StandUp_Chen_Tiandian

Tiandian Chen

3/4/2019

Abstract

The main objective of the project study is to determine the responses changes to various emotional states, such as fatigue, stress, and motivation, by reducing the sitting time during a three-month period. In this article, we only discuss one response variable of EMA survey measures, which is the focus level. The data is formally analyzed through conducting linear regression models to indicate the relationship between focus level and other factors that may have effects on it. To estimate the quality of statistical models, the Akaike information criterion (AIC) is applied. According to our final results, the model with changes in standing time as explanatory variable appears to be the best-fitted one. In addition, We conclude that reducing sitting time does not significantly affect focus level.

Introduction

In the wake of high-speed social progress, an increasing number of office jobs with high salaries is provided for people to keep hasty life rhythm and place themselves in a favorable position in the economic-boom society. However, working in an office environment enforce individuals to sit for a long time during work days, which may cause negative impacts on their health. A recent study claims that prolonged sitting that involves very low energy expenditure is adversely associated with physical health outcomes, including premature mortality, type 2 diabetes, and so forth (David, Bethany, Genevieve, Healy & Neville, 2012). In this research, we are interested in examining whether or not a diminution of workplace sitting in an office can result in any changes in mental health.

We use statistical charts, such as box-plots and scatterplots, and summary function in R studio to offer an overall understanding of the data set. The statistical technique, linear regression, is chosen to construct our models with multiple possible explanatory variables combinations, where the best model is provided by AIC backward selection method.

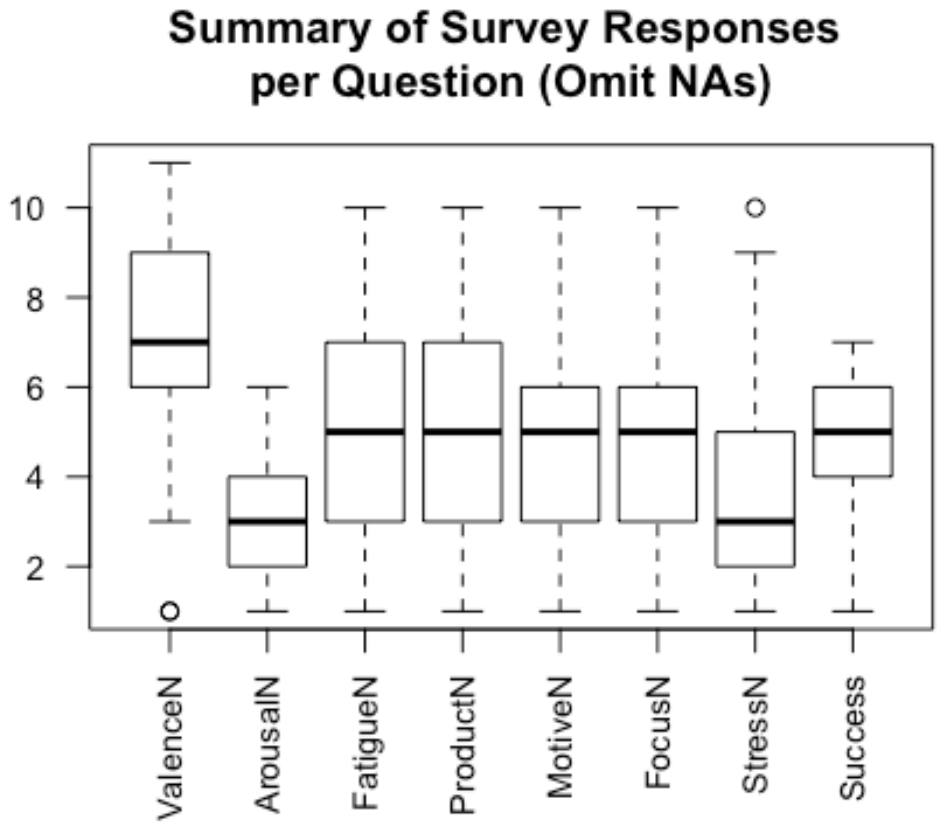
Results

The experiment was conducted with a two-arm randomized waitlist-controlled trail that contains one intervention group and one control group. The two groups were evaluated over a 6-month period, with data collected at baseline, 3- and 6-months. The experimental measurements included 50 participants' daily activity levels and their responses to various emotional states using a mobile-delivered ecological

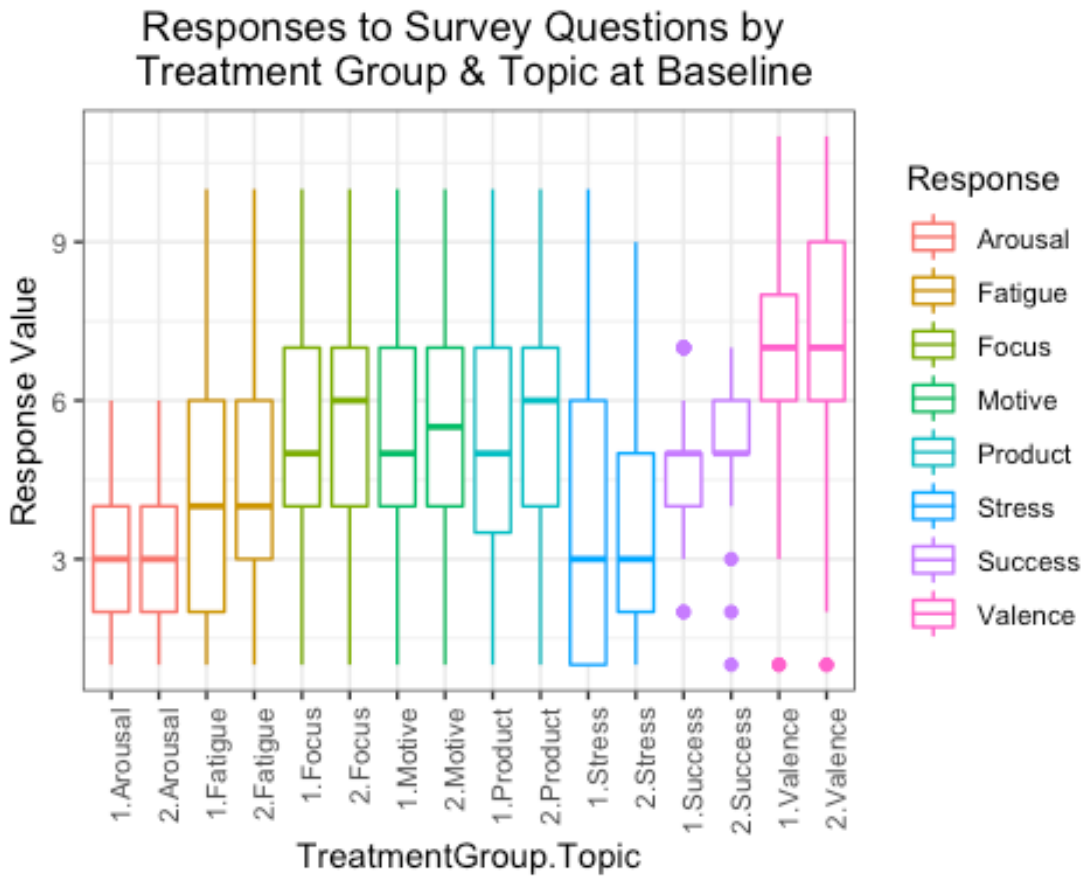
momentary assessment (EMA). In this research, we only compared the differences from baseline to 3-months results.

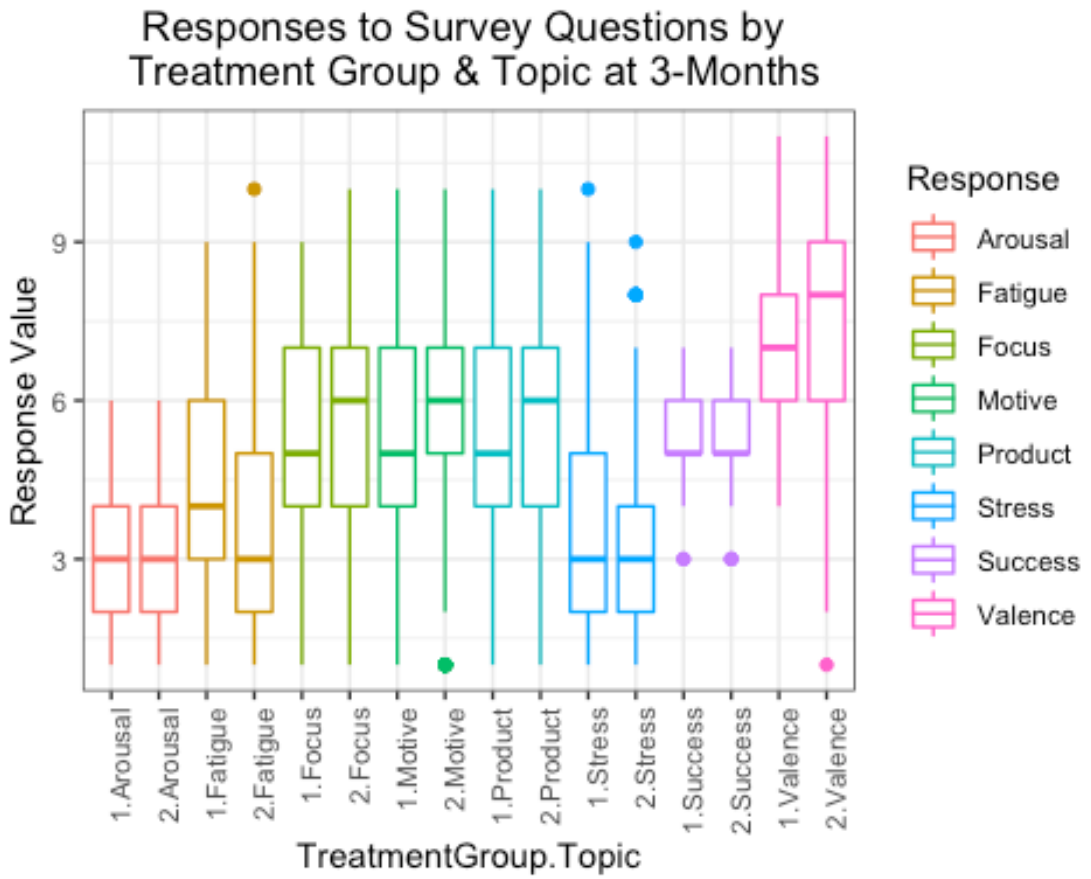
Data Description

```
#boxplot for the summary of survey responses
boxplot(na.omit(dat[18:25]),main = "Summary of Survey Responses \nper Q
uestion (Omit NAs)", cex.axis = 0.9, las = 2)
```



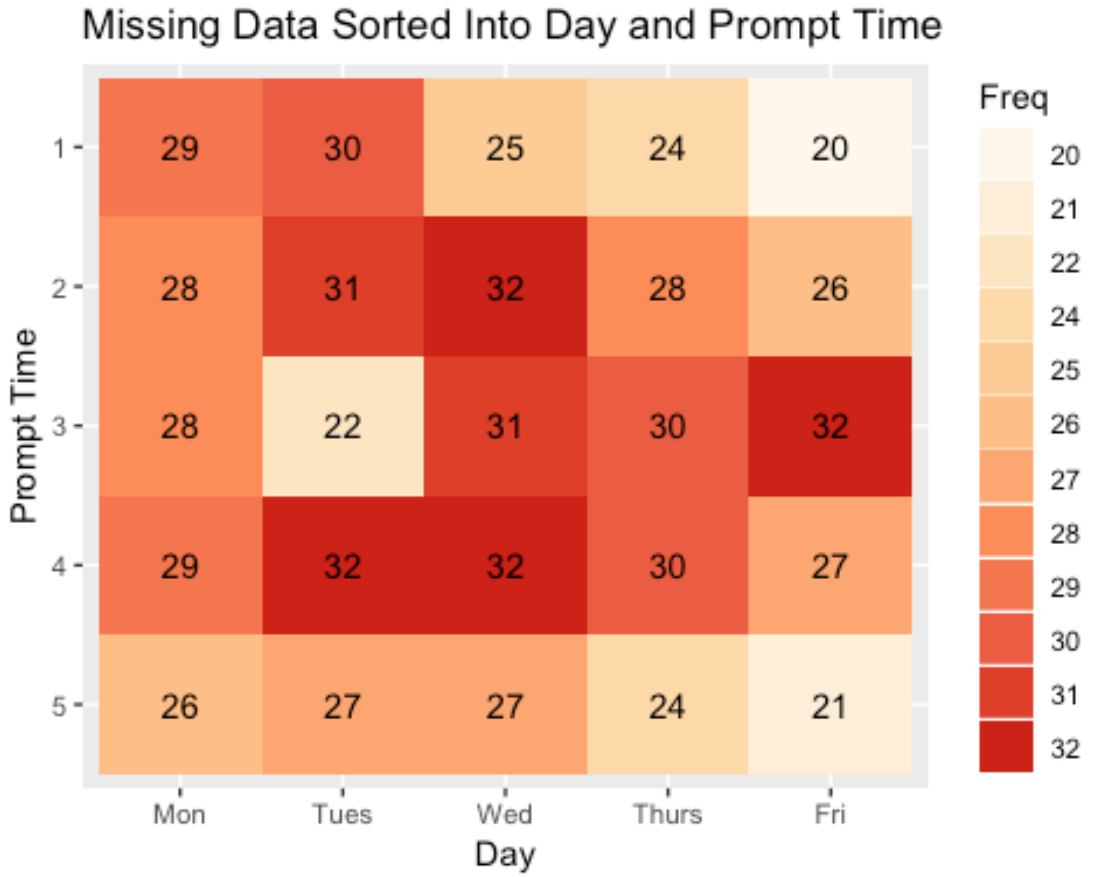
The above box-plot displays the distribution of emotional measures after deleting all of the missing values, providing a big picture of the overall data collection. Both Arousal and Stress are observed with an outlier, which may require a further discussion on it. Note that for variables Arousal, Fatigue, and Stress, a larger level refers to a lower emotional state.


Responses to Survey Questions by
Treatment Group & Topic at Baseline
Response Value
9
6
3
1.Arousal
2.Arousal
1.Fatigue
2.Fatigue
1.Focus
2.Focus
1.Motive
2.Motive
1.Product
2.Product
1.Stress
2.Stress
1.Success
2.Success
1.Valence
2.Valence
TreatmentGroup.Topic
Response
Arousal
Fatigue
Focus
Motive
Product
Stress
Success
Valence



The treatment and control groups' responses to EMA measures at the baseline and

the thrid month are shown respectively by the two boxplots.



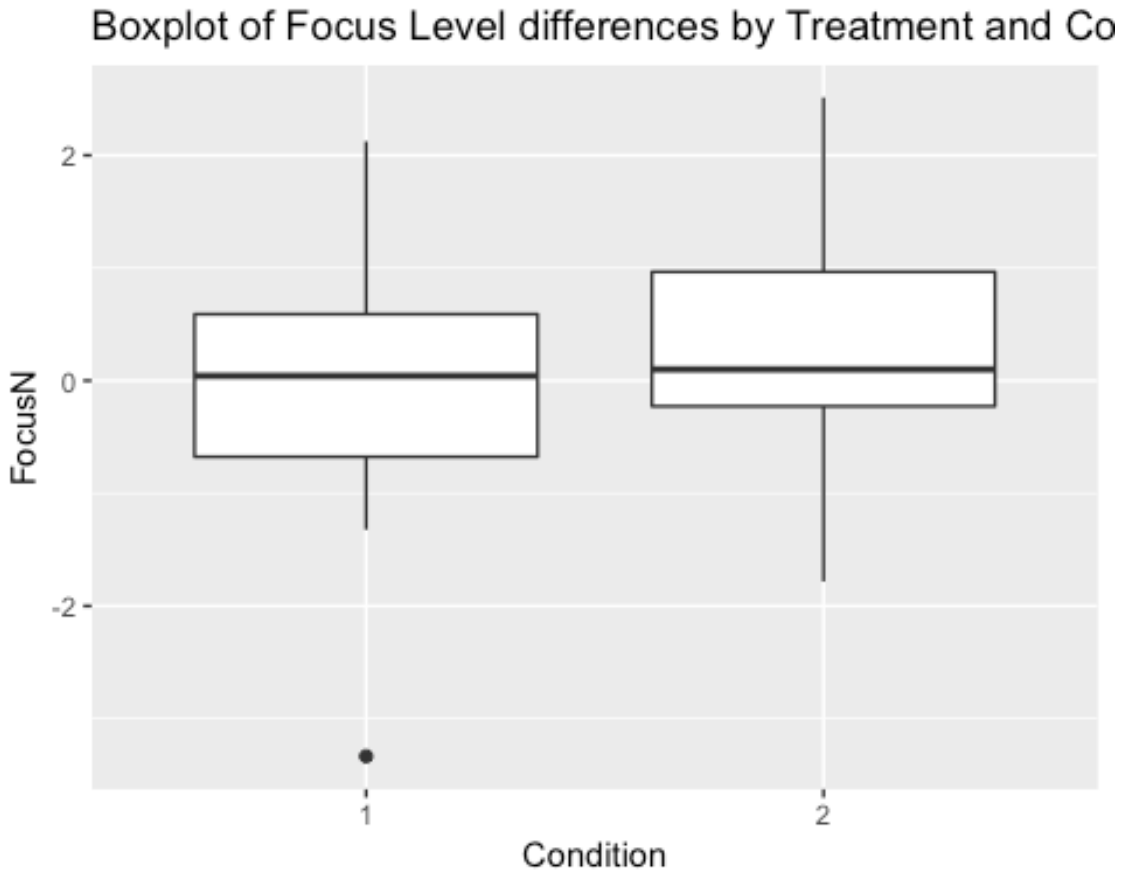
The chart shows the frequency for missing data of focus level that sorted into day and prompt time. Note that a lighter color refers to a lower frequency.

```
summary(summ_dat_focus_diff)
```

```
##         ID     Condition GENDER OVERWEIGHT      AGE             BMI
##  1      : 1   1:18      0:36   0:29       Min.   :22.00   Min.   :18.
37
##  2      : 1   2:21      1: 3   1:10       1st Qu.:32.00   1st Qu.:20.
34
##  3      : 1             2: 0              Median :39.00   Median :22.
21
##  4      : 1                               Mean   :38.97   Mean   :23.
35
##  5      : 1                               3rd Qu.:44.50   3rd Qu.:24.
82
##  6      : 1                               Max.   :59.00   Max.   :35.
38
##  (Other):33
##
##      FocusN           Work_Sit         Work_Stand         Work_Step
```

```
##  Min.   :-3.3333   Min.   :-5.4721   Min.   :-1.31538   Min.   :-0.6
8752
##  1st Qu.:-0.3285   1st Qu.:-1.0121   1st Qu.:-0.08767   1st Qu.:-0.1
4280
##  Median : 0.0850   Median :-0.5024   Median : 0.14778   Median : 0.0
3513
##  Mean   : 0.1256   Mean   :-0.4996   Mean   : 0.45713   Mean   :-0.0
1565
##  3rd Qu.: 0.9234   3rd Qu.: 0.2016   3rd Qu.: 0.72089   3rd Qu.: 0.0
9544
##  Max.   : 2.5145   Max.   : 1.3550   Max.   : 5.48058   Max.   : 0.5
5830
##
```

```
ggplot(summ_dat_diff, aes(x = Condition, y = FocusN)) +
  geom_boxplot() +
  ggtitle("Boxplot of Focus Level differences by Treatment and Control
Group")
```

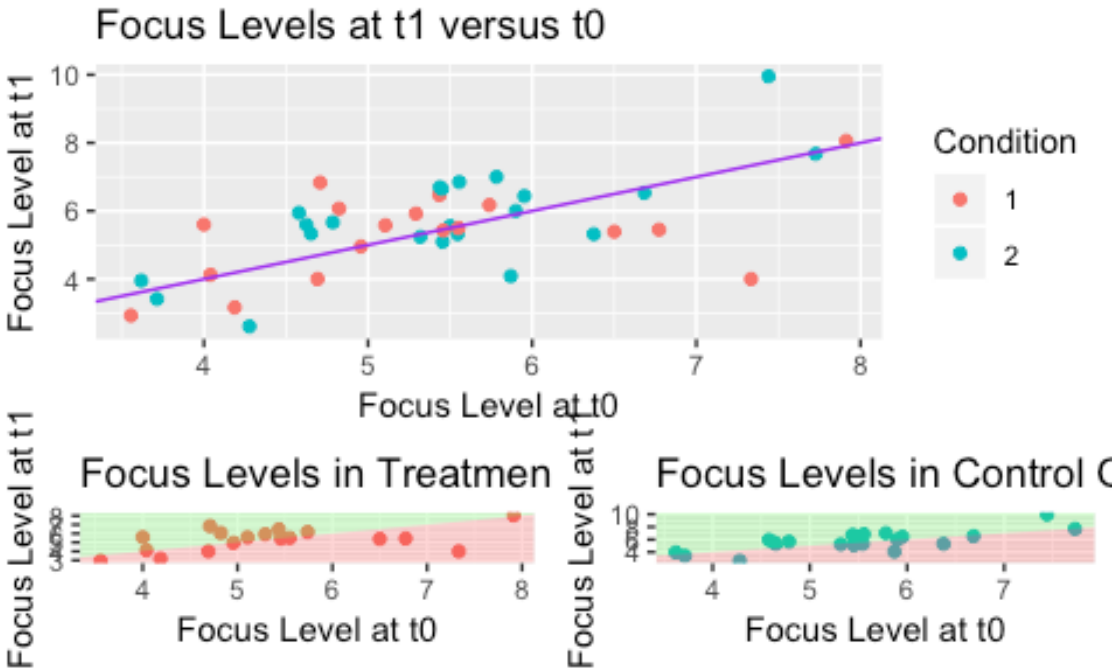


The chart summarizes the three-month changes in focus level and activity level of participants in either treatment or control group. As stated in the result, their average sitting and stepping time is reduced, the average standing time is raised,

and the mean value of focus level is increased by 0.1256 after three months. The boxplot depicts the focus level differences distribution by treatment group and control group, where condition 1 refers to the treatment group and condition 2 refers to the control group.

```
summary(summ_dat_focus_cont)

##        ID        FocusN0          FocusN1
##  1      : 1   Min.   :3.556   Min.   :2.929
##  3      : 1   1st Qu.:4.696   1st Qu.:4.333
##  4      : 1   Median :5.198   Median :5.475
##  5      : 1   Mean   :5.337   Mean   :5.313
##  7      : 1   3rd Qu.:5.692   3rd Qu.:6.030
##  20     : 1   Max.   :7.913   Max.   :8.056
##  (Other):12

summary(summ_dat_focus_diff_treat)

##        ID     Condition GENDER OVERWEIGHT      AGE             BMI

##  1      : 1   1:18      0:18   0:13       Min.   :23.00   Min.   :19.
07
##  3      : 1   2: 0      1: 0   1: 5       1st Qu.:34.00   1st Qu.:20.
26
##  4      : 1             2: 0              Median :39.50   Median :21.
82
##  5      : 1                                Mean   :39.67   Mean   :23.
29
##  7      : 1                                3rd Qu.:43.00   3rd Qu.:25.
19
##  20     : 1                                Max.   :59.00   Max.   :35.
38
##  (Other):12

##      FocusN            Work_Sit         Work_Stand        Work_Step

##  Min.   :-3.33333   Min.   :-5.4721   Min.   :-0.3604   Min.   :-0.5
9980
##  1st Qu.:-0.67598   1st Qu.:-1.3344   1st Qu.: 0.1005   1st Qu.:-0.1
8881
##  Median : 0.04250   Median :-0.9261   Median : 0.6135   Median : 0.0
3231
##  Mean   :-0.02335   Mean   :-1.1382   Mean   : 1.0488   Mean   :-0.0
1344
##  3rd Qu.: 0.58967   3rd Qu.:-0.2881   3rd Qu.: 1.2733   3rd Qu.: 0.1
1705
##  Max.   : 2.12500   Max.   : 0.8432   Max.   : 5.4806   Max.   : 0.5
5830
##
```

The first summary illustrates the responses of focus level in the treatment group at baseline (FocusN0) and the third month (FocusN1). The changes in focus state along with the differences in sitting, standing, and stepping time are indicated in the second summary. From the result, we find that in the treatment group, there is a reduction in average sitting time and an improvement of average standing time, while the mean value of focus level is decreasing.



The first scatterplot shows the focusability of participants in both groups at t1 and t0, where the treatment group and the control group are distinguished by different colors. The following two scatterplots demonstrate the treatment or control group's focus level performance respectively at t1 versus t0. Note that point in the green area implies that the participant has a positive emotional change from t0 to t1, while the point in the red area refers to a negative emotional change.

Data Analysis

```
full <- lm(FocusN ~ Work_Sit + Work_Stand + Work_Step +
             Condition + GENDER + AGE + BMI + OVERWEIGHT,
           data = summ_dat_diff)
summary(full)

##
## Call:
## lm(formula = FocusN ~ Work_Sit + Work_Stand + Work_Step + Condition
```

```
+
##     GENDER + AGE + BMI + OVERWEIGHT, data = summ_dat_diff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.3832 -0.7927  0.0944  0.7251  1.8931
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.19684    2.25378   0.531    0.599
## Work_Sit     0.14890    0.34230   0.435    0.667
## Work_Stand   0.42966    0.36297   1.184    0.246
## Work_Step    1.15612    0.87260   1.325    0.195
## Condition2   0.65095    0.42985   1.514    0.140
## GENDER1     -0.67088    0.73943  -0.907    0.371
## AGE         -0.02519    0.02519  -1.000    0.325
## BMI         -0.02322    0.10625  -0.219    0.829
## OVERWEIGHT1  0.19404    0.84828   0.229    0.821
##
## Residual standard error: 1.153 on 30 degrees of freedom
## Multiple R-squared:  0.1748, Adjusted R-squared:  -0.0453
## F-statistic: 0.7942 on 8 and 30 DF,  p-value: 0.612

step(full, direction = "backward")

## Start:  AIC=18.88
## FocusN ~ Work_Sit + Work_Stand + Work_Step + Condition + GENDER +
##     AGE + BMI + OVERWEIGHT
##
##              Df Sum of Sq    RSS    AIC
## - BMI         1   0.06348 39.950 16.938
## - OVERWEIGHT  1   0.06957 39.956 16.944
## - Work_Sit    1   0.25157 40.138 17.121
## - GENDER      1   1.09445 40.980 17.932
## - AGE         1   1.32897 41.215 18.154
## - Work_Stand  1   1.86296 41.749 18.656
## <none>                    39.886 18.876
## - Work_Step   1   2.33385 42.220 19.094
## - Condition   1   3.04906 42.935 19.749
##
## Step:  AIC=16.94
## FocusN ~ Work_Sit + Work_Stand + Work_Step + Condition + GENDER +
##     AGE + OVERWEIGHT
##
##              Df Sum of Sq    RSS    AIC
## - OVERWEIGHT  1   0.01098 39.960 14.949
## - Work_Sit    1   0.20320 40.153 15.136
## - GENDER      1   1.08857 41.038 15.987
## - AGE         1   1.62898 41.578 16.497
## - Work_Stand  1   1.80403 41.754 16.661
```

```
## <none>                     39.950 16.938
## - Work_Step   1   2.36661 42.316 17.183
## - Condition   1   3.05665 43.006 17.814
##
## Step:  AIC=14.95
## FocusN ~ Work_Sit + Work_Stand + Work_Step + Condition + GENDER +
##     AGE
##
##              Df Sum of Sq    RSS    AIC
## - Work_Sit    1   0.23731 40.198 13.180
## - GENDER      1   1.09703 41.058 14.005
## - Work_Stand  1   1.92468 41.885 14.783
## - AGE         1   2.06132 42.022 14.911
## <none>                    39.960 14.949
## - Work_Step   1   2.46575 42.426 15.284
## - Condition   1   3.04770 43.008 15.815
##
## Step:  AIC=13.18
## FocusN ~ Work_Stand + Work_Step + Condition + GENDER + AGE
##
##              Df Sum of Sq    RSS    AIC
## - GENDER      1    0.9753 41.173 12.115
## - AGE         1    1.9550 42.153 13.032
## <none>                    40.198 13.180
## - Work_Step   1    2.2604 42.458 13.313
## - Condition   1    3.1814 43.379 14.150
## - Work_Stand  1    3.7868 43.985 14.691
##
## Step:  AIC=12.11
## FocusN ~ Work_Stand + Work_Step + Condition + AGE
##
##              Df Sum of Sq    RSS    AIC
## - Work_Step   1    1.9088 43.082 11.882
## <none>                    41.173 12.115
## - AGE         1    2.2843 43.457 12.221
## - Condition   1    2.5464 43.720 12.455
## - Work_Stand  1    3.9249 45.098 13.666
##
## Step:  AIC=11.88
## FocusN ~ Work_Stand + Condition + AGE
##
##              Df Sum of Sq    RSS    AIC
## - AGE         1    2.0141 45.096 11.664
## - Condition   1    2.2121 45.294 11.835
## <none>                    43.082 11.882
## - Work_Stand  1    3.1374 46.219 12.624
##
## Step:  AIC=11.66
## FocusN ~ Work_Stand + Condition
##
```

```
##               Df Sum of Sq    RSS    AIC
## - Condition    1    2.1945 47.291 11.517
## <none>                     45.096 11.664
## - Work_Stand   1    2.4955 47.592 11.765
##
## Step:  AIC=11.52
## FocusN ~ Work_Stand
##
##               Df Sum of Sq    RSS    AIC
## - Work_Stand   1    1.0425 48.333 10.368
## <none>                     47.291 11.517
##
## Step:  AIC=10.37
## FocusN ~ 1

##
## Call:
## lm(formula = FocusN ~ 1, data = summ_dat_diff)
##
## Coefficients:
## (Intercept)
##      0.1256
```

A global linear regression model was created with focus level as the response variable and all of the factors that might have effects on the emotional state, including demographic effects and differences in sitting time, standing time, and stepping time while working in an office, as explanatory variables.

The output shows that the p-values of the test are all greater than a 5% significance level, indicating that the explanatory variables do not affect focus level.

AIC in a stepwise algorithm was applied to determine the best-fitted model that minimized the information loss for our analysis. As a lower AIC value implied a better model, here we chose the model with the lowest AIC value.

According to the result, Work_Stand that represents the changes in standing time is the only explanatory variable selected to form the model.

```
model_focus_pick <- lm(FocusN ~ Work_Stand, data = summ_dat_diff)
model_focus_pick

##
## Call:
## lm(formula = FocusN ~ Work_Stand, data = summ_dat_diff)
##
## Coefficients:
## (Intercept)   Work_Stand
##      0.0640       0.1347

summary(model_focus_pick)
```

```
##
## Call:
## lm(formula = FocusN ~ Work_Stand, data = summ_dat_diff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4083 -0.4486 -0.0754  0.8220  2.4190
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0640     0.1935   0.331    0.743
## Work_Stand    0.1347     0.1492   0.903    0.372
##
## Residual standard error: 1.131 on 37 degrees of freedom
## Multiple R-squared:  0.02157,    Adjusted R-squared:  -0.004876
## F-statistic: 0.8156 on 1 and 37 DF,  p-value: 0.3723
```

From the linear regression model presented above, we have the equation FocusN = 0.064 + 0.1347 * Work_Stand, which implies that the changes in patterns of focus level are positively correlated with improvement in standing time. However, the effect of Work_Stand variable is found to be not significant at a 5% significance level, and the null hypothesis that the average differences in focus level are the same before and after applying changes in standing time is not rejected.

Conclusion and Further Discussions

To analyze if reducing sitting time in a three-month scale can cause a significant change in focusability, multiple linear regression models are conducted. The Akaike information criterion (AIC) is applied to the global model that constructed with the response variable, focus level, as well as all the possible significant explanatory variables, including activity levels and demographic effects, to obtain the model with the highest relative quality. It is found that the model which only takes changes in standing time as the explanatory variable is the most preferred one due to its minimum AIC value. Nevertheless, a large p-value, which is greater than the 5% significance level, indicates that modification in standing time does not have a significant effect on the adjustment of focus state. In other words, we conclude that a reduction in sitting time does not result in changes in responses to focus level.

To acquire a more accurate and persuasive conclusion and a further understanding of the subject, a larger data set with sufficient participants and completed record of measurements is needed. In addition, the missing data is suggested to be filled with a better imputation method instead of removing all of it.

Reference

David, W.D., Bethany, H., Genevieve, N.H., Neville,O.(2012). Too much sitting – A health hazard