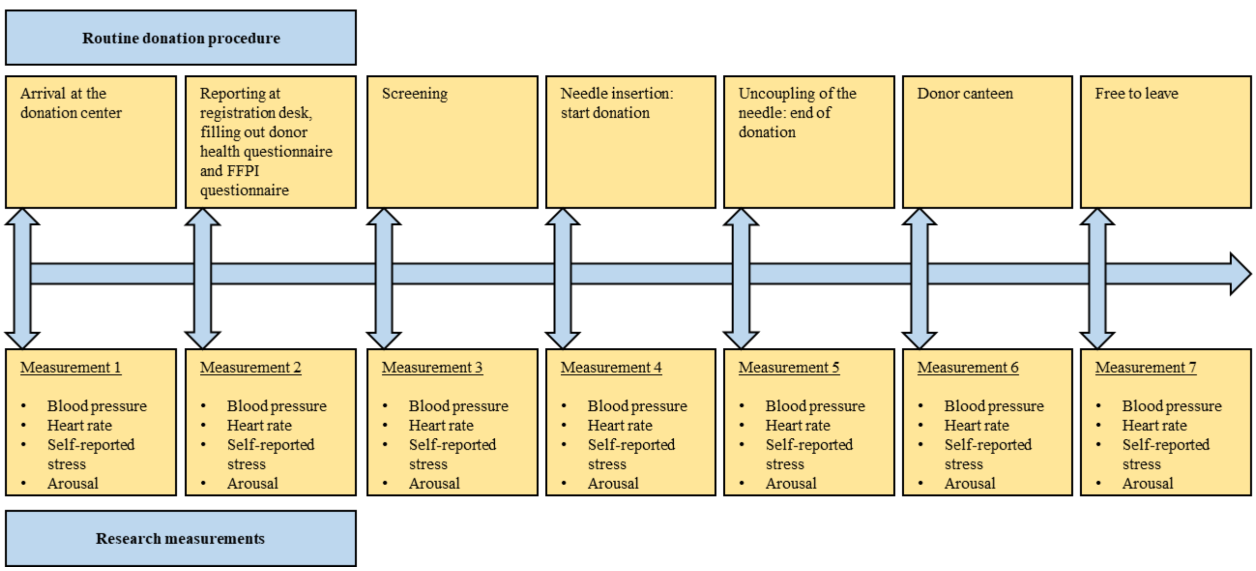
> The experience of stress can indeed be predicted on the basis of personality. Especially donors that score low on the personality trait FFPI emotional stability (high level of neuroticism) will experience more stress during a blood donation compared to other FFPI personality traits.

3 Methodology

3.1 Original dataset description

This study uses the dataset of the DISTRESS study by Hoogerwerf et al. (2017; 2018). Initially, the experimental setup of the DISTRESS study was to invite 1,502 donors to participate. After certain inclusion criteria, the final group for taking measurements consisted of 399 donors. This random sample (N = 399) represents both first-time (N = 199) as experienced donors (N = 200) that were invited to participate. All these donors went through the same donation procedure where multiple measurements were taken, running from October 2014 to April 2016. The following figure illustrates the set-up of the DISTRESS study where solely the relevant measurements for this study are taken into account. Sanquin has provided the DISTRESS dataset to us to conduct an additional study to predict which donors will experience less or more stress during a blood donation on the basis of personality. The DISTRESS dataset consists of 1,239 instances with 828 features, where 399 instances represent the donors that have participated in the DISTRESS study.

Figure 1 Set-up of the study procedure.



3.2 Feature extraction

3.2.1 Predictor variables

The FFPI consists of 100 brief statements and can be administered in 10 to 15 minutes (Hendriks et al. 1999). Based on the responses on these statements, a sum score ranging between 0 and 100 is calculated for each donor on every personality trait. A full overview of all personality traits, accompanied by a description and exact name in the DISTRESS dataset can be consulted in Table

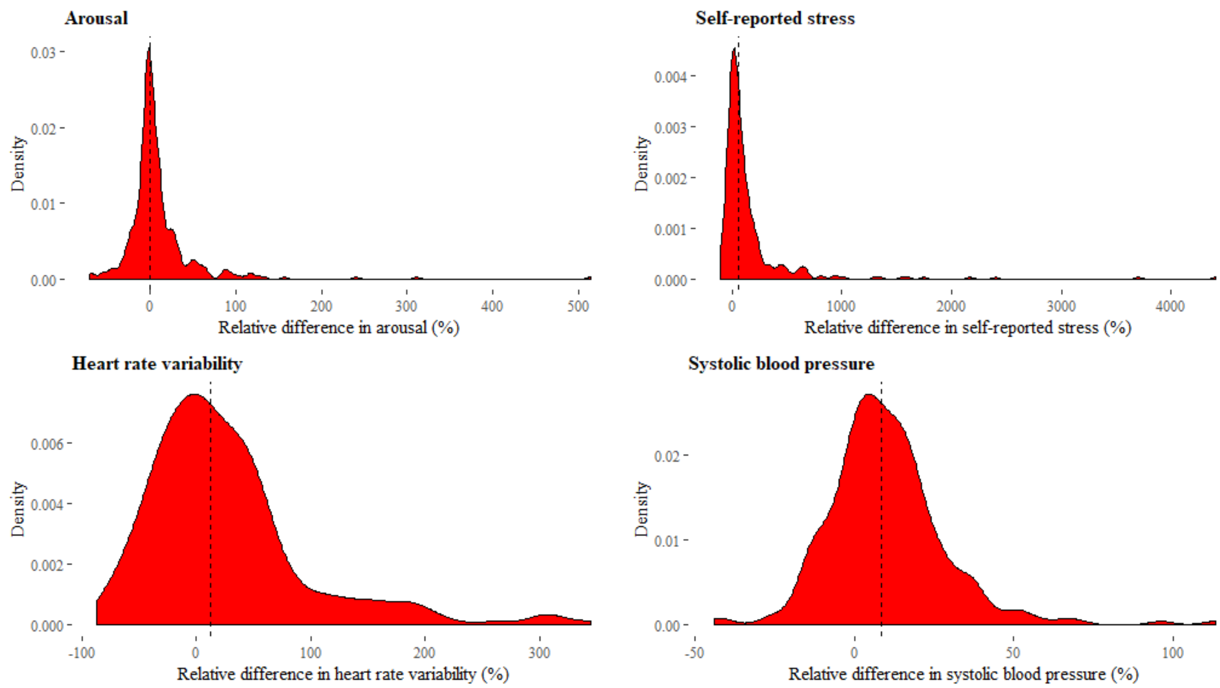
3 of Appendix A. These FFPI personality traits will be used as the predictor variables for the analysis.

3.2.2 Target variables

It is essential to assign donors to different groups of perceived stress during a blood donation because no labels have so far been assigned to which donors perceive “high” levels of stress versus those who do not. Assigning donors to a particular group is a critical adjustment to the DISTRESS dataset due to the different classification algorithms that will be used. As it is not possible to perform classification algorithms on non-labeled data, the labelling process is a requirement for a successful experimental procedure. Based on the previous research, the moment of needle insertion is a highly salient event for physiological and psychological stress responses during a blood donation (Hoogerwerf et al. 2017; 2018), Therefore, the relative difference between measurement one (arrival at the donation center) and measurement four (needle insertion) will be taken as the value to categorize donors in different groups of perceived stress. Self-reported stress and arousal will be used as the target variables that represent psychological stress during the blood donation, and heart rate variability and systolic blood pressure are the target variables representing physiological stress (Table 4 of Appendix A). By having these four target variables, this study could eventually give an overview of similarities or differences between the different outcomes of predicting physiological stress and psychological during a blood donation on the basis of personality.

The median split is regarded as a method that could be used to convert continuous variables (arousal, self-reported stress, heart rate variability and systolic blood pressure) into dichotomous variables (categorical variables with two groups). In accordance with DeCoster et al. (2011), the median split tends to give the best results when the data show in general a symmetric distribution. Figure 2 illustrates that all target variables show an overall symmetric distribution of the relative difference between measurement one and measurement four. The dashed vertical line indicates the median value of the target variables. This vertical line shows that the right-hand side and left-hand side of the median look quite symmetric. Therefore, it is justified to categorize the donors for each target variable into two groups (less stress / more stress) by using the median split.

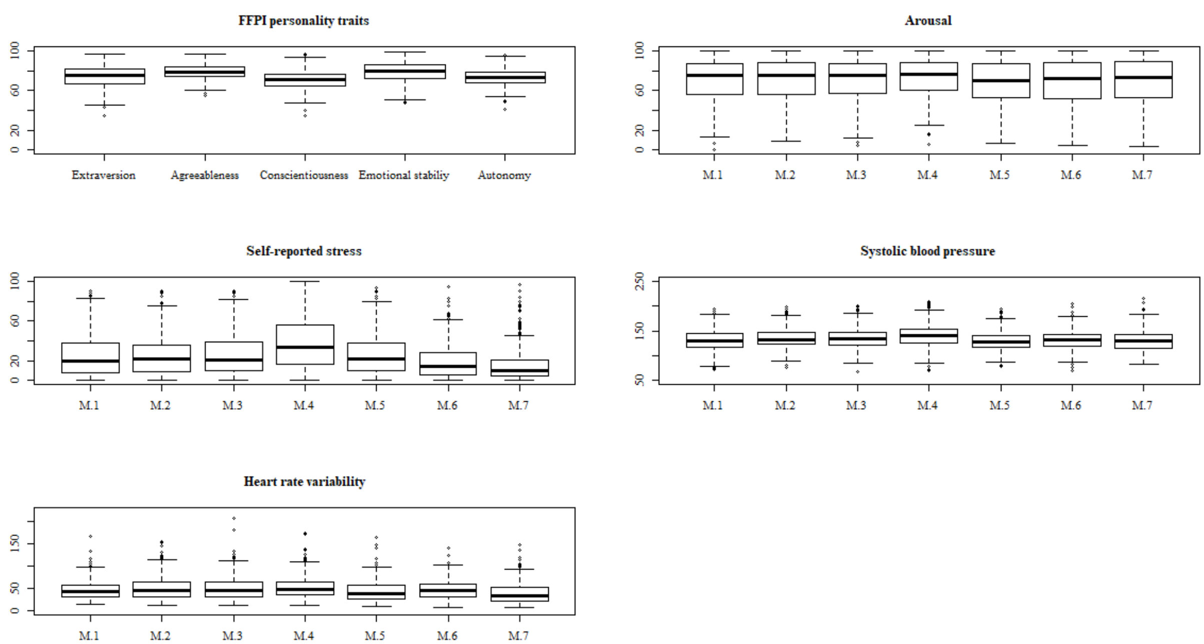
Figure 2 Distribution of all target variables towards needle insertion (the relative difference between measurement 1 and measurement 4).



3.3 Outliers

It often happens that some observations differ from the majority of observations in a certain dataset. Such observations are called outliers. Outliers may be classified as errors, could have been recorded under exceptional circumstances, or belong to a different population (Rousseeuw & Hubert 2011). As a result, these outliers can be the cause of poor performances of an algorithm. It is important to detect such outliers in order to understand the eventual outcomes in a more precise way. Besides, the detection of outliers can result into the removal of certain observations. It is possible that some observations in the DISTRESS dataset are regarded as impossible values that should be excluded from the analysis. Tukey's boxplot is one way of detecting certain outliers. In this plot, a box is drawn from the first quartile (Q_1) to the third quartile (Q_3) of the dataset. The points outside the interval $[Q_1 - 1.5\ IQR,\ Q_3 + 1.5\ IQR]$ are traditionally marked as outliers. By using Tukey's boxplot analysis on the DISTRESS dataset, eleven observations have been found that qualify as impossible values or extreme values (+ 5 standard deviations from the mean). Two observations of arousal and two observations of self-reported stress have been replaced by missing values because they were outside the range of the possible interval [0 - 100]. Three observations of systolic blood pressure have been replaced because they were equal to zero. Finally, four observations of heart rate variability have been removed because their values were higher than + 5 standard deviations from the mean. Figure 3 shows multiple boxplots of the relevant features, without taking the impossible values and extreme values into account. When interpreting this figure, it becomes clear that in particular self-reported stress, systolic blood pressure and heart rate variability have many outliers. However, these outliers have been analyzed and can still be considered as realistic observations. Therefore, these outliers have not been excluded from the analysis.

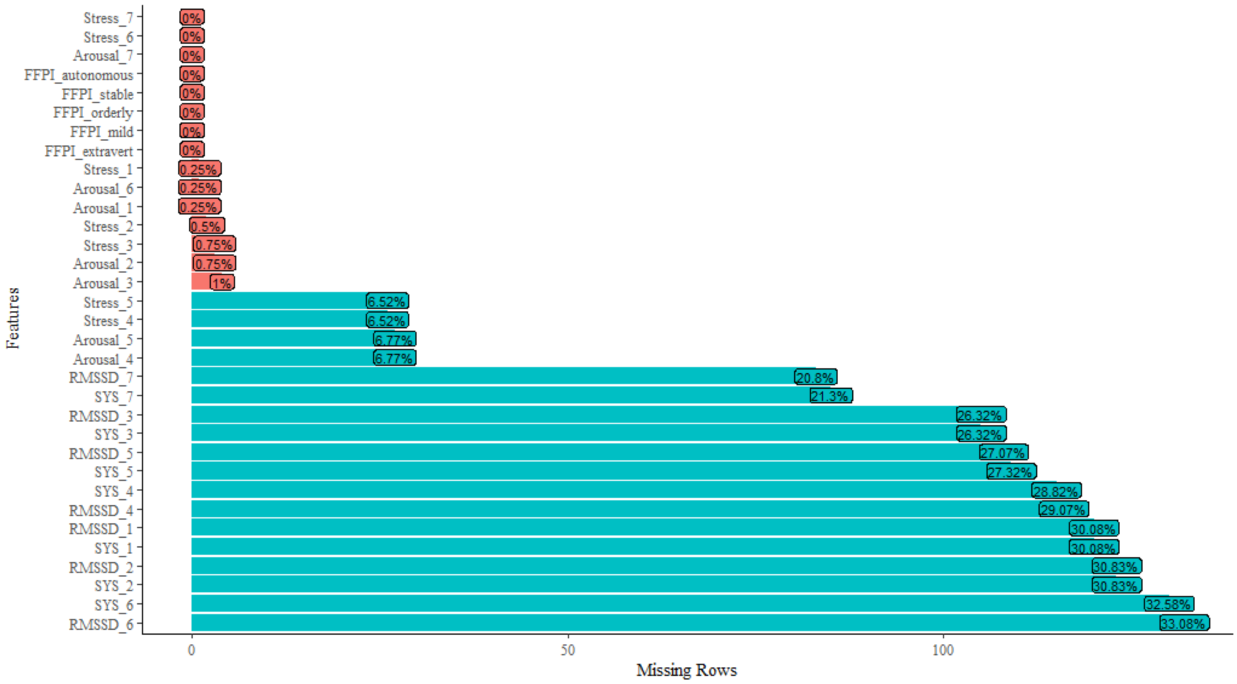
Figure 3 Detect outliers by using boxplots for all relevant features in the DISTRESS dataset.



3.4 Missing data

Multiple measurements have been taken during the process of the blood donation. However, when looking at the DISTRESS dataset, quite some features have missing data. The following figure illustrates the missing data for all features (target variables and predictor variables) among all donors that have participated in the DISTRESS study. The cause of these missing data is that measurements were not correctly taken at the key moments of the blood donation or that the data has not properly been recorded in the DISTRESS dataset.

Figure 4 Missing data in the DISTRESS dataset (N = 399).



When interpreting Figure 4, the missing values reduce the available data up to 33.08 %. Especially measurements of systolic blood pressure (SYS) and heart rate variability (RMSSD) contain many missing data. A solution that could be used to limit the reduction of the available dataset is to use missing data imputation techniques (Schafer & Graham 2002). However, considering that the missing data for some target variables covers up to 33.08 % of the entire dataset, missing data imputation techniques will result in a considerable bias. Therefore, when predicting certain target variables, this study solely uses the complete cases of the target variables. This results into a reduction of the DISTRESS data for analysis. The following table gives an overview of the data that is available for predicting certain target variables on the basis of personality traits.

Table 1 Overview of the available data for predicting the target variables.

Target variable (less stress / more stress)	**Available data from the DISTRESS study (N = 399)**
Arousal	N = 371 (93 %)
Self-reported stress	N = 360 (90 %)
Heart rate variability	N = 218 (55 %)
Systolic blood pressure	N = 218 (55 %)

3.5 Machine learning algorithms

3.5.1 Justification of the algorithms used

Since this study is focusing on predicting experienced stress during the blood donation on the basis of personality traits, binary classification algorithms will be used on the DISTRESS data. This means that the output will be classified into two groups (less stress / more stress).

The 'No Free Lunch' theorem states that, given no prior knowledge of the prediction problem, no single algorithm works best for every problem (Domingos, 2012). Therefore, this study has performed multiple classification algorithms to analyze whether experienced stress during a blood donation can be predicted on the basis of personality. Also, the dilemma of Occam's razor is used to give a justification for the algorithms used. Occam's razor is usually described as the problem-solving principle which states that simpler solutions are better than complex ones. Unfortunately, in prediction, accuracy and simplicity (interpretability) usually come into conflict with each other (Breiman 2001). An example is the use of decision trees for prediction. A decision tree is excellent for interpretation but is not as good in prediction compared to more complex models like random forests or support vector machines. As a result, this study will both use algorithms that are good for interpretability (decision trees) as algorithms that usually give in general higher accuracy rates, but are more challenging to interpret (random forests and support vector machines).

3.5.2 Software

The different machine learning algorithms will be operated in the software environment R (version 3.5.1), where the caret() package will be used to perform the algorithms. The caret() package is regarded as the ideal package in the process of training, testing, tuning and evaluating machine learning algorithms.

3.5.3 Decision trees

A decision tree is regarded as a supervised machine learning algorithm that could be used to solve a binary classification problem. The main concept of a decision tree is that it takes one predictor variable at a time and tests a binary condition. The decision tree starts with a root node that represents the entire population of the data sample (all blood donors). In order to build the decision tree, there are decision nodes that represent tests of a binary condition (yes / no). By testing the different conditions, multiple splits are made in the data that result into subsets that contain instances (donors) with similar values. Eventually, decision nodes do not split anymore, which are called the terminal nodes. These terminal nodes represent a certain class label (less stress / more stress). The main advantages of decision trees are that they can be displayed graphically, are easily interpreted and can handle qualitative predictors without the need of creating dummy variables (James at al. 2013). However, decision trees do not have the same level of predictive accuracy as other classification algorithms that will be described in the next sections.

3.5.4 Random forests

Decision trees are regarded as notoriously noisy models (Hastie et al. 2009). The problem is that decision trees can suffer from high variance. This means that if we would split up the dataset into two parts at random, and fit a decision tree on both parts, the results would be quite different. Approaches like bagging and random forests can reduce this risk of high variance by producing multiple trees on the same data which are then combined to give a single consensus prediction (James et al. 2013). However, random forests are regarded as an improvement of bagging because it decorrelates the many trees that have been produced. In this approach, successive decision trees are grown by introducing a random element into their construction (Breiman 2001). At each decision node, the algorithm will choose a random sample of m predictor variables that are considered as the split candidates of the full set of p predictor variables (5 in the DISTRESS data). In general, $m \approx \sqrt{p}$ ($\sqrt{5}$ in the DISTRESS data) is used as the number of predictor variables that split the decision node of the tree (James et al. 2013). Selecting a random sample of predictor variables at each decision node results into decorrelating the many trees that have been grown. Therefore, this study will perform random forests on the DISTRESS data to see whether it produces higher accuracy rates in comparison to decision trees.

3.5.5 Support vector machines

The support vector machines are regarded as a more sophisticated classification algorithm. By using this approach, each data item (donor) will be plotted in an n-dimensional space, where n is the number of personality traits. A support vector machine algorithm is a representation of the data items (donors) as points in an n-dimensional space, mapped so that the data items (donors) of the separate categories are divided by a hyperplane. New data items (donors) are then mapped into that same space and predicted to belong to a category (less stress / more stress) based on which side of the hyperplane they fall. The beautiful thing about support vector machines is that it can be both used as a linear and non-linear classifier. In order to decide whether a linear or non-linear support vector machines should be used, is to plot the data. However, since it is impossible to visualize the data when the dimension of the feature space is more than three (which is the case with the five personalities in the DISTRESS data), the target variables are plotted for each personality trait. Figures 9, 10, 11 and 12 of Appendix B show that there is no linear relationship between the predictor and target variables when plotting all target variables (stress responses) on the prediction variables (personality traits). Based on these visualizations, the assumption is made that a linear hyperplane cannot separate the two classes within the target variables (less stress / more stress). Therefore, non-linear support vector machines will be performed on the DISTRESS data.

3.6 Evaluation method

3.6.1 Metric

Many metrics could be used to evaluate the performance of the model. Accuracy is regarded as a metric that is commonly used to assess the performance of classification models. Therefore, accuracy will be used to find the best-fitted classification model on the data. The accuracy is a metric that displays the overall performance of the model by making the following calculation:

$$\frac{True\ Positives + True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives}$$

3.6.2 Cross-validation

In machine learning, the most common approach to estimate the prediction error is the use of cross-validation. Cross-validation can be used to estimate the prediction error rate of an algorithm by holding out a subset of the training observations (validation set) from the training process, and then applying the algorithm on those held out observations (James et al. 2013). The goal of cross-validation is to test the algorithm's ability to predict unseen data that was not used during the training and validation procedure. By using cross-validation, the risk of overfitting will largely disappear, and the estimate of the algorithm's performance will be very close to the actual out-of-sample performance (Yarkoni & Westfall, 2017). In this study, leave one out cross-validation ("LOOCV") will be used when performing the classification algorithms. By using the LOOCV approach, one data point will be held out of the training data and is used as the validation set. This will be repeated for all possible separations in the training data, and then the misclassification error (sum of misclassified observations) is averaged for all trials. In this study, first, a split will be made in the DISTRESS dataset, where the proportion of the training set and test set (unseen data) will respectively consist of 70% and 30% of the data. Consequently, the classification algorithms will be trained and validated on the training set by using LOOCV. By using this approach, the accuracy of the model on the training set can eventually be compared with the accuracy of the model on the test set.

4 Results

4.1 Prediction results

The following table presents the results of the three binary classification algorithms that have been performed to predict which donors are experiencing less or more stress during a blood donation. All three classification algorithms have been performed by using LOOCV.

Table 2 Accuracy rates for predicting who is experiencing less or more stress during a blood donation for different psychological and physiological stress responses. DT = Decision trees, RF = Random forests, SVM = Support vector machines.

Algorithms	**Arousal**		**Self-reported stress**		**Heart rate variability**		**Systolic blood pressure**	
Accuracy on	Train	Test	Train	Test	Train	Test	Train	Test
DT	0.5172	0.4909	0.6548	0.4722	0.5974	0.4375	0.4805	0.4844
RF	0.5479	0.5272	0.5516	0.4815	0.5584	0.4219	0.5390	0.4688
SVM	0.5402	0.5545	0.5476	0.5000	0.5974	0.4219	0.5390	0.4531

First, in order to find the best possible model for the decision trees, Table 5 of Appendix C illustrates what the most optimal depths of the decision trees were for predicting experienced stress during a blood donation. By performing the best possible decision trees, Table 2 demonstrates that the accuracy rates of the decision trees fluctuate around 0.5000 on both the training set and the test set.

Second, to find the best possible model for the random forests, the number of randomly selected variables and the number of trees have been tuned for predicting experienced stress during a blood donation. Table 6 of Appendix C illustrates what the most optimal randomly selected predictor variables were for predicting experienced stress during a blood donation. By performing the best possible models, Table 2 shows that the accuracy rates of the random forests also fluctuate around 0.5000 on both the training set and the test set for all stress responses. Therefore, the random forests produce approximately the same results as the decision trees.

Finally, the support vector machines have been performed to predict experienced stress during a blood donation. To find the most optimal model for the support vector machines, Table 7 of Appendix C demonstrates the most optimal values of the parameters. Besides, Table 2 shows that the accuracy rates of the support vector machines fluctuate around 0.5000 for both the training as the test set.

When looking at the overall accuracy rates of the three classification algorithms, the performances are approximately the same (around 0.5000). There are no results to be found which indicate that the experience of stress during a blood donation can be predicted on the basis of personality. This means that all machine learning algorithms perform as good as a ‘flipping a coin’ (pure chance) when predicting which donors will experience less or more stress during a blood donation.

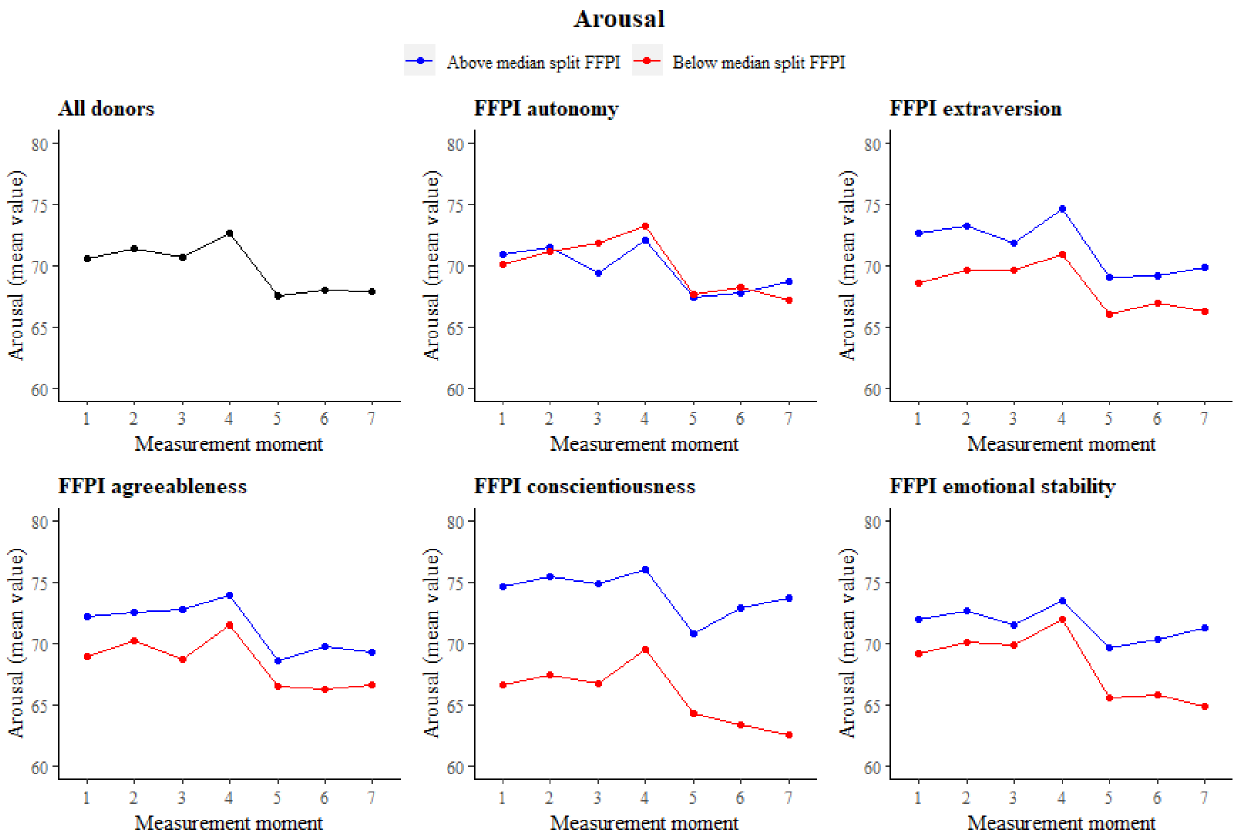
4.2 Post-hoc analysis

A post-hoc analysis is done on the DISTRESS dataset to get a better view on how to interpret the results of the binary classification algorithms. Since the prediction results indicate that the experience of stress during a blood donation cannot be predicted on the basis of personality traits, this section will demonstrate a descriptive analysis on how the different stress responses fluctuate during the whole course of a blood donation (taking all seven measurements into account). This might give a better understanding of why the accuracy rates are low when predicting experienced stress during a blood donation on the basis of personality (where only measurement one and four have been taken into account).

4.2.1 Arousal

Figure 5 demonstrates that donors who score above the median split on extraversion, agreeableness, conscientiousness and emotional stability have higher levels of arousal during the blood donation in comparison to donors that score below the median split on these personality traits. When looking at autonomy, there are little differences in perceived arousal during the whole course of a blood donation. The biggest difference in perceived arousal is seen at conscientiousness. Donors who score above the median split on conscientiousness have a higher score on arousal during the whole course of a blood donation compared to the donors who score below the median split. The course of arousal after the moment of needle insertion (between measurement five and seven) shows different patterns for donors who score above the median split on all personality traits in comparison to the donors that score below the median split on these personality traits. When looking at the overall course of arousal during the blood donation, all personality traits show that there is an increase in arousal towards needle insertion and a decrease of arousal after the moment of needle insertion.

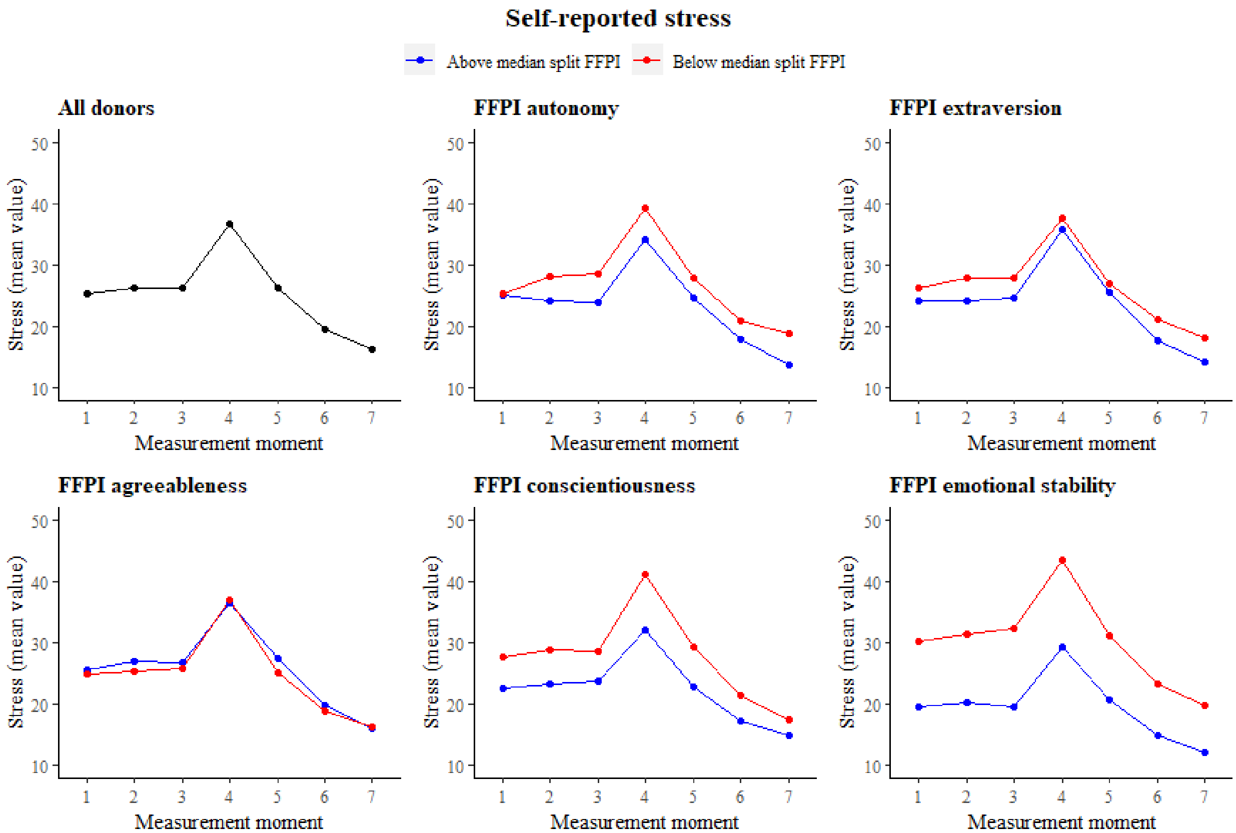
Figure 5 Mean scores of arousal during a blood donation for all FFPI personality traits.



4.2.2 Self-reported stress

Figure 6 illustrates the course of self-reported stress during a blood donation for different personalities. This figure shows that all donors who score below the median split on autonomy, conscientiousness and emotional stability have a higher score on self-reported stress during the whole course of the blood donation in comparison to donors who score below the median split on these personality traits. For extraversion and agreeableness, the course of self-reported stress during a blood donation looks almost identical. By examining the overall course of self-reported stress during the blood donation for different personalities, all personality traits show that there is a steady increase in arousal towards needle insertion and a solid decline in self-reported stress after the moment of needle insertion.

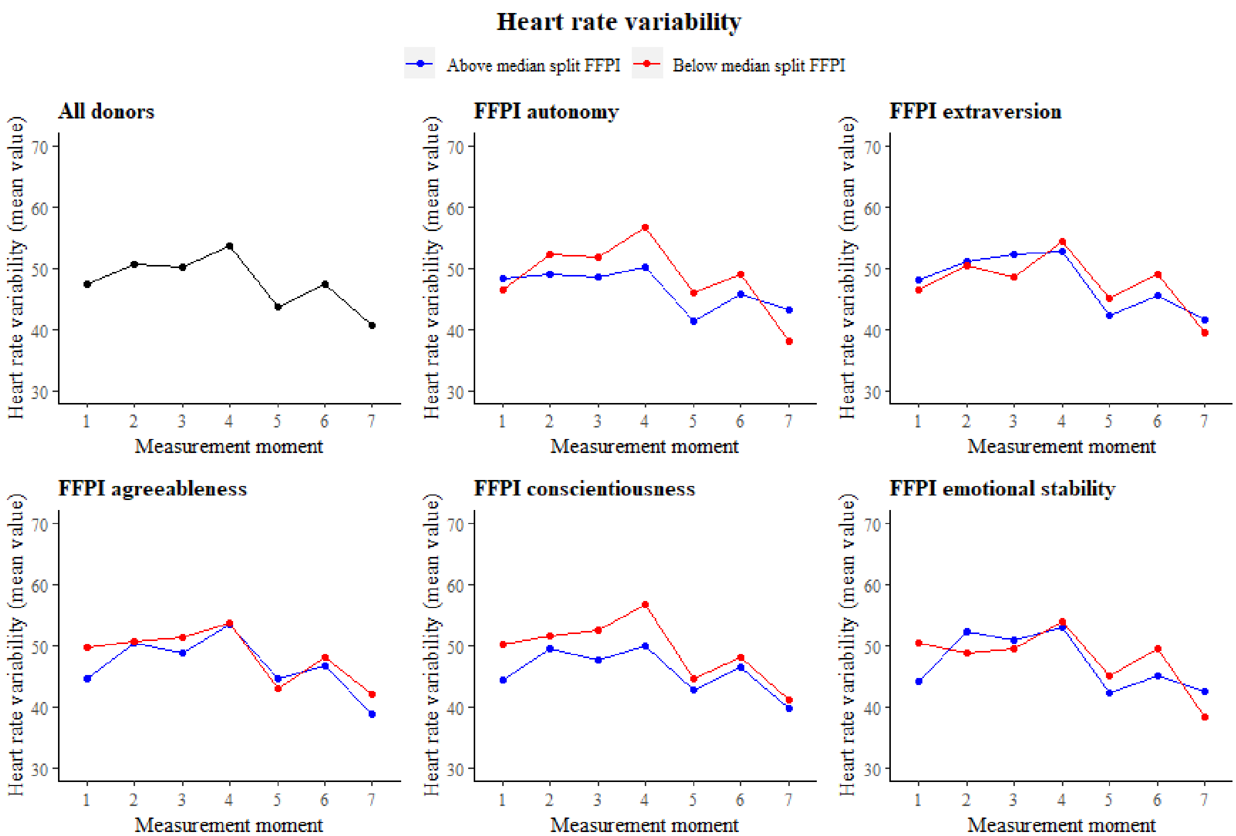
Figure 6 Mean scores of self-reported stress during a blood donation for all FFPI personality traits.



4.2.3 Heart rate variability

When looking at the course of heart rate variability during a blood donation in Figure 7, it is noticeable that the course of heart rate variability for almost all personality traits shows a different pattern towards needle insertion (measurement four), whereas the course of heart rate variability after needle insertion shows a pattern that is generally the same among the different groups of personality. In addition, the peaks in heart rate variability are steeper between measurement three and four for donors who score above the median split on conscientiousness, autonomy and extraversion compared to those who score below the median split on these personality traits.

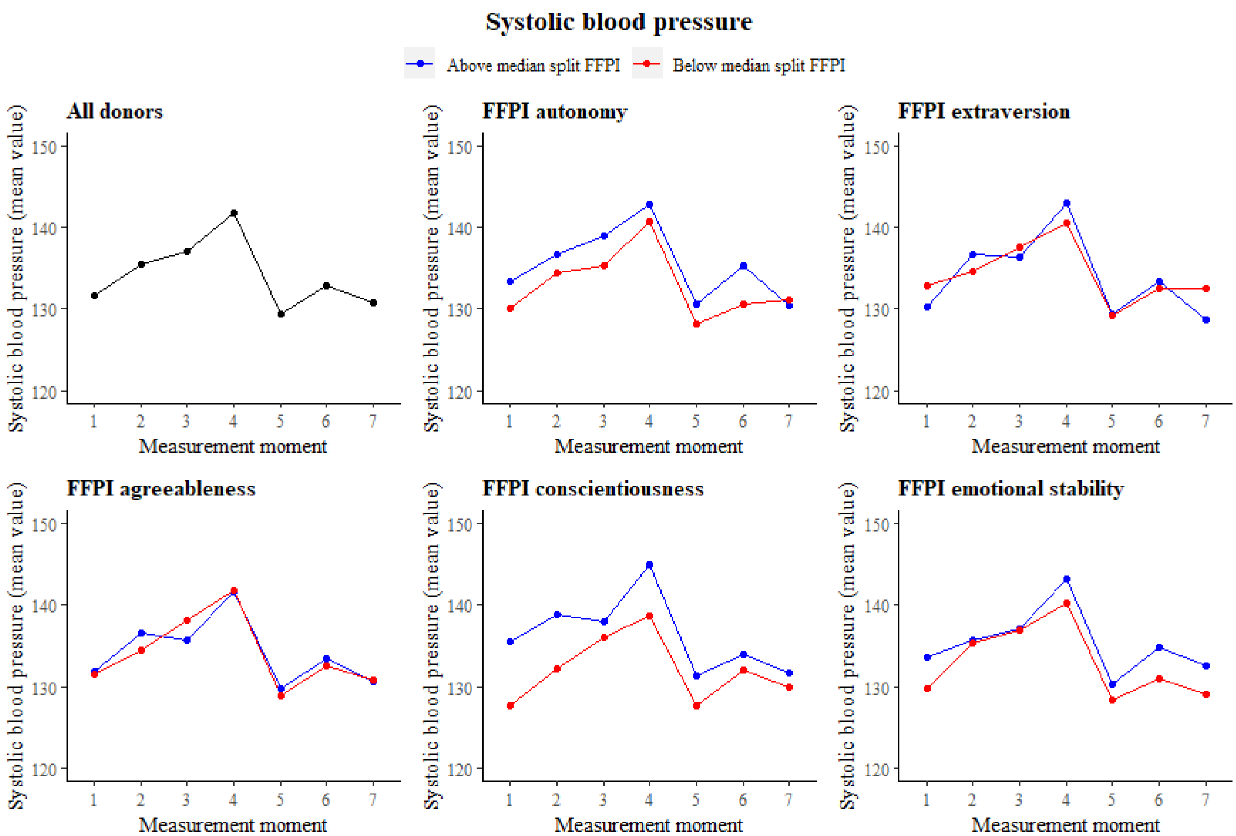
Figure 7 Mean scores of heart rate variability during a blood donation for all FFPI personality traits.



4.2.4 Systolic blood pressure

Figure 8 displays the course of systolic blood pressure during blood donation. When looking at the differences for systolic blood pressure during the blood donation, virtually no differences can be found between donors who score above the median split on the personality traits and those who score below the median split on the personality traits. The two groups lie very close to each other during the whole course of the blood donation, and the pattern looks approximately the same.

Figure 8 Mean scores of systolic blood pressure during a blood donation for all FFPI personality traits.



5 Discussion

The goal of this study was to investigate the influence of personality traits on the experience of stress during a blood donation. By performing several machine learning models, the results indicate that the experience of psychological and physiological stress during a blood donation cannot be predicted on the basis of personality. Several potential explanations for these results are described below.

A first explanation that could be given for the low prediction rates of personality on the experience of stress during a blood donation is that a blood donation is experienced as such a highly stressful event for all donors, that personality traits do not make any difference in the experience of stress.

A second explanation for the low prediction value of personality on the experience of stress during a blood donation is that this study tries to predict human behavior (stress). Human behavior is regarded as a structurally highly complex concept. This is because human behavior is influenced by physical, emotional, cognitive and social factors that interact with each other in complex ways (Schmidt 2005). Therefore, it is perhaps not surprising that the predictive value of personality is low when predicting experienced stress during a blood donation

However, when considering the previous literature on this topic, the findings of this study are inconsistent with the general expectations that follow from previous research. By examining the previous research, several studies have investigated the relationship between personality and the experience of stress outside the context of a blood donation setting (Ebstrup et al. 2011; Garbarino et al. 2014, Roohafza et al. 2016; Saklofske et al. 2011). The overall findings of previous research indicate that there is a relationship between certain personality traits and the experience of stress. Especially neuroticism is regarded as a personality trait that is positively related to the experience of stress. On the contrary, Garbarino et al. (2014) found very weak relationships between personality and the experience of stress. Therefore, the findings of this study accompanied by the findings of Garbarino et al. (2014) give more reasons for conducting future research on the influence of personality on the experience of stress.

5.1 Study limitations

5.1.1 Limitations of the dataset

A third explanation for the low prediction rates is that the DISTRESS dataset consists of some limitations.

The first limitation of the DISTRESS dataset is related to the proportion of missing data. As visualized in section 3.4, the proportion of missing data for some measurements of target variables is 33.05 %. These missing data have not been imputed by imputation techniques to avoid the risk of creating a considerable bias in the data. However, by not imputing these missing data, there is the risk of losing datapoints with valuable information. In this study, much valuable information is lost when predicting the physiological stress responses (N = 218 instead of N = 399).

The second limitation is related to the size of the DISTRESS dataset. The original dataset had 399 donors available for analysis. When applying machine learning algorithms, it would be ideal to have more donors available in order to train the model in a better way. However, since the results indicate that there is no problem of overfitting (the accuracy rates on the training and test set are approximately the same), there is no need for increasing the number of data points.

The third limitation of the DISTRESS dataset is related to measuring personality traits by conducting an FPPI questionnaire. This could be seen as a limitation because personality traits are in general measured by conducting a Big Five personality questionnaire (Ebstrup et al. 2011; Garbarino et al. 2014, Roohafza et al. 2016; Saklofske et al. 2011). As a result, the use of FFPI personality traits could contribute to diverging results when comparing the findings of this study with the findings of previous research. Nonetheless, the FFPI personality traits show many similarities with the Big Five personality traits (Hendriks et al. 1999). Therefore, it is still justified to compare the findings of this study with the findings of previous research. Still, for future research, it is recommended to use the Big Five personality questionnaire in order to reduce the risk of not being able to compare the findings with previous research.

5.1.2 Limitations of the analysis

A fourth explanation for the low prediction rates is that the analysis consists of several limitations.

The first limitation of the analysis relates to the use of the median split for classifying donors into two groups (less stress / more stress). This was done for the sake of simplicity and for applying several binary classification algorithms. However, much information is lost by cutting the dataset in half by using the median split (Altman & Royston 2006). This results into a reduction of the predictive power to detect a relationship between the predictor variables and the target variables. Additionally, donors that are close to but on opposite sides of the median split are characterized as being very different rather than very similar. This is precisely the case when classifying donors into two groups for arousal and self-reported stress (see Figure 2 of section 3.2.2).

Furthermore, this study has defined the experience of stress during a blood donation as the relative difference of stress responses (psychological and physiological) between arrival at the donation center and the moment of needle insertion. This definition was given due to the findings of Hoogerwerf et al. (2017; 2018). However, when looking at the results of the machine learning algorithms, no relation can be found between personality and the experience of stress during a blood donation. When looking at the post-hoc analysis in section 4.2, an analysis that takes all seven measurements into account would be more suitable instead of only using two measurements. This study reduces much of the information by only using the two measurements instead of the seven measurements available in the DISTRESS dataset. The machine learning models that have been applied to this data can therefore be regarded as models that are too simple to explain the underlying complexities of the DISTRESS data. Other longitudinal analyses like multilevel models (Steele 2008) could provide more reliable results because they are better equipped to longitudinal data.

5.2 Future research

The findings of this study demand for conducting future research on the relationship between personality and the experience of stress during a blood donation. Since psychological and physiological stress responses fluctuate during the whole process of a blood donation, these additional studies should conduct longitudinal analyses on the relationship between personality and the experience of stress during a blood donation.

Additionally, this study also demands for doing future research on other possible causes of the experience of stress during a blood donation. These subsequent studies should include more predictor variables (not only personality traits) to predict the experience of stress during a blood donation. It is also important that these future studies apply longitudinal analyses to search for possible causes of the experience of stress during a blood donation.

6 Conclusion

The main goal of this study was to investigate what the influence of personality traits were on the experience of stress during a blood donation. By performing several binary classification algorithms, we tried to predict which donors experience less or more stress during a blood donation. When performing these algorithms, the results indicate that these binary classification algorithms are as good as ‘flipping a coin’ (pure chance) when predicting who is experiencing less or more stress during a blood donation. Therefore, this study concludes that the experience of stress cannot be predicted on the basis of personality.

7 References

Altman, D., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal*, 1080.

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, 199-231.

Byrne, N., & Ditto, B. (2005). Alexithymia, cardiovascular reactivity, and symptom reporting during blood donation. *Psychosomatic Medicine*, 471-475.

DeCoster, J., Gallucci, M., & Iselin, A. (2011). Best Practices for Using Median Splits, Artificial Categorization, and their Continuous Alternatives. *Journal of Experimental Psychopathology*, 197-209.

Ditto, B., & France, C. (2006). Vasovagal symptoms mediate the relationship between predonation anxiety and subsequent blood donation in female volunteers. *Transfusion*, 1006-1010.

Domingos, P. (2012). A Few Useful Things to Know about Machine Learning. *Communications of the ACM*, 78-87.

Ebstrup, J., Eplov, L., Pisinger, C., & Jørgensen, T. (2011). Association between the Five Factor personality traits and percieved stress: is the effect mediated by general self-efficacy? *Anxiety, Stress, & Coping*, 407-419.

Garbarino, S., Chiorri, C., & Magnavita, N. (2014). Personality traits and the Five-Factor Model are associated with work-related stress in special force police officers. *Int Arch Occup Environ Health*, 295-306.

Hanson, S., & France, C. (2009). Social support attenuates presyncopal reactions to blood donation. *Transfusion*, 843-850.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning.* New York: Springer-Verlag New York Inc.

Hendriks, J., Hofstee, W., & De Raad, B. (1999). The Five-Factor Personality Inventory (FFPI). *Personality and Individual Differences*, 207-325.

Hoogerwerf, M., Veldhuizen, I., Merz, E., de Kort, W., Frings-Dresen, M., & Sluiter, J. (2017). Psychological and hormonal stress response patterns during a blood donation. *Vox Sanguinis*, 733-743.

Hoogerwerf, M., Veldhuizen, I., Tarvainen, M., Merz, E., Huis in 't Veld, E., de Kort, W., . . . Frings-Dresen, M. (2018). Physiological stress response patterns during a blood donation. *Vox Sanguinis*, 357-367.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* New York: Springer.

Kumar, U. (2016). *The Wiley Handbook of Personality Assessment.* Chichester: John Wiley & Sons, Ltd.

Roohafza, H., Feizi, A., Afshar, H., Mazaheri, M., Behnamfar, O., Hassanzadeh-Keshsteli, A., & Adibi, P. (2016). Path analysis of relationship among personality, perceived stress, coping, social support, and psychological outcomes. *World Journal of Psychiatry*, 248-256.

Rousseeuw, P., & Hubert, M. (2011). Robust statistics for outlier detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 73-79.

Saklofske, D., Austin, E., Mastoras, S., Beaton, L., & Osborne, S. (2012). Relationship of personality, affect, emotional intelligence and coping with student stress and academic success: Different patterns of association for stress and success. *Learning and Individual Differences*, 251-257.

Sanquin. (2019, April 9). *Missie en visie*. Retrieved from Sanquin: https://www.sanquin.nl/over-sanquin/ons-verhaal/missie-en-visie?r178:u_u_i_d=3b22d105-601d-48ff-9d62-87242acada7d

Schafer, J., & Graham, J. (2002). Missing Data: Our View of the State of Art. *Psychological Methods*, 147-177.

Schmidt, B. (2005). Human factors in complex systmes: The modelling of human behaviour. *Conference Proceedings, 19th European Conference on Modelling and Simulation*, (pp. 5-14). Riga, Latvia.

Steele, F. (2008). Multilevel models for longitudinal data. *Journal of the Royal Statistical Scoiety: series A (statistics in society)*, 5-19.

Ulrich, R., Simons, R., & Miles, M. (2003). Effects of environmental simulations and television on blood donor stress. *Journal of Architectural and Planning Research*, 38-47.

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, 1-23.

Appendix A

Table 3 Predictor variables in the DISTRESS dataset.

Predictor variable	Description	Name in dataset
Extraversion	A scale from 0 to 100 (introvert to extrovert).	FFPI_extravert
Agreeableness	A scale from 0 to 100 (bossy to mild).	FFPI_mild
Conscientiousness	A scale from 0 to 100 (disorderly to orderly).	FFPI_orderly
Emotional stability	A scale from 0 to 100 (unstable to stable).	FFPI_stable
Autonomy	A scale from 0 to 100 (non-autonomous to autonomous).	FFPI_autonomous

Table 4 Target variables in the DISTRESS dataset.

Target variable	Description	Name in dataset
Systolic blood pressure	Measurements of systolic blood pressure during seven key moments of a routine blood donation.	SYS_1, SYS_2, SYS_3, SYS_4, SYS_5, SYS_6, SYS_7
Heart rate variability	Measurements of heart rate variability during seven key moments of a routine blood donation, by taking the root mean square of successive differences in heart rate.	RMSSD_1, RMSSD_2, RMSSD_3, RMSSD_4, RMSSD_5, RMSSD_6, RMSSD_7,
Self-reported stress	Measurements of self-reported stress during seven key moments of a routine blood donation.	Stress_1, Stress_2, Stress_3, Stress_4, Stress_5, Stress_6. Stress_7
Arousal	Measurements of self-reported arousal during seven key moments of a routine blood donation	Arousal_1, Arousal_2, Arousal_3, Arousal_4, Arousal_5, Arousal_6, Arousal_7.

Appendix B

Figure 9 Relationship between arousal towards needle insertion (Arousal_14) and different personality traits.

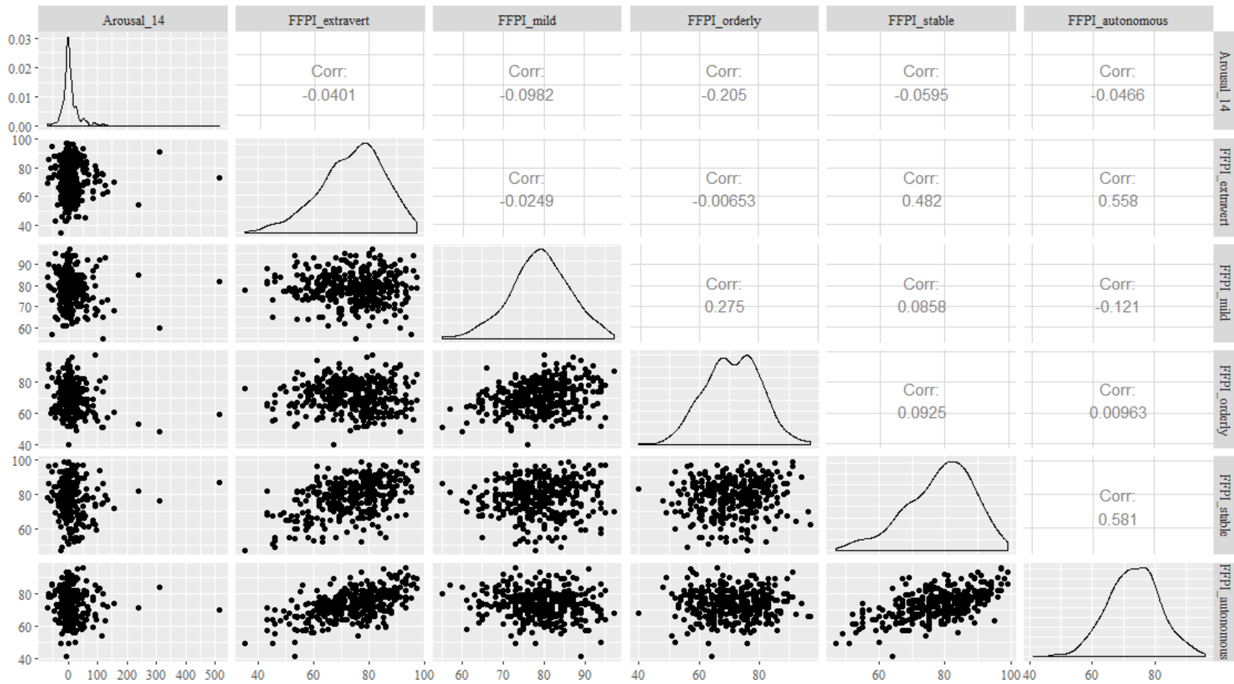


Figure 10 Relationship between self-reported stress towards needle insertion (Stress_14) and different personality traits.

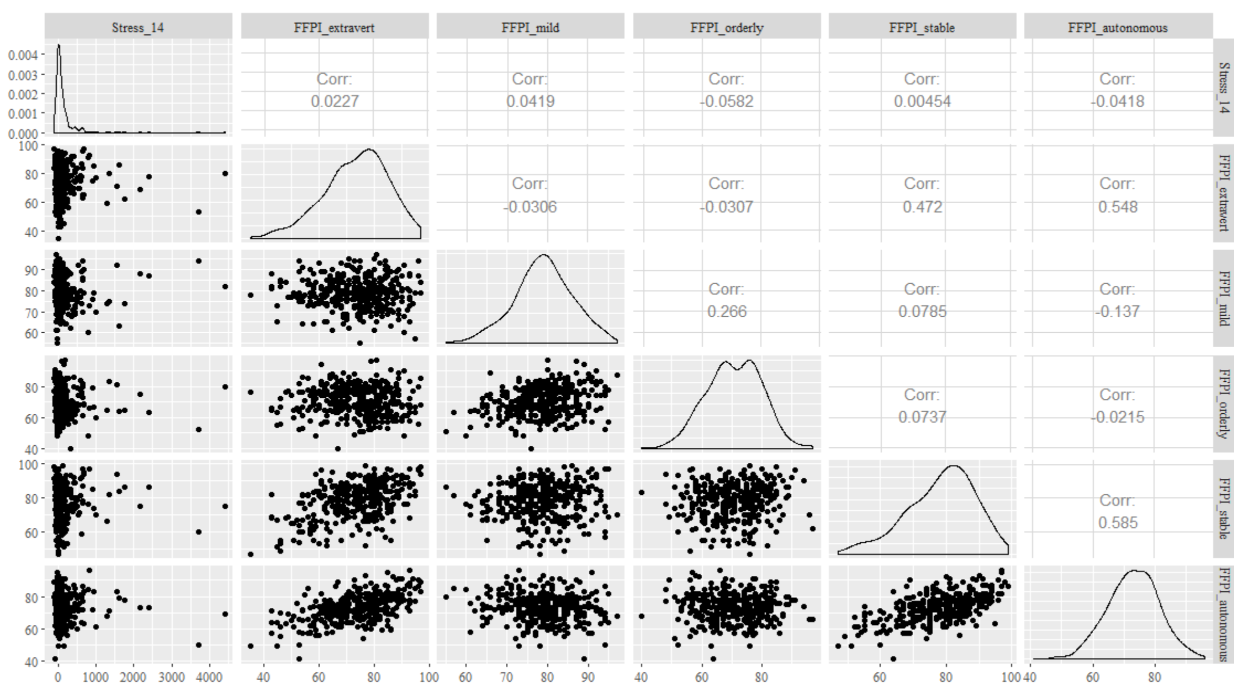


Figure 11 Relationship between heart rate variability towards needle insertion (RMSSD_14) and different personality traits.

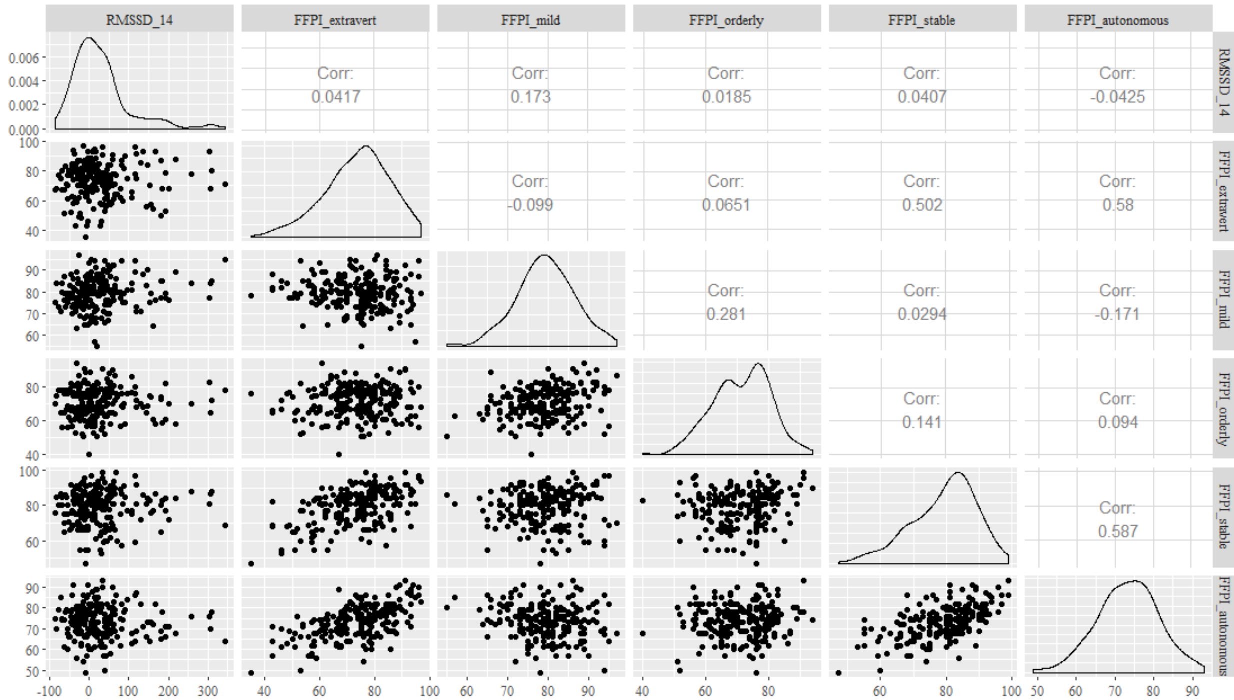
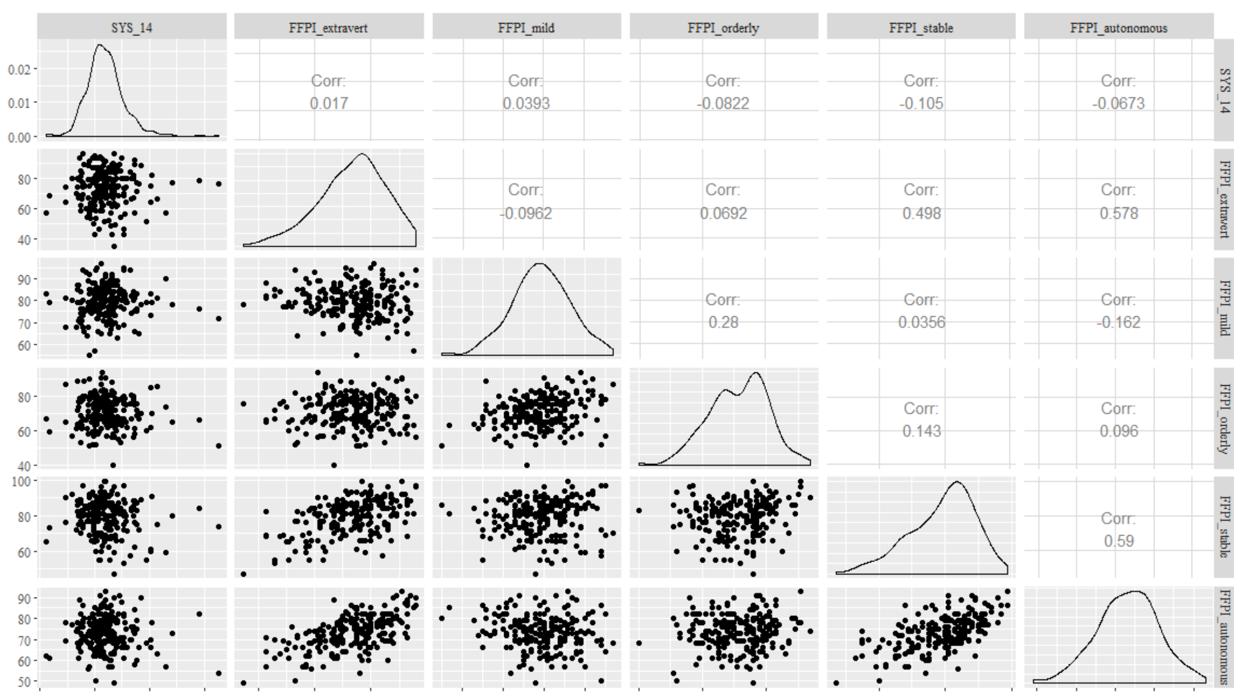


Figure 12 Relationship between systolic blood pressure towards needle insertion (SYS_14) and different personality traits.



Appendix C

To find the best possible model for the decision tree, the depth of the decision tree has been tuned for predicting groups of low stress and more stress for different stress responses during a blood donation. This is done by finding the smallest decision tree (lowest depth of the tree) with the highest LOOCV accuracy. The following table represents the accuracy rates of the decision trees for all stress responses accompanied by the most optimal depth of the tree (depth).

Table 5 Accuracy rates of the most optimal decision trees for predicting psychological and physiological stress responses.

Stress response	depth	Train set	Test set
Arousal	10	0.5172	0.4909
Self-reported stress	3	0.6548	0.4722
Heart rate variability	1	0.5974	0.4375
Systolic blood pressure	1	0.4805	0.4844

In order to find the best possible model for the random forests, the number of randomly selected variables and the number of trees have been tuned for predicting groups of low stress and more stress for different stress responses during a blood donation. The following table represents the accuracy rates of the random forests for all stress responses accompanied by the most optimal number of trees (ntree) and the most optimal number of randomly selected prediction variables (mtry).

Table 6 Accuracy rates of the most optimal random forests for predicting psychological and physiological stress responses.

Stress response	(ntree, mtry)	Train set	Test set
Arousal	(30, 2)	0.5479	0.5272
Self-reported stress	(10, 4)	0.5516	0.4815
Heart rate variability	(100, 3)	0.5584	0.4219
Systolic blood pressure	(30, 3)	0.5390	0.4688

In order to find the best possible model for the support vector machines, the parameters of the support vector machines have been tuned: C and γ. The objective was to identify the best (C, γ) to accurately predict unknown data (i.e. the test set). The most optimal pairs of (C, γ) have been found by performing a 'grid search' on C and γ using LOOCV. Various pairs of (C, γ) have been tried and the one with the highest LOOCV accuracy was chosen. The following table represents the accuracy rates of the support vector machines for all stress responses accompanied by the most optimal values of C and γ.

Table 7 Accuracy rates of the most optimal support vector machines for predicting psychological and physiological stress responses.

Stress response	**(C, γ)**	**Train set**	**Test set**
Arousal	(0.6, 0.2)	0.5402	0.5545
Self-reported stress	(1.1, 0.2)	0.5476	0.5000
Heart rate variability	(0.7, 0.1)	0.5974	0.4219
Systolic blood pressure	(1, 0.025)	0.5390	0.4531