Final Project

Morgan Godley, Priom Mahmud, Qin Xu, Chaobo Yan

Method and Analysis

Model Assumption

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_k + \epsilon_{ijk}$$

where μ is the overall mean, τ_i is the ith level of activity (walking and meditation), β_j is the jth level of time (morning and afternoon) and γ_k is the kth team member. $\epsilon_{ijk} \sim N\left(0, \sigma^2_{ijk}\right)$ is the random error.

Team member as a blocking factor can be controlled as random effect because our four team members were sampled from all the potential 571B students who are interested in the same topic and it happens to be us. Nevertheless, team member can also be fixed effect due to the reason that only four of the current 571B students interested in our public health topic. We will compare both of scenarios in our model and the difference between those two scenarios.

In the case of team member as a random effect, we assume $k \sim N\left(\gamma_k, \sigma^2_{\gamma_k}\right)$. Therefore, $\epsilon_{ijk} \sim N\left(0, \sigma^2_{ijk} + \sigma^2_{\gamma_k}\right)$ is this case, where the model could be more volatile.

Experiment Design

Due to the availability of our experiment are easy to be executed and the cost is free, we show a 2^2 full factorial design to each team member. Here is the design table:

```
##      Team member Activity Time
## 1  Team member 1        -    -
## 2  Team member 2        -    -
## 3  Team member 3        -    -
## 4  Team member 4        -    -
## 5  Team member 1        +    -
## 6  Team member 2        +    -
## 7  Team member 3        +    -
## 8  Team member 4        +    -
## 9  Team member 1        -    +
## 10 Team member 2        -    +
## 11 Team member 3        -    +
## 12 Team member 4        -    +
## 13 Team member 1        +    +
## 14 Team member 2        +    +
## 15 Team member 3        +    +
## 16 Team member 4        +    +
```

We mark walking as +, meditation as - in Activity and morning as +, afternoon as - in Time.

Test Hypotheses

The above design and statistical model can lead to the following test hypothesis

H_0 : The mean change of the mood before/after taking the activity is the same

H_a : The mean change of the mood before/after taking the activity is not the same

First of all, we want to include activity τ_i and time β_j two main effects to reveal whether they do have significant effect on the mood change. We also want to include the two-way interaction $(\tau\beta)_{ij}$ between activity and time to see if there is any correlation between them.

If our assumption about effect of activity (meditation or walking) truly depends on the time of day, we should see terms have significant impact on our model.

Even though the interactions among blocking factor team member γ_k and two treatment factors are not assumed in our statistical model, we also discuss them in our result section to show some interesting facts across personal perspective of mood change.

However, we might find there is not too much evidence to show our assumption from above is correct. In that case, we can start dropping the terms who have less evidence with hierarchical orders.

Except for testing the hypothesis, we also investigate which activity has a higher average effect on the mood change, how the time of day affect the mood change and its interaction with the activity, the impact of personal favorite and interactions etc.

Test Power

To balance the executability of the experiment (sample size/replicate) and the power of the test, we can pre-set the difference of mean change and the power of the test to a certain level and calculate the number of samples we need to reach such a level. Of course, more samples means higher power but we might have to reduce the optimal sample size for sake of availability of each team member. Since we haven't started the experiment yet, we can assume we have some pilot data from any other four UA graduate students by randomly sampling the uniform random samples from -10 to 10 to get some approximating and general ideas of the power of the test. Here is the table of power of the test for some reasonable difference value and sample size are set at the significance level 0.05:

Power of the test	Difference in mean change	Sample size
0.95	0.1/1/3/5	74542/747/84/31
0.9	0.1/1/3/5	60275/604/68/26
0.85	0.1/1/3/5	51504/516/59/22
0.8	0.1/1/3/5	45024/452/52/20

We do not consider any design with sample size over 100 because of the availability and time constraints. There are considerable sample sizes, for example, to test the difference in mean change within 3 with power 0.8, we only need 64 samples for each treatment in total meaning 4 replicates for each team member in the morning or afternoon per treatment roughly.

Random effect of not studying the interaction of blocking factor

```
library(lmerTest)
```

```
## Loading required package: lme4
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'lmerTest'
```

```
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
## The following object is masked from 'package:stats':
##
##     step
```

```
wm_lmer <- lmer(Score ~ Time * Activity + (1|User),
                data = wm,
                contrasts = list(Time = "contr.helmert",
                Activity = "contr.helmert"))
```

```
summary(wm_lmer)
```

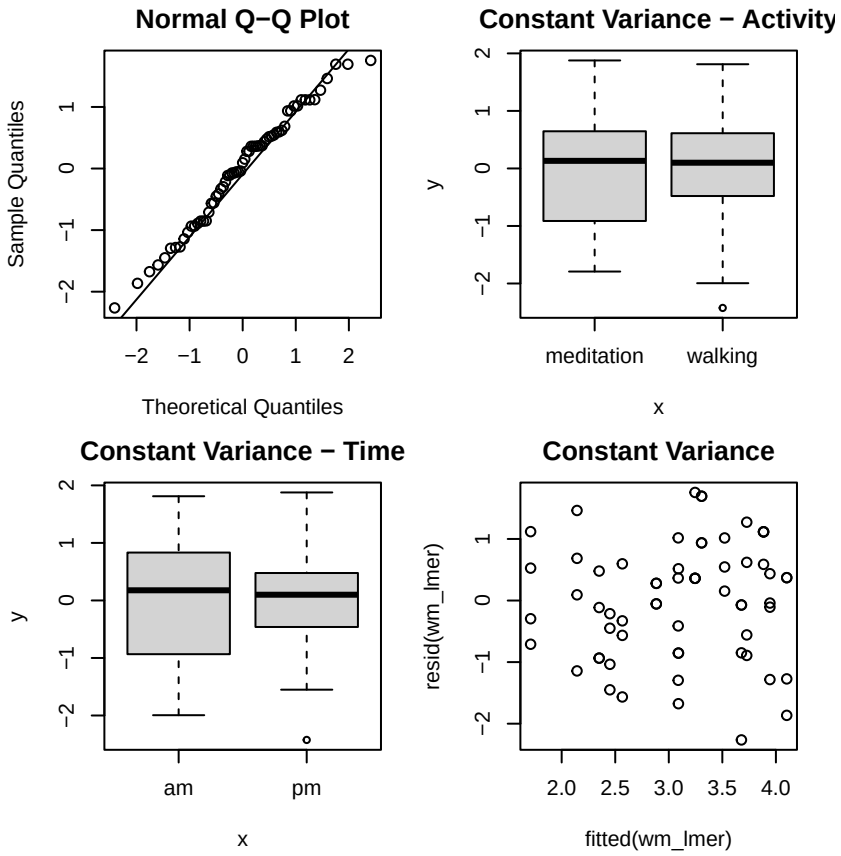
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Score ~ Time * Activity + (1 | User)
##    Data: wm
##
## REML criterion at convergence: 189.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.29089 -0.78917  0.09393  0.59871  1.77593
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  User     (Intercept) 0.5446   0.7379
##  Residual             0.9766   0.9882
## Number of obs: 63, groups:  User, 4
##
## Fixed effects:
##                 Estimate Std. Error      df t value Pr(>|t|)
## (Intercept)       3.1056     0.3894  3.0042   7.975  0.00409 **
## Time1            -0.2657     0.1246 56.0144  -2.133  0.03736 *
## Activity1         0.1624     0.1246 56.0144   1.303  0.19781
## Time1:Activity1   0.0542     0.1246 56.0144   0.435  0.66523
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Time1 Actvt1
## Time1       0.005
## Activity1   0.005  0.017
## Tm1:Actvty1 0.005  0.017 0.017
```

```
ranova(wm_lmer)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Score ~ Time + Activity + (1 | User) + Time:Activity
##            npar   logLik    AIC    LRT Df Pr(>Chisq)
## <none>        6  -94.724 201.45
## (1 | User)    5 -102.242 214.48 15.037  1  0.0001054 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
layout(matrix(1:6, 2, 3)); par(mar = c(4, 4, 2.5, 1.5))
# Check normality
qqnorm(residuals(wm_lmer))
qqline(residuals(wm_lmer))

#Residuals
plot(wm$Time, rstudent(wm_lmer), main = "Constant Variance - Time")
plot(wm$Activity, rstudent(wm_lmer), main = "Constant Variance - Activity")
plot(fitted(wm_lmer), resid(wm_lmer), main = "Constant Variance")
#plot(wm_lmer,which=4)
```



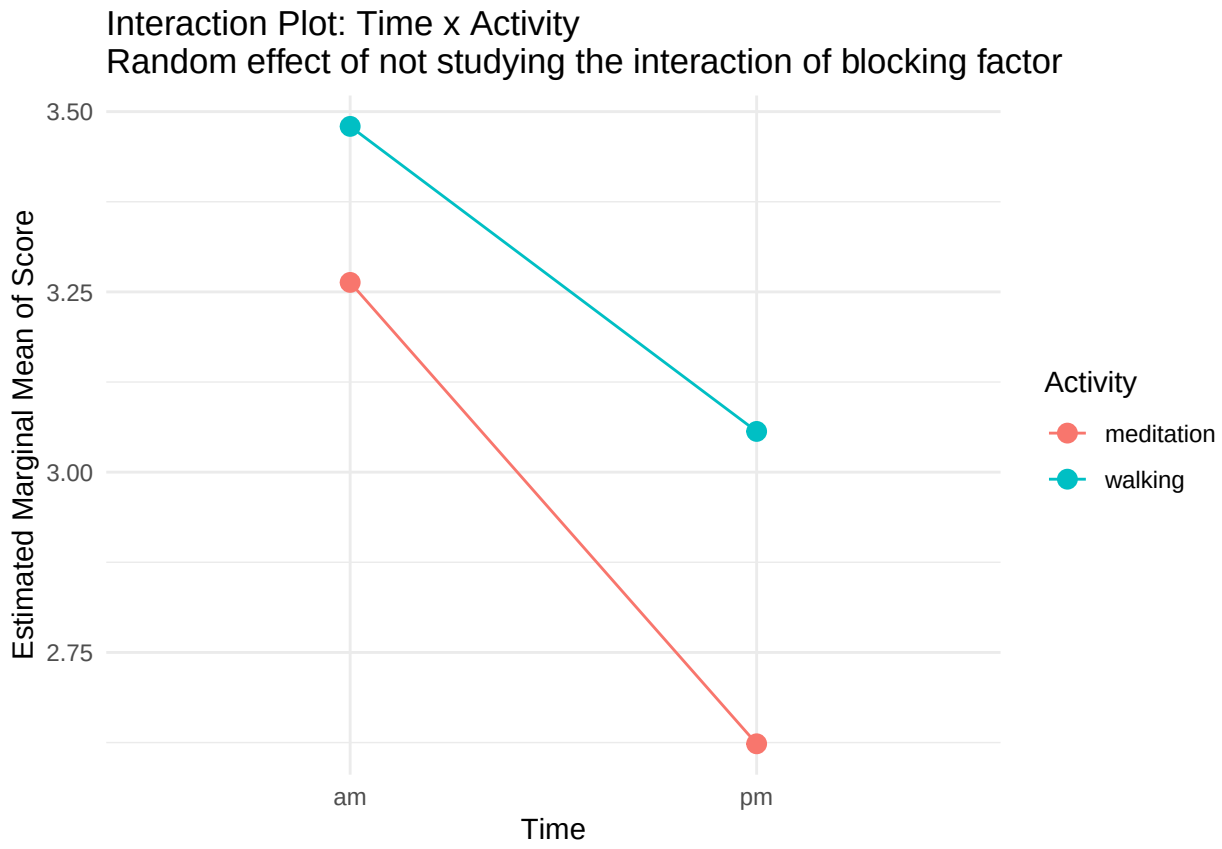

```
# Interaction Plot

# Estimate Marginal Means
library(emmeans)
```

```
## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'
```

```
emm_wm_lmer <- emmeans(wm_lmer, ~ Time * Activity)
```

```
## Cannot use mode = "kenward-roger" because *pbkrtest* package is not installed
```

```r
emm_df_wm_lmer <- as.data.frame(emm_wm_lmer) #as DF

# Use ggplot2 for the interaction plot
library(ggplot2)
ggplot(emm_df_wm_lmer, aes(x = Time, y = emmean, group = Activity, color = Activity)) +
  geom_point(size = 3) +
  geom_line() +
  labs(title = "Interaction Plot: Time x Activity\nRandom effect of not studying the interaction of bloc
       y = "Estimated Marginal Mean of Score") +
  theme_minimal()
```



Fixed effect of not studying the interaction of blocking factor

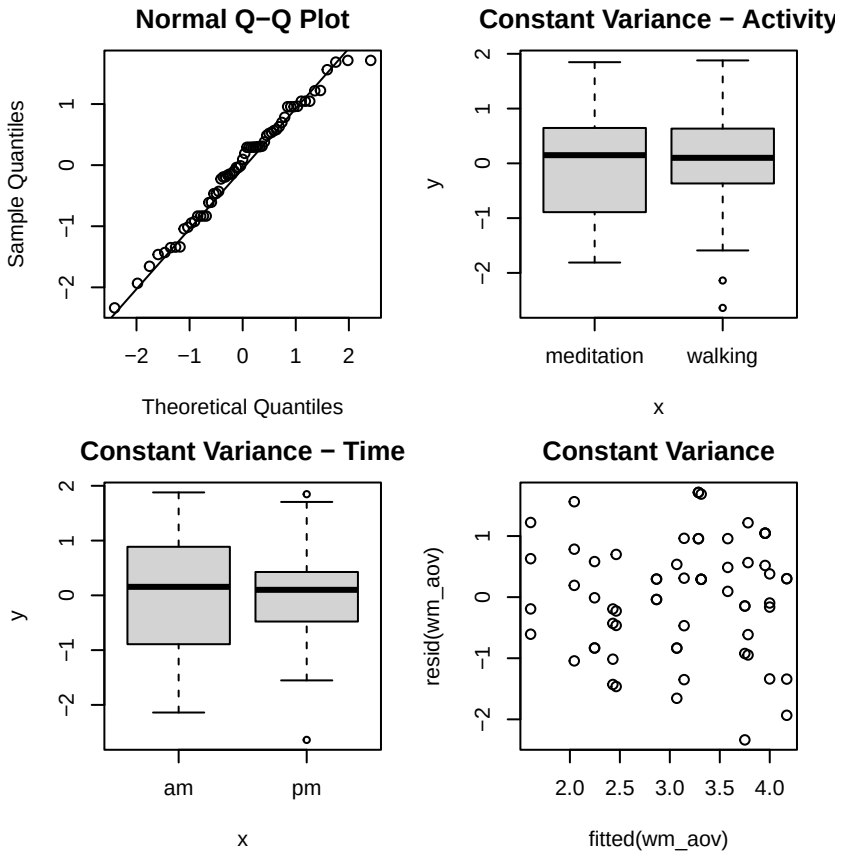
```r
wm_aov <- aov(Score ~ Time * Activity + User,
              data = wm,
              contrasts = list(Time = "contr.helmert",
              Activity = "contr.helmert"))
```

```r
summary(wm_aov)
```

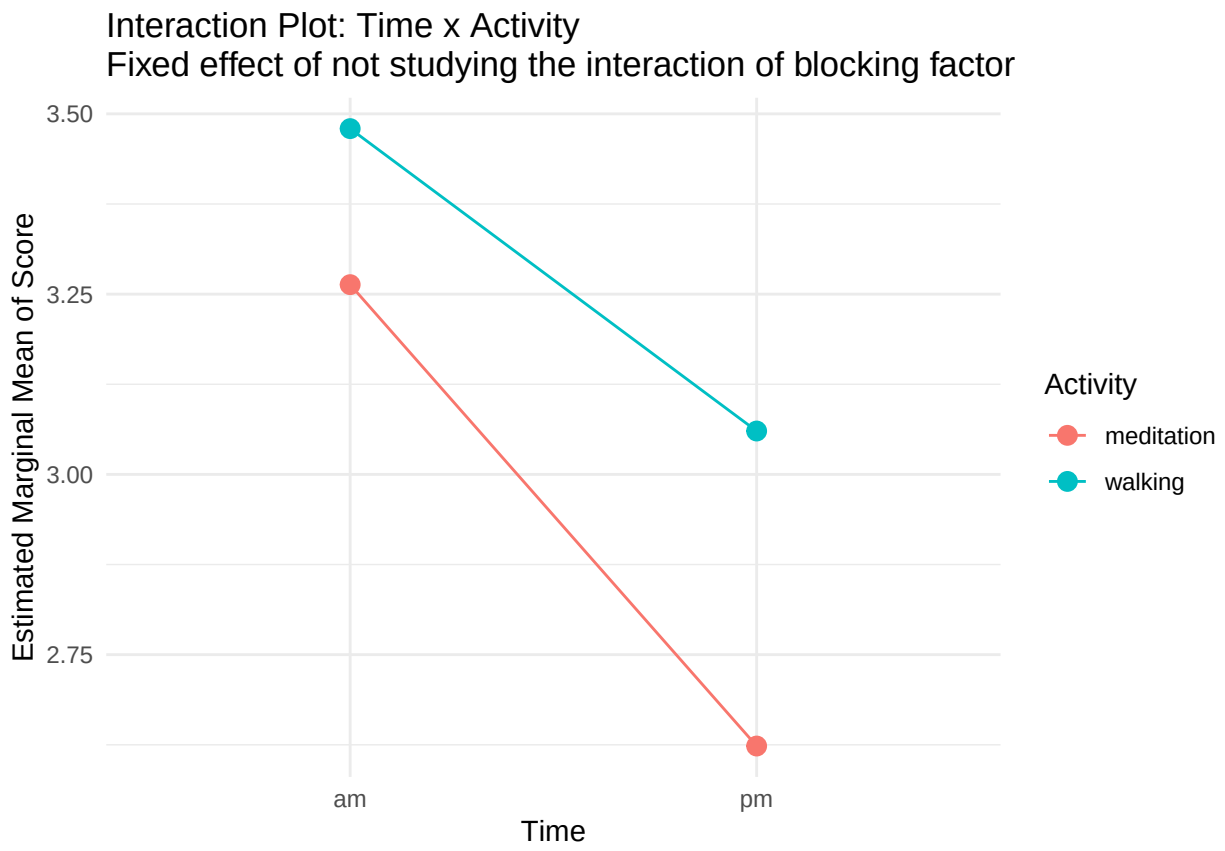
```
##               Df Sum Sq Mean Sq F value   Pr(>F)
## Time           1   4.82   4.822   4.938   0.0303 *
## Activity       1   1.49   1.491   1.527   0.2218
## User           3  28.73   9.578   9.807 2.69e-05 ***
## Time:Activity  1   0.19   0.191   0.196   0.6600
## Residuals     56  54.69   0.977
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
layout(matrix(1:6, 2, 3)); par(mar = c(4, 4, 2.5, 1.5))
# Check normality
qqnorm(residuals(wm_aov))
qqline(residuals(wm_aov))

# Check constant variance
plot(wm$Time, rstudent(wm_aov), main = "Constant Variance - Time")
plot(wm$Activity, rstudent(wm_aov), main = "Constant Variance - Activity")
plot(fitted(wm_aov), resid(wm_aov), main = "Constant Variance")
```




```r
# Interaction Plot

# Estimate Marginal Means
emm_wm_aov <- emmeans(wm_aov, ~ Time * Activity)
emm_df_wm_aov <- as.data.frame(emm_wm_aov) #as DF

# Interaction plot
ggplot(emm_df_wm_aov, aes(x = Time, y = emmean, group = Activity, color = Activity)) +
  geom_point(size = 3) +
  geom_line() +
  labs(title = "Interaction Plot: Time x Activity\nFixed effect of not studying the interaction of block
       y = "Estimated Marginal Mean of Score") +
  theme_minimal()
```



Fixed effect of study the interaction of blocking factor

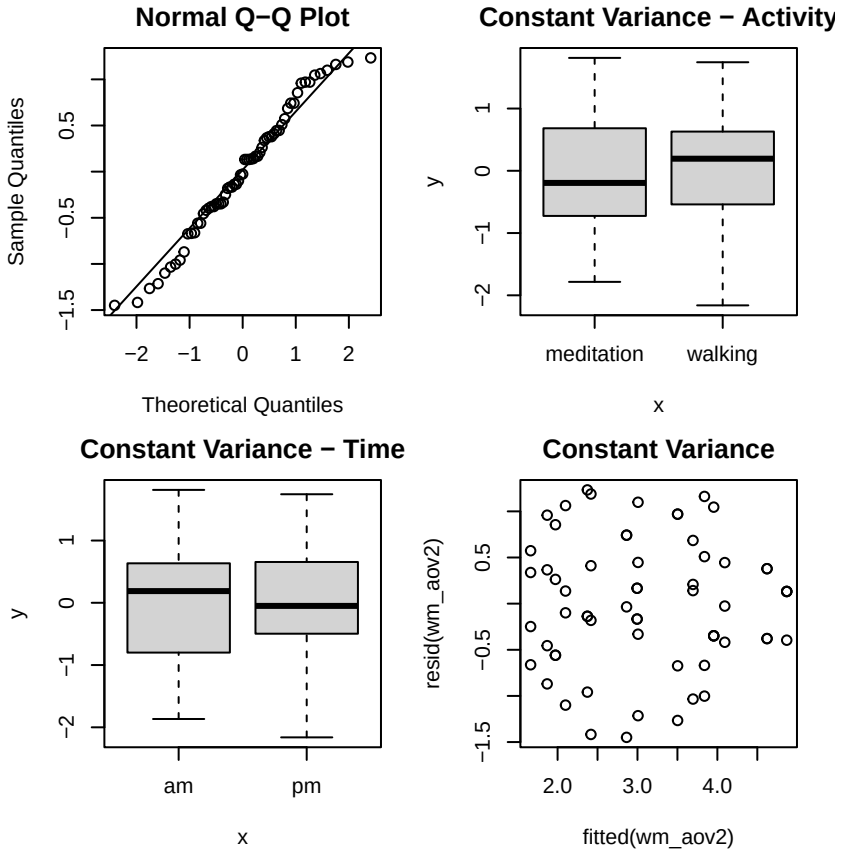
```
wm_aov2 <- aov(Score ~ Time * Activity * User,
               data = wm,
               contrasts = list(Time = "contr.helmert",
               Activity = "contr.helmert"))
```

```
anova(wm_aov2)
```

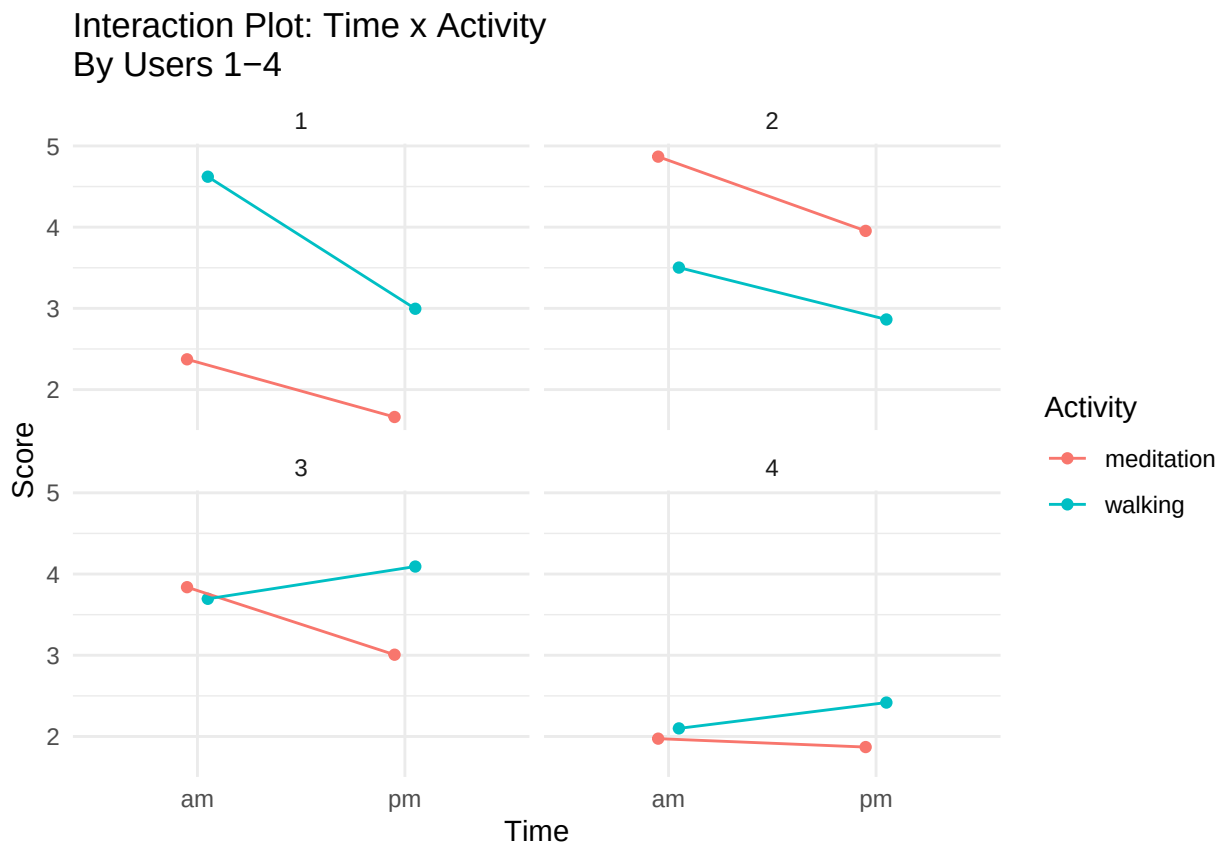
```
## Analysis of Variance Table
##
## Response: Score
##                     Df  Sum Sq Mean Sq F value    Pr(>F)
## Time                 1  4.8225  4.8225  7.4626  0.008845 **
## Activity             1  1.4911  1.4911  2.3074  0.135454
## User                 3 28.7346  9.5782 14.8218 6.340e-07 ***
## Time:Activity        1  0.1911  0.1911  0.2957  0.589179
## Time:User            3  3.7619  1.2540  1.9404  0.135950
## Activity:User        3 18.2854  6.0951  9.4319 5.452e-05 ***
## Time:Activity:User   3  2.2722  0.7574  1.1721  0.330432
## Residuals           47 30.3726  0.6462
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
layout(matrix(1:6, 2, 3)); par(mar = c(4, 4, 2.5, 1.5))
# Check normality
```

```r
qqnorm(residuals(wm_aov2))
qqline(residuals(wm_aov2))

# Check constant variance
plot(wm$Time, rstudent(wm_aov2), main = "Constant Variance - Time")
plot(wm$Activity, rstudent(wm_aov2), main = "Constant Variance - Activity")
plot(fitted(wm_aov2), resid(wm_aov2), main = "Constant Variance")
```




```r
ggplot(wm, aes(x = Time, y = Score, color = Activity, group = Activity)) +
  stat_summary(fun = mean, geom = "point", position = position_dodge(0.2)) +
  stat_summary(fun = mean, geom = "line", position = position_dodge(0.2)) +
  facet_wrap(~ User) +
  labs(title = "Interaction Plot: Time x Activity\nBy Users 1-4") +
  theme_minimal()
```



Random effect of studying the interaction of blocking factor

```
wm_lmer2 <- lmer(Score ~ Time * Activity
                 + (Time + Activity|User),
                data = wm,
                contrasts = list(Time = "contr.helmert",
                Activity = "contr.helmert"))
```

```
summary(wm_lmer2)
```

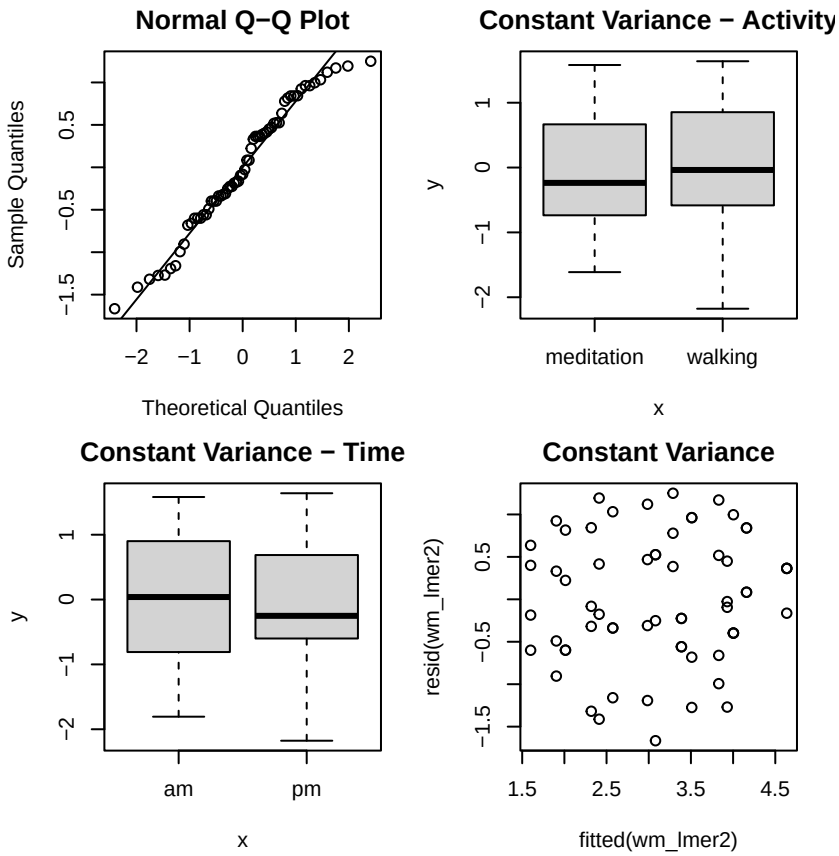
```
## Linear mixed model fit by REML. t-tests use Satterthwaite's method [
## lmerModLmerTest]
## Formula: Score ~ Time * Activity + (Time + Activity | User)
##    Data: wm
##
## REML criterion at convergence: 173
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.0675 -0.6511 -0.1026  0.6520  1.5508
##
## Random effects:
##  Groups   Name            Variance Std.Dev. Corr
##  User     (Intercept)     1.4954   1.2229
##           Timepm          0.1678   0.4097   -0.38
##           Activitywalking 1.3629   1.1674   -0.73 -0.35
```

```
##  Residual                 0.6493   0.8058
## Number of obs: 63, groups:  User, 4
##
## Fixed effects:
##                 Estimate Std. Error       df t value Pr(>|t|)
## (Intercept)      3.10195    0.38768  3.00532   8.001  0.00405 **
## Time1           -0.26937    0.14426  3.64206  -1.867  0.14224
## Activity1        0.15874    0.30903  3.01353   0.514  0.64277
## Time1:Activity1  0.05055    0.10159 53.01640   0.498  0.62080
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) Time1  Actvt1
## Time1       -0.425
## Activity1   -0.464 -0.233
## Tm1:Actvty1  0.004  0.012  0.006
```

```
ranova(wm_lmer2)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## Score ~ Time + Activity + (Time + Activity | User) + Time:Activity
##                                      npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>                                 11 -86.501 195.00
## Time in (Time + Activity | User)        8 -87.778 191.56  2.5531  3   0.465778
## Activity in (Time + Activity | User)    8 -94.355 204.71 15.7068  3   0.001302
##
## <none>
## Time in (Time + Activity | User)
## Activity in (Time + Activity | User) **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

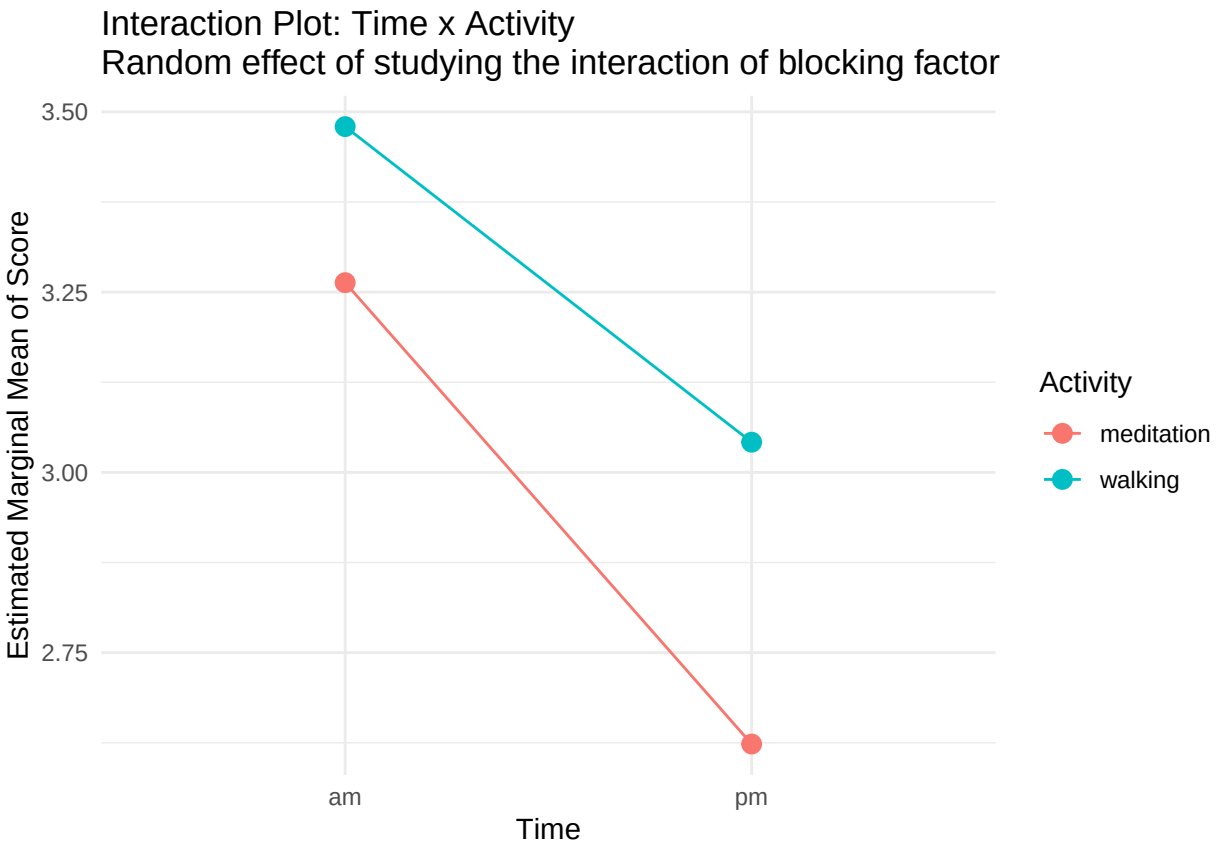
```
layout(matrix(1:6, 2, 3)); par(mar = c(4, 4, 2.5, 1.5))
# Check normality
qqnorm(residuals(wm_lmer2))
qqline(residuals(wm_lmer2))

# Check constant variance
plot(wm$Time, rstudent(wm_lmer2), main = "Constant Variance - Time")
plot(wm$Activity, rstudent(wm_lmer2), main = "Constant Variance - Activity")
plot(fitted(wm_lmer2), resid(wm_lmer2), main = "Constant Variance")
```




```
# Interaction Plot

# Estimate Marginal Means
emm_wm_lmer2 <- emmeans(wm_lmer2, ~ Time * Activity)
```

```
## Cannot use mode = "kenward-roger" because *pbkrtest* package is not installed
```

```
emm_df_wm_lmer2 <- as.data.frame(emm_wm_lmer2) # as DF

# Interaction plot
ggplot(emm_df_wm_lmer2, aes(x = Time, y = emmean, group = Activity, color = Activity)) +
  geom_point(size = 3) +
  geom_line() +
  labs(title = "Interaction Plot: Time x Activity\nRandom effect of studying the interaction of blockin
       y = "Estimated Marginal Mean of Score") +
  theme_minimal()
```


Interaction Plot: Time x Activity
Random effect of studying the interaction of blocking factor
Estimated Marginal Mean of Score
3.50
3.25
3.00
2.75
am
pm
Time
Activity
meditation
walking