Statistical Analysis of the Impact of Stresses and Distractions on Subjects' Physiological Responses

- By team 8: Chonghua Li and Charanteja Sakamuri

1. Introduction

Driving safety is an important issue that everyone has to face in our daily life. Among all the causes of driving incidents, distraction is a very common cause. A research on traffic safety facts published in 2013 states that, 10% of the fatal crashes occurred on U.S. roadways involved distraction, such as manipulating cell phone, adjusting audio and/or climate control, eating or drinking, conversation with other occupants in the car, etc. [1]

In this project, we are going to analyze the impact of different types of stresses on people's physiological responses while driving on a highway. The goal is to find a physiological variable that could be served as a measure of stress and could be used as the explanatory variable in the subsequent analysis. The hypothesis for this experiment is:

- H0: applying stresses do not affect the means of the physiological variables in LD sessions when compared to the ND session.
- H1: applying stresses significantly affect the means of the physiological variables in the LD sessions when compared to the ND session.

2. Experimental Design

This experiment adopts the crossover design. There are 68 subjects. The stressors are categorized into 3 types: cognitive, emotional and sensorimotor. The experimental scenarios are setup in a simulated world where subjects drive along the same freeway for four times with each time under one of the following conditions: no distraction (Normal Drive, ND), cognitive distraction (Cognitive Drive, CD), emotional distraction (Emotional Drive, ED), sensorimotor distraction (Sensorimotor Drive, MD). During the experimental drives, key response variables and several explanatory variables were continuously recorded. The response variables included speed, acceleration, brake force, steering, and lane position signals, while the possible explanatory variables included perinasal EDA, palm EDA, heart rate, and breathing rate. In our project, we will only focus on the effect of stress on the physiological responses.

3. Data Quality Control and Visualization

Data quality is the assessment of data's fitness to serve its purpose in a given context. The data quality control is one of the most essential steps in data analytics. In this project, our goal is to visualize the data, detect outliers and perform proper corrections. The guideline for data cleanness for this project is listed as follow:

- **Palm EDA signal**: [10 - 4,700] kOhm.
- **Perinasal EDA signal**: already cleaned.
- **Heart Rate**: [40 – 120] bpm.
- **Breathing Rate**: [4 – 70] bpm.

The subjects with values outside of the range are excluded from analyzing this signal.

4. Data analysis and results

In CD, ED and MD sessions, the stressors were applied in two separate periods, which divided the entire drive into 5 phases. There is no stress applied in the ND session, therefore, in order to compare with the

loaded sessions, this session was also divided into 5 phases using the time points from the loaded sessions, respectively.

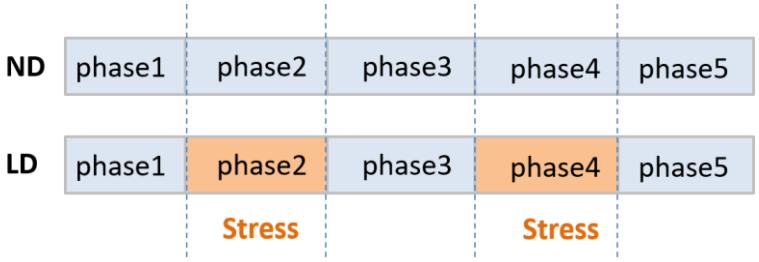


Figure 1. Representation of the experimental design. ND – Normal Drive; LD – Loaded Drive (including CD, MD and ED).

To determine whether the stress has a significant impact on the subject's physiological response, the means of each variable in each phase in the LD session and ND session were calculated for each subject, and the difference of means between the two sessions were also calculated. If the Δmean is significantly different from 0 in phase 2 and phase4 but not significantly different from 0 in the rest of the phases, we could conclude that this variable is a good indicator for stress.

4.1 Normality test

We will use one sample t-test to test the significance of the Δmeans. Since there are four variables, the alpha for t-test need to be subject to the Bonferroni correction. The adjusted alpha = 0.05/4 = 0.125. We first visualized the normality of each set of Δmeans by qq-plot (Figure 2 -5). From the plots, we can see that both the Δmeans for HR and BR are quite normal, which is also supported by Pearson's test (for BR) and Shapiro test (for HR). However, the data for perinasal perspiration signals and palm EDA signals are not normal (Figure 2 and 3, left panels). We next transformed the data by adding a value of abs(min_value)*1.01 to make all the Δmeans positive, and then applied logarithms on the data. The transformed data was then subjected to qq-plot and Shapiro and Pearson's test. Unfortunately, transformation could not make either of the data normal (Figure 2 and 3, right panels; Shapiro test and Pearson's test, data not shown). It's hard to determine if the normality comes from the outliers in the original data or it is a nature of the data.

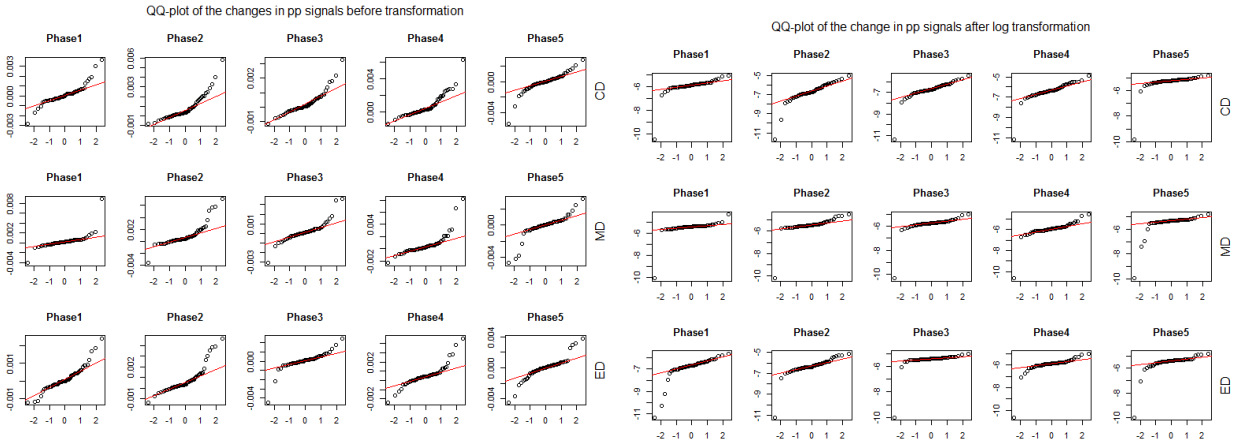


Figure 2. QQ-plots for difference between the means of the perinasal perspiration signals in the ND session and the loaded sessions. Left panel, before transformation; right panel, after logarithms transformation.

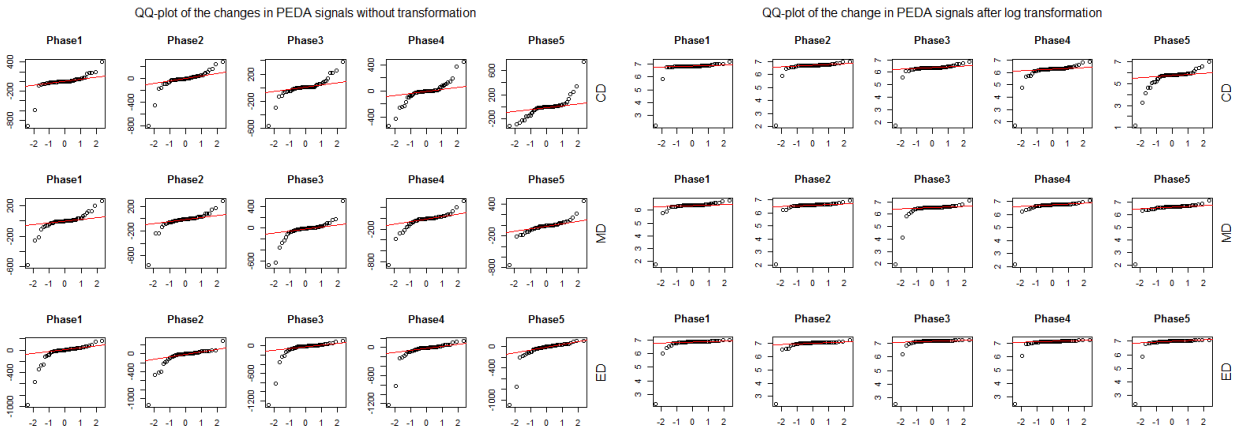


Figure 3. QQ-plots for difference between the means of the palm EDA signals in the ND session and the loaded sessions. Left panel, before transformation; right panel, after log transformation.

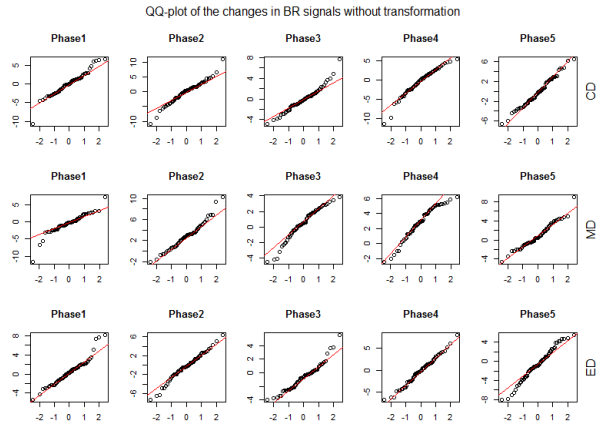


Figure 4. QQ-plots for difference between the means of the Breathing Rate signals in the ND session and the loaded sessions. The data are normally distributed.

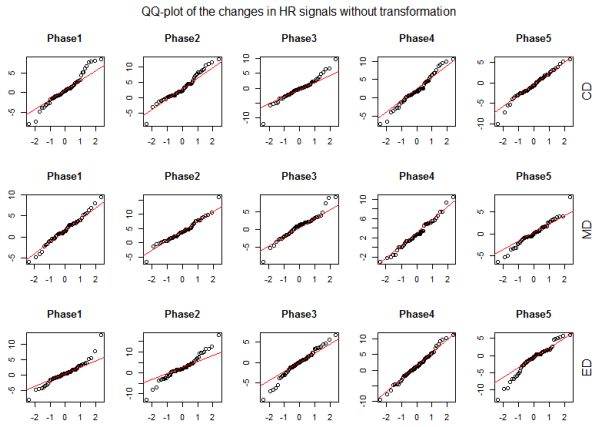


Figure 5. QQ-plots for difference between the means of the Heart Rate signals in the ND session and the loaded sessions. Data are normally distributed.

4.2 T-test and Wilcoxon one-sample signed rank test

Since both the data for HR and BR normally distributed, we applied one sample t-test on the difference of means. The adjusted alpha is equal to 0.0125. The significance of t-test is denoted as follow: 0.0125-0.01, *; 0.01 – 0.001, **; <0.001, ***. The significance of the t-test together with the boxplot for Δmeans in each phase of each session is shown in Figure 6 (BR; p-values in Appendix) and Figure 7 (HR; p-values in Appendix).

For perinasal perspiration signal, both t-test and wilcox test were applied on the untransformed data. The significance of each phase in each session obtained from both tests are consistent (Figure 8; p-values and boxplot for the log transformed data in Appendix). We've also applied t-test on the transformed data, and p-values for all the phases become 0 (see Appendix).

For palm EDA signal, both t-test and wilcox test showed that none of the phases in the untransformed data are significantly different from 0 (Figure 9; p-values in Appendix). We've also applied t-test on the log transformed data, and it turned out that the ∆means in all phases are significantly different from 0 (p-values and the boxplot for the log transformed data in Appendix). Neither of the results indicates that stress has specific impact on palm EDA signals.

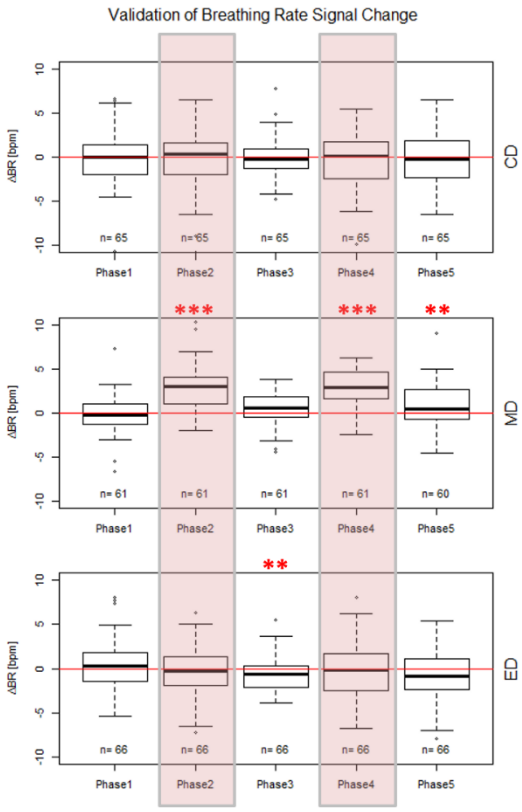


Figure 6. Boxplot for the change of Breathing Rate signal changes between the ND session and the loaded sessions. Y-axis, ∆means between LD session and ND seesion.

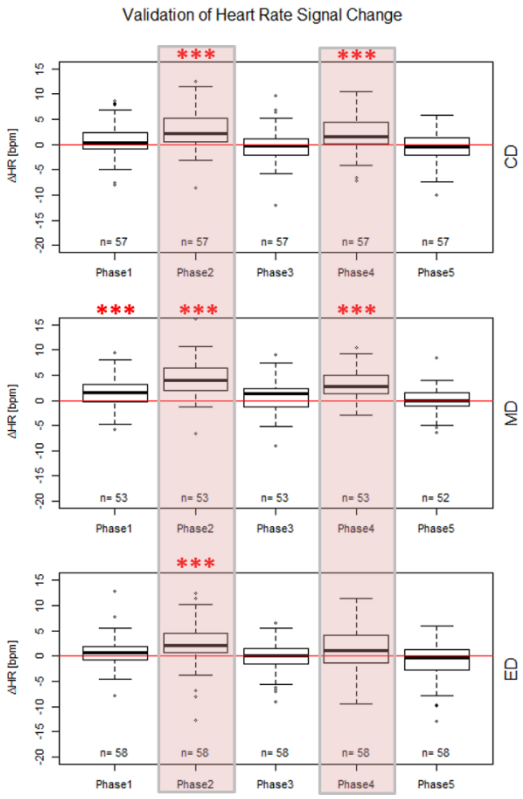


Figure 7. Boxplot for the change of Heart Rate signal changes between the ND session and the loaded sessions. Y-axis, ∆means between LD session and ND seesion.

The results of the significance tests are summarized in Figure 10. We can see that, the perinasal perspiration signal response well to stress, with the mean difference in phase 2 and phase 4 in all three sessions significantly different from 0, while the rest of unstressed phases remain insignificant. The palm EDA signal has no clear correlation to the stress, as discuss above. The breathing rate and heart rate have some response to stress; however, they are not as consistent as the perinasal perspiration signal.

5. Conclusion

In our experiment, we tested if stress can change a subject's physiological response. We found that applying stress to the subjects when they are driving can cause physiological changes in heart rate, breathing rate and perinasal perspiration signals. Among the three responsible variables, perinasal perspiration signal has the most accurate response to stress. We also found that palm EDA has no clear correlation to the stress. In conclusion, perinasal perspiration signal can be used as an indicator of

stress; we could use it as the explanatory variables in the subsequent analysis to determine if stress affect subjects' driving behaviors.

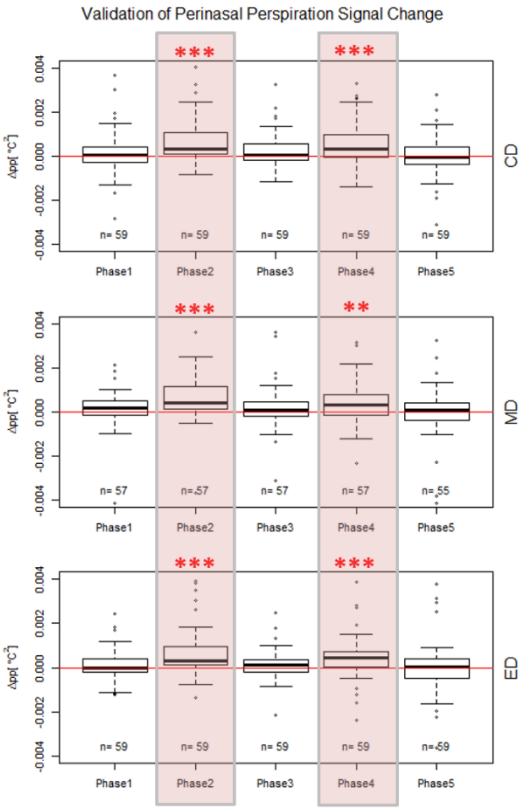


Figure 8. Boxplot for the change of perinasal persipration signal changes between the ND session and the loaded sessions. Y-axis, ∆means between LD session and ND seesion.

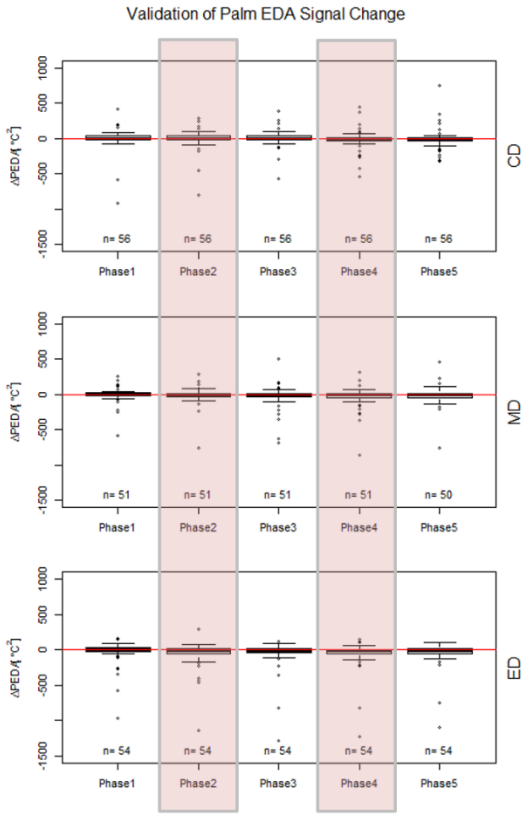


Figure 9. Boxplot for the change of Palm EDA signal changes between the ND session and the loaded sessions. Y-axis, ∆means between LD session and ND seesion.

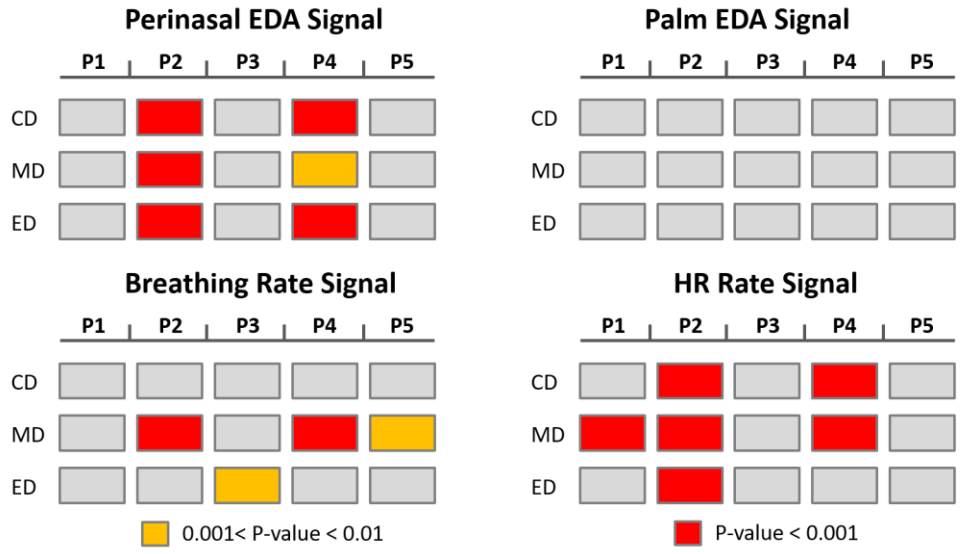


Figure 10. Summary of the significance of t-test on the difference of means in each phase in each session.

Reference:

1. National Center for Statistics and Analysis, Distracted Driving: 2013 Data, in Traffic Safety Research Notes. DOT HS 812 132. April 2015, National Highway Traffic Safety Administration: Washington, D.C.

Appendix

Figure A1. Validation of transformed ΔPP signals

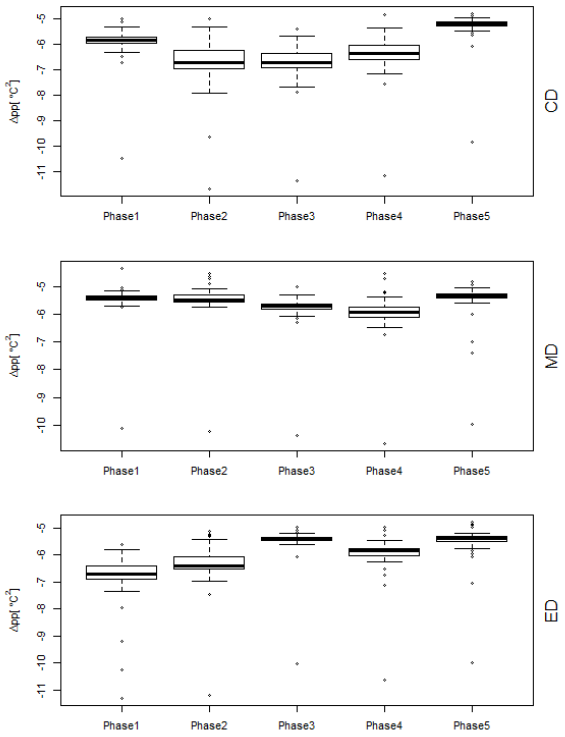
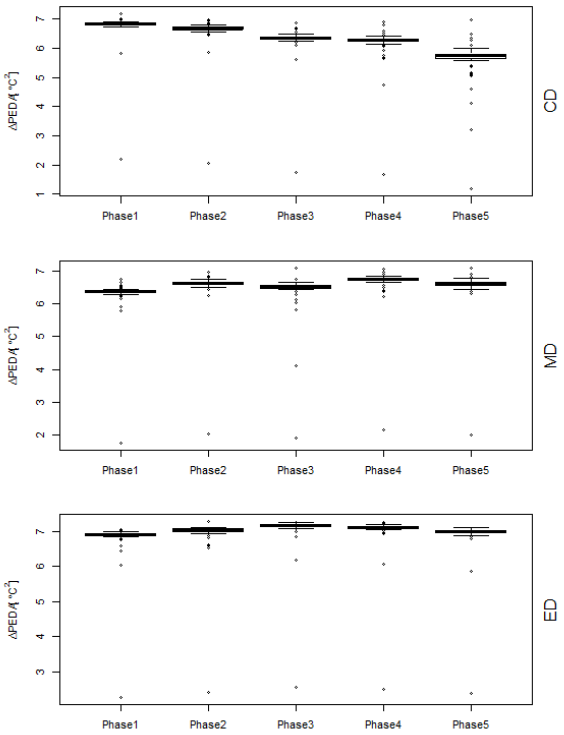


Figure A2. Validation of transformed ΔPEDA signals



p-values of each test (α = 0.0125):

1. pp signals

	Session	phase 1	phase 2	phase 3	phase 4	phase 5
original Δpp Wilcox	CD	0.4171	0.0000	0.0650	0.0001	0.9910
	MD	0.0307	0.0000	0.1550	0.0014	0.7344
	ED	0.3136	0.0000	0.1341	0.0000	0.9368
original Δpp t-test	CD	0.2842	0.0000	0.0135	0.0002	0.6408
	MD	0.1493	0.0001	0.2003	0.0056	0.6318
	ED	0.1456	0.0000	0.6237	0.0006	0.8218
transformed Dpp t-test	CD	0.0000	0.0000	0.0000	0.0000	0.0000
	MD	0.0000	0.0000	0.0000	0.0000	0.0000
	ED	0.0000	0.0000	0.0000	0.0000	0.0000

2. BR signals

Session	phase 1	phase 2	phase 3	phase 4	phase 5
CD	0.8455	0.8226	0.7267	0.3112	0.5269
MD	0.3199	0	0.0639	0	0.0058
ED	0.2724	0.3606	0.0052	0.5786	0.0834

3. HR signals

Session	phase 1	phase 2	phase 3	phase 4	phase 5
CD	0.1255	0	0.6306	0.0001	0.2258
MD	0.0003	0	0.0718	0	0.8299
ED	0.128	0.0004	0.7404	0.0307	0.0416

4. PEDA signals

	Session	phase 1	phase 2	phase 3	phase 4	phase 5
original Δpeda Wilcox	CD	0.6804	0.8416	0.2353	0.1302	0.0479
	MD	0.8257	0.1208	0.1277	0.0300	0.0823
	ED	0.8633	0.1877	0.0733	0.0866	0.1505
original Δpeda t-test	CD	0.6804	0.8416	0.2353	0.1302	0.0479
	MD	0.8257	0.1208	0.1277	0.0300	0.0823
	ED	0.8633	0.1877	0.0733	0.0866	0.1505
transformed Δpeda t-test	CD	0.0000	0.0000	0.0000	0.0000	0.0000
	MD	0.0000	0.0000	0.0000	0.0000	0.0000
	ED	0.0000	0.0000	0.0000	0.0000	0.0000