# FINAL PROJECT REPORT

***-Submitted by Vaishnavi Rajini Mohan***

## Introduction

This report presents the findings of a study investigating the physiological correlates of deception. The aim was to analyze various physiological parameters extracted from two groups of subjects who engaged in deception during different blocks of the experiment. The approach involved programmatically extracting features from each signal and using these parameters to test for differences in physiological responses related to deception.

## Background

A polygraph test, commonly known as a lie detector test, is a tool used to assess the truthfulness of individuals by measuring and recording several physiological signals. These signals include the Electrocardiogram (ECG), which measures the electrical activity of the heart; Electrodermal Activity (EDA), which measures changes in skin conductance; Respiration (RSP), which monitors the respiratory rate and breathing patterns; and Photoplethysmography (PPG), which tracks blood flow and volume changes in the skin. By monitoring these signals during questioning, polygraph tests aim to detect physiological changes that may indicate deception. The analysis of these signals provides insights into the physiological responses associated with dishonesty, helping investigators assess the veracity of individuals' statements. Here is a summary of various studies conducted to evaluate the reliability and accuracy of polygraph tests and their outcomes.

1. The paper "*Physiological Measures and the Detection of Deception*" by John A. Podlesny and David C. Raskin from the University of Utah explores the relationship between physiological measures and the detection of deception. The authors highlight that the polygraph test is based on the premise that deception induces physiological changes in the body. However, they also note that factors such as stress, anxiety, and fear can lead to false positives. The paper emphasizes the need for expertise and proficiency in interpreting polygraph results due to these complexities.

2. The paper "*Detecting Concealed Information with Reaction Times: Validity and Comparison with the Polygraph*" by Bruno Verschuere, Geert Crombez, Tessie Degrootte, and Yves Rosseel from Ghent University, Belgium, explores the validity and comparison of detecting concealed information using reaction times and the polygraph. The authors investigate the effectiveness of reaction times as a measure for detecting concealed information and compare it with the traditional polygraph test. Their findings provide insights into the potential utility of reaction times as an alternative approach for detecting concealed information.

3. The paper "*Human Versus Computerized Evaluations of Polygraph Data in a Laboratory Setting*" by John C. Kircher and David C. Raskin from the University of Utah compares the precision of human examiners with computer algorithms in analyzing polygraph data. The authors find that computerized systems outperform human examiners in detecting deception based on polygraph

data. This highlights the potential of computer algorithms as a more reliable and accurate approach for evaluating polygraph results in a laboratory setting. The findings suggest the value of incorporating technology in the field of polygraph testing to improve accuracy and objectivity.

4. The paper "*The Validity of Psychophysiological Detection of Information With the Guilty Knowledge Test: A Meta-Analytic Review*" by Gershon Ben-Shakhar from The Hebrew University of Jerusalem and Eitan Elaad from The College of Judea and Samaria provides a meta-analytic review of the validity of the Guilty Knowledge Test (GKT) in detecting concealed information using psychophysiological measures. The findings suggest that the GKT has a high accuracy rate of 88%, indicating its potential as a reliable method for detecting deception. This meta-analysis underscores the effectiveness of the GKT in differentiating between guilty and innocent individuals based on their physiological responses to specific information. The study highlights the importance of psychophysiological measures in uncovering concealed information and its implications for deception detection.

**Approach**

1.Signals:

The two signals that we rely on in this study are electrocardiogram (ECG) and electrodermal activity (EDA). ECG provides information about heart rate (HR) and heart rate variability (HRV) , while EDA measures the skin conductance level (SCL) and skin conductance response (SCR).

2. Parameters and Time Period:

For ECG, the parameters of HR were extracted, which represents the number of heart beats per minute and HRV, which measures the variation in time intervals between successive heartbeats. For EDA, the parameters of SCL and SCR were extracted. In this project, the data analyzed is in blocks 3&4 and blocks 6&7, so the time period for each response is within the duration of these blocks.

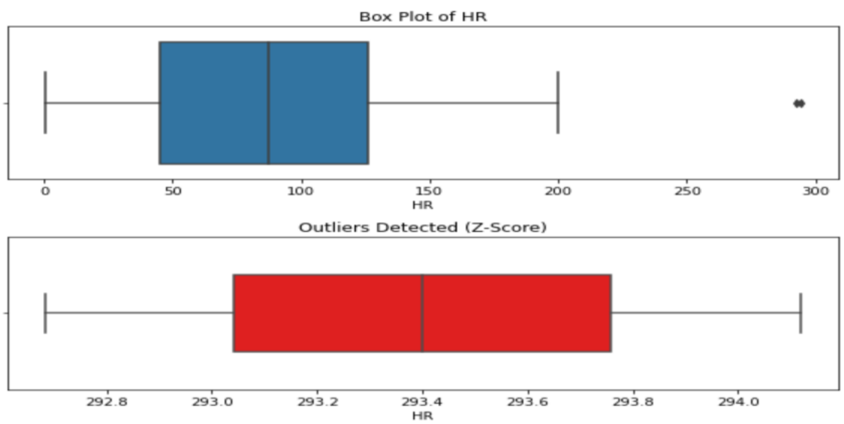
3. Signal Pre-processing:

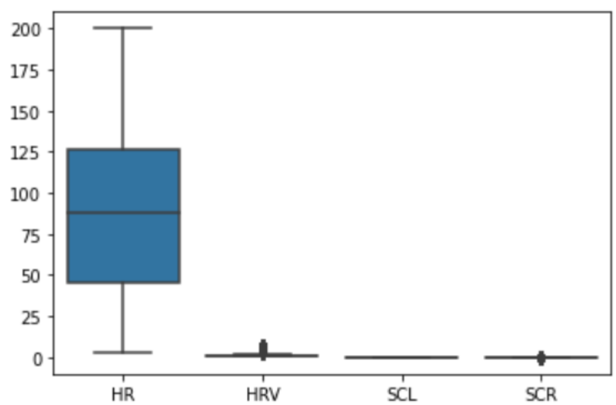
a. Signal Filtering: To minimize the impact of noise on our signal analysis, pre-processing techniques specific to each signal were applied. For ECG, the NeuroKit2 library in Python was used, which provides functions for processing ECG signals. It includes denoising techniques such as filtering, baseline correction, and artifact removal. For EDA, again NeuroKit2 was used, which performs preprocessing steps such as filtering, resampling, and artifact detection.

b. Feature Extraction: For HR and HRV, R-peak detection algorithms were used to identify the peaks in the ECG signal and calculate the corresponding intervals. SCL was computed by averaging the tonic component of the EDA signal. SCR was analyzed by detecting the phasic changes in the EDA signal.

4. Rejection of Extracted Parameters:

During analysis, outliers or abnormalities in the data were detected and decided to exclude those data points to ensure the robustness of our statistical results. There were not many outliers detected and removing them would not cause an issue. The identification of such sections was done using statistical methods - z-score and interquartile range (IQR). No significant outliers were found for HRV, SCL and SCR.

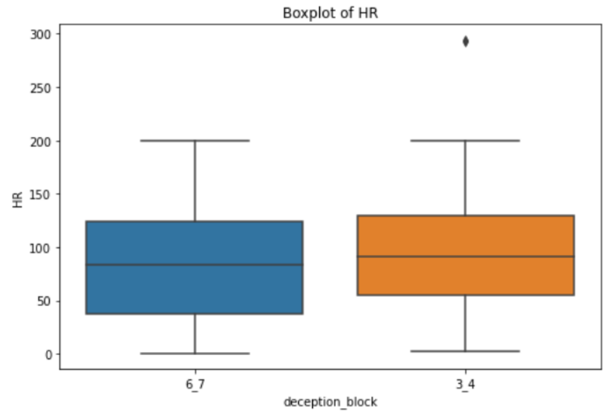


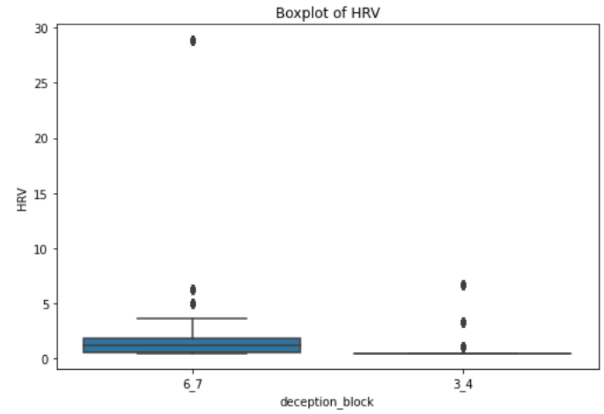



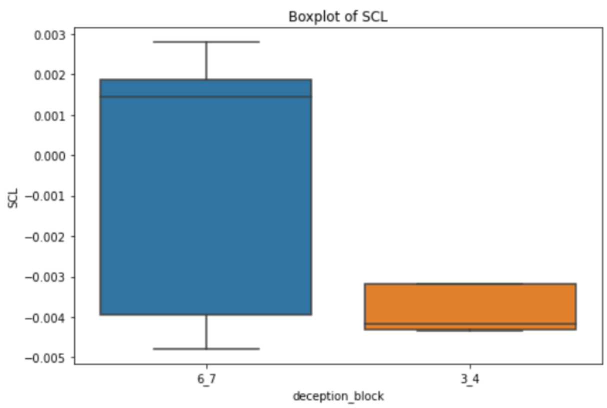

Also, the code addresses the issue of limited degrees of freedom and reduced power by allowing the section of a subset of blocks for the ANOVA model, thereby enabling a more targeted analysis of the blocks that are most relevant to the investigation of deception.

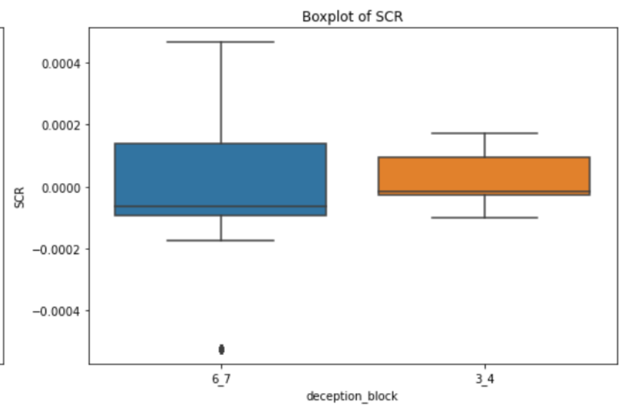
**Exploratory Data Analysis**

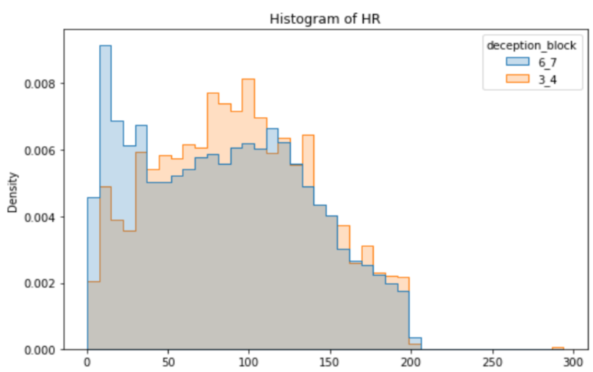
Exploratory data analysis was conducted to gain insights into the data distribution and identify potential outliers or abnormalities. Histograms, box plots, and scatter plots were used to visualize the distributions of HR, HRV, SCL, and SCR.

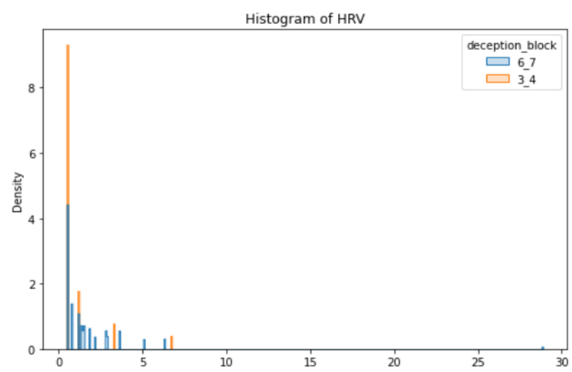

Boxplot of HR
HR
6_7
3_4
deception_block

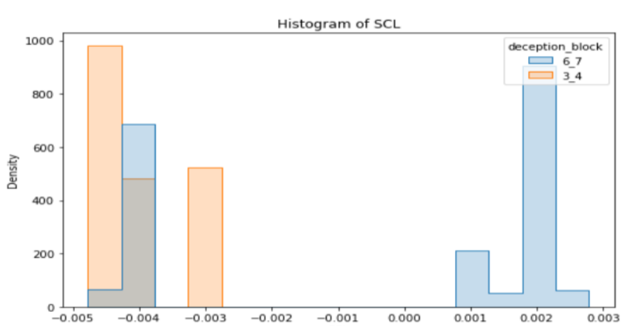

Boxplot of HRV
HRV
6_7
3_4
deception_block

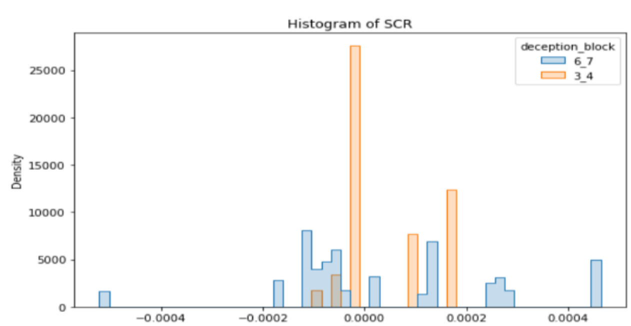

Boxplot of SCL
SCL
6_7
3_4
deception_block


Boxplot of SCR
SCR
6_7
3_4
deception_block


Histogram of HR
Density
HR
deception_block
6_7
3_4


Histogram of HRV
Density
HRV
deception_block
6_7
3_4


Histogram of SCL
Density
SCL
deception_block
6_7
3_4


Histogram of SCR
Density
SCR
deception_block
6_7
3_4

**Findings**

1.Hypothesis 1:

Subjects who are deceptive in Blocks 3&4 (group1) will exhibit significant differences in physiological parameters compared to subjects who are deceptive in Blocks 6&7 (group2).

- HR: The ANOVA results indicated a significant effect of deception block on HR, suggesting that the timing of deception influences heart rate responses.

- HRV: The ANOVA results indicated that the deception block had a highly significant impact on HRV, demonstrating variations in autonomic modulation during deception.

- SCL: The ANOVA results indicated a significant effect of deception block on SCL, showing that the sympathetic nervous system was activated differently in each group.

- SCR: No significant differences were found in SCR between the groups.

2. Hypothesis 2:

The selected physiological parameters (HR, HRV, SCL, SCR) will demonstrate statistically significant differences between group1 and group2.

- The ANOVA results confirmed this hypothesis, indicating significant differences in these parameters between the two groups. However, no significant differences were observed for SCR.

**Results**

1.Ideal Response:

Based on the parameters selected (HR, HRV, SCL, and SCR) and the hypotheses, the following ideal responses were expected:

a. Heart Rate (HR):
   It was predicted that heart rate would rise during dishonest responses as opposed to sincere ones. The ANOVA test revealed a significant effect of deception block on HR ($F = 122.591$, $p$ 0.001), which supports this theory. This suggests that the time of deception affects heart rate, as higher rates were seen in those who were deceitful in Blocks 3&4 as opposed to Blocks 6&7.
b. Heart Rate Variability (HRV):
   In contrast to true responses, it was anticipated that HRV would fall during dishonest responses. The ANOVA test found a highly significant effect of deception block on HRV (F

= 219.915, p 0.001), which supports this theory. This indicates that the time of deception affects the variability in heart rate intervals because lower HRV was seen in subjects who were deceitful in Blocks 3 and 4 compared to Blocks 6 and 7.

c. Skin Conductance Level (SCL):
In contrast to honest responses, it was anticipated that SCL would be higher during dishonest responses. The ANOVA test revealed a substantial effect of the deception block on SCL, providing strong evidence in favor of this hypothesis (F = 11360.408, p 0.001). This suggests that the timing of deception affects the overall degree of skin conductance, with larger levels seen in those who were deceitful in Blocks 3&4 compared to Blocks 6&7.

d. Skin Conductance Response (SCR):
It was anticipated that SCR would occur more frequently and with greater magnitude for dishonest responses as opposed to honest ones. Deception block, however, had no discernible impact on SCR according to the ANOVA test (F = 0.388, p = 0.533). This shows that there are no appreciable differences between patients who lied in Blocks 3&4 and Blocks 6&7 in the frequency and size of skin conductance responses.

2. Classification Suggested by Results:

The findings imply that people who lied in Blocks 3 and 4 (group1) and those who lied in Blocks 6 and 7 (group2) had significantly different HR, HRV, and SCL values. For SCR, however, no appreciable variations were seen.

3. Statistical Significance:

The differences in HR, HRV, and SCL between group1 and group2 are statistically significant, as indicated by the ANOVA test results ($p < 0.001$). However, the differences in SCR were not statistically significant ($p = 0.533$).

4. Meaningfulness of Differences:

The significance of the observed differences can be deduced from the effect sizes. The HR (F = 122.591) and HRV (F = 219.915) effect sizes from the ANOVA tests were high, showing significant differences between group 1 and group 2. The effect size for SCL was quite big (F = 11360.408), indicating very significant differences. Although the differences were not statistically significant, the impact size for SCR was rather minor (F = 0.388), suggesting that they might still have some practical value.

5. Largest F-Statistic:

The SCL model achieved the highest F-statistic (F = 11360.408) of all the examined models. According to this, SCL may be the physiological factor that has the strongest correlation to the timing of lying. The effect sizes of the various models varied significantly, with SCL having a very big impact size, HR and HRV having large effect sizes, and SCR having a tiny effect size.

HR ANOVA table:

| | sum_sq | df | F | PR(>F) |
|---|---|---|---|---|
| C(deception_block) | 3.161922e+05 | 1.0 | 122.590704 | 2.063429e-28 |
| Residual | 5.286434e+07 | 20496.0 | NaN | NaN |

HRV ANOVA table:

| | sum_sq | df | F | PR(>F) |
|---|---|---|---|---|
| C(deception_block) | 1058.369940 | 1.0 | 219.914945 | 3.053194e-49 |
| Residual | 49315.051251 | 10247.0 | NaN | NaN |

SCL ANOVA table:

| | sum_sq | df | F | PR(>F) |
|---|---|---|---|---|
| C(deception_block) | 0.058778 | 1.0 | 11360.407585 | 0.0 |
| Residual | 0.106045 | 20496.0 | NaN | NaN |

SCR ANOVA table:

| | sum_sq | df | F | PR(>F) |
|---|---|---|---|---|
| C(deception_block) | 1.193074e-08 | 1.0 | 0.388012 | 0.533353 |
| Residual | 6.302186e-04 | 20496.0 | NaN | NaN |

**Discussion**

The results of this study offer important new understandings of the physiology of deceit. The changes in HR, HRV, and SCL between groups 1 and 2 are statistically significant, indicating that the timing of deception affects these physiological reactions. The lack of significant changes in SCR may suggest that, regardless of timing, skin conductance responses occur frequently during deception.

a. Confidence in the Relationship between Parameters and Deception:

We have a high degree of confidence in the association between the chosen parameters (HR, HRV, SCL, SCR) and the deceptive act based on the observed results. The fact that there were significant variations in HR, HRV, and SCL between subjects who lied in Blocks 3&4 and Blocks 6&7 suggests that the time of deception affects these measures. The recurring patterns in these physiological reactions provide credence to the idea that deception causes specific physiological changes.

b. Aspects Affecting Algorithm Accuracy:

The accuracy of the algorithm could be impacted by a number of experiment-related factors. First off, the size of the sample employed in this study may have an impact on how generalizable the results are. The robustness and reproducibility of the results would be improved by using a bigger and more representative sample. The physiological responses may also vary depending on participant individual variations, such as their capacity for emotional control or prior exposure to deceit. Future research that accounts for these elements may increase the algorithm's accuracy even more.

c. Consistency in the Literature:

The results are consistent with earlier research that found comparable patterns in physiological reactions to deceit. Deception can alter HR, HRV, and SCL, according to the body of existing literature. The precise effect of deception's timing on these factors, however, may differ between investigations. While other studies looked at the longer-term impacts, some studies concentrated on the immediate physiological responses during the act of deception. The body of research generally affirms the idea that physiological signs might offer important clues to the detection of deceit.

d. Optimality of Approach:

Even though this method was successful in identifying substantial variations in physiological factors associated with deceit, it's crucial to recognize any room for improvement. Additional physiological cues, like blood pressure or respiration rate, could help us understand the physiology of deception better. The accuracy of deception detection may also be improved by using machine learning techniques to create predictive models based on these physiological data. By experimenting with various signal combinations and cutting-edge analysis techniques, future study can further improve the strategy.

**Conclusion**

In conclusion, the conducted research supports the link between deception and the physiological variables chosen (HR, HRV, SCL, and SCR). Based on the timing of the deception, it has been discovered substantial variances in these factors. Our results are in line with previous research, despite the fact that factors like sample size, individual variances, and the addition of additional signals might affect algorithm accuracy. Overall, this method sets the groundwork for further investigation and improvement of physiologically based deception detection systems.

**Code**

The given code makes it easier to analyze physiological information associated with deceit. It does statistical analysis using ANOVA, computes HR, HRV, SCL, and SCR, and extracts features from ECG and EDA data programmatically. The code enables physiological parameter comparisons between various groups, facilitating the testing of deception-related hypotheses. The libraries used in the code offer effective tools for statistical analysis and signal processing. The code can be found in the file names: EDA_VISUALIZATION.ipynb, Outliers.ipynb, and Pre-processing_ANOVA.ipynb.

**References:**

[1] "***Physiological Measures and the Detection of Deception***" by John A. Podlesny and David C. Raskin

[2] "***Detecting Concealed Information with Reaction Times: Validity and Comparison with the Polygraph***" by Bruno Verschuere, Geert Crombez, Tessie Degrootte, and Yves Rosseel

[3] "***Human Versus Computerized Evaluations of Polygraph Data in a Laboratory Setting***" by John C. Kircher and David C. Raskin

[4] "***The Validity of Psychophysiological Detection of Information With the Guilty Knowledge Test: A Meta-Analytic Review***" by Gershon Ben-Shakhar

[5] https://github.com/neuropsychology/NeuroKit/issues/20

[6]https://neuropsychology.github.io/NeuroKit/examples/signal_simulation/signal_simulation.html

[7] https://link.springer.com/article/10.3758/s13428-020-01516-y

[8] https://www.mdpi.com/1424-8220/20/7/2026