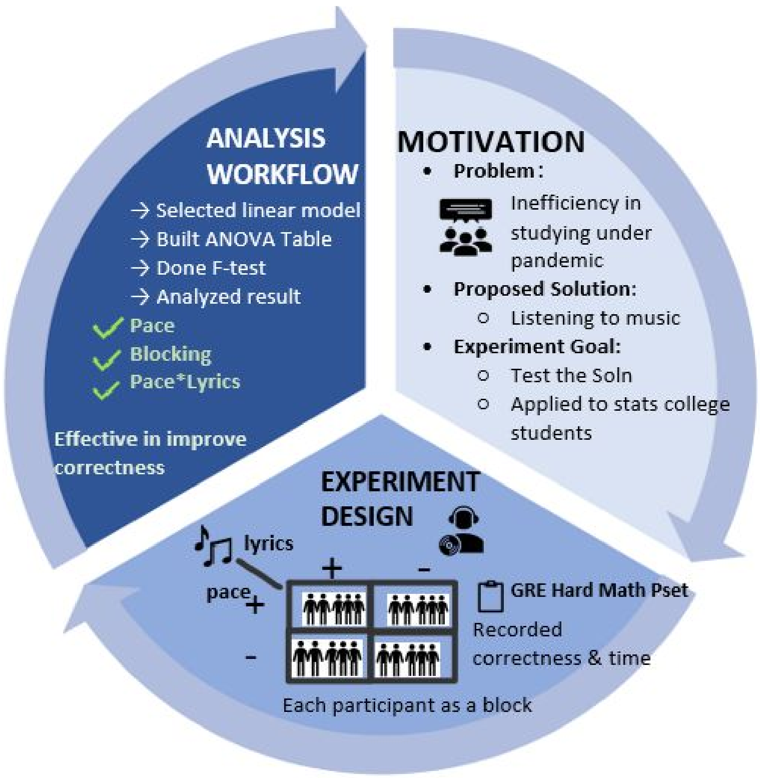
Stat 424 Group Project Milestone 4
Group 7 Ruofan Yao, Zimo Wan, Yuqi Niu, Jing Duan, Yanxu Guo

I. Introduction

During the global pandemic, teachers and students have to change the teaching/learning mode to remote one. While studying at home allows students to have more flexible schedules, it also brings challenges to them, such as how to stay focused and improve the efficiency of finishing schoolwork or preparing for standardized tests. It could be difficult for college students to seek guidance from tutors or advisors without face-to-face communication. Therefore, our group decided to find a way to ensure studying quality that could benefit students like us who are concerned about their deficiency in concentrating and adjusting their study mode. Since music is accessible to most students, and it always brings fondness to oneself, which may potentially cause positive changes in students' studying outcomes, our team decided to study whether listening to music would influence college students' correctness and efficiency in doing homework. Since all of our participants (group members) are majoring in statistics, we specified our focus on the impacts on stats college students in doing mathematics problems to allow the test results to be generalized. Which specific types of music that impacts students' performance are also something we were eager to explore. Our hypothesis was that the fast pace music could increase the time of completing math problems while decreasing the correctness of that.

Previous research had hypothesized and tested the relationship between external stimulus (i.e. music) and cognitive performance. For instance, Lehmann and Seufert (2017) had conducted a study including 81 college students with half of the students listening to pop music and half of students studying in silence indicated that the interaction of working memory capacity and music would have a positive relationship with students' performance, while there is no relationship solely between students' performance and the music they heard (p. 8). However, this study didn't take individual differences into account (i.e. not blocking) and it also did not specify what kinds of music could positively influence students' performance, which we would test in our study later. Therefore, our experiment may provide useful information to give a better understanding of the impacts of music on study.



II. Design and Methodology

Since we had two factors with only two levels for each factor: fast-pace (+) or slow-pace (-) and lyrics (+) or no lyrics (-), we decided to conduct a blocked 2^2 factorial design. To eliminate the individual effect, we chose to run the experiment in five blocks and compared the effect within the block. An alternative design was to give questions to more stats students and see the differences in their efficiency of doing math problems. In this way, we were able to draw a more general conclusion of what kinds of music affect students' efficiency the most by increasing the sample size. However, it is hard to let each individual other than our group members do 4 problem sets because it is time-consuming. Additionally, if we ask each individual to randomly do one of the problem sets, we would fail to eliminate the individual effect and the residual would be large.

During the experiment, participants would need to individually finish math questions under four conditions, which are the four corners of the factorial design: fast-pace song having lyrics, slow-pace song having lyrics, fast-pace song with no lyrics, and slow-pace song with no lyrics, and then collect their rate of correctness and time to finish the questions. We prepared 4 hard-level sections of the GRE Quantitative Reasoning tests, and each section contained 20 questions, with 80 questions per person in total. In this case, we assumed that the difficulty level will not be a nuisance factor in our experiment. There were three different songs in each combination, and each participant was randomly assigned to one song within each category. Although an individual song might still cause a single effect to a participant, we would expect that effect to be averaged out across all the participants.

III. Implementation and Results

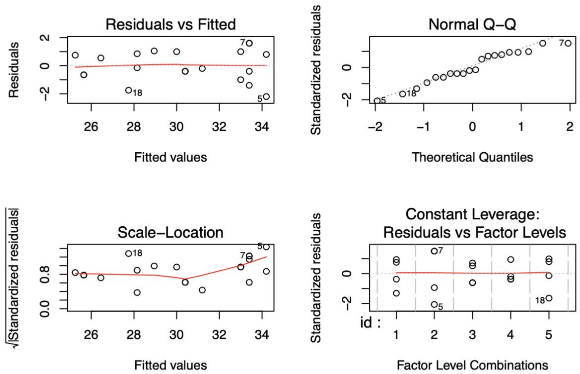
All the data were collected through our five group members, basically, students who are stats majors and have similar knowledge in math. The participants were randomly assigned to four songs belonging to four different categories and had access to four hard-level sections of the GRE math questions online. Five participants conducted the experiment on the same day without break as well as following the same section order and treatment order to eliminate the effects of the potentially confounding variable.

The model we choose for time was time = $ß_0 + ß_{id\_i} *id_i + ß_1 *pace +ß_2* lyric + ß_3*pace*lyric$. Based on the F test performed on each variable in the ANOVA table, we could see that only the individual effect was significant with a p-value of 0.00001188, while pace, lyric, and the interaction term could not significantly affect the time to finish math questions.

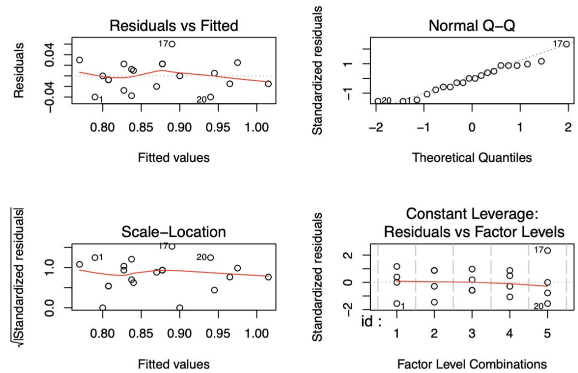
The model we choose for correctness was correctness = $ß_0 + ß_1$ *id +$ß_2$* lyric + $ß_3$*pace*lyric + $ß_4$*pace. According to the ANOVA table, the blocking effect and pace are significant with p-values of 0.00006649 and 0.01961. Also, the interaction term is significant at the 10% significance level. Therefore, we can conclude that blocking, pace, and interaction between pace and lyric would significantly affect stats major students' correctness on solving GRE math problems.

The summary output for the time model showed none of the coefficients of pace, lyrics, or interaction term were significant. While the blocking effect was large and the coefficients of ids were all significant. Based on the summary output for the correctness model, the coefficient of pace equals -0.02, showing that the correctness of doing math problems when listening to fast-paced music is mostly lower compared to slow-paced music for each participant. At the same time, different individuals also had divergent correctness (especially for the first and the third participant). For the third, and fifth participant, the individual effect reflects on the high correctness and short time spent on the problems. This is because these two individuals have done those problems before. For the first participant, the individual effect reflects on the low correctness and long time spent on the problems, this is because she did it early in the morning. Thus, all of the individual effects make sense.

From the diagnostic plots, we could see that both models fit well. There was no pattern in the residual plots, no heteroskedasticity, and the normality assumption was not violated. Hence no transformation is needed to implement for either model.



Residual Plots for Model 1



Residual Plots for Model 2

IV. Conclusion

From our first model, we found that neither the pace nor the lyrics of music have any effects on the time participants spent on the math problems while we also found each individual had different time patterns. From the second model, we found that when doing math problems, listening to fast-paced music would cause correctness to reduce by 0.02 which proved part of our hypothesis that the fast pace music could decrease the correctness of completing math problems. Also, different individuals had divergent correctness (especially for the first and the third participant). Thus, we would recommend statistics students to listen to slow-paced songs, which might help increase their correctness.

Ambiguity also exists in the experiment. Firstly, the randomization process needs to be improved. During the process we did not randomize the order of treatments, which may cause performance to systematically deteriorate over the course of the study, and that might look like it's being caused by each of the treatments. Secondly, other factors, including the time of the day doing the test, the characteristics of carefulness of individuals may influence the result of time students spend on doing the problems and their correctness are counting towards individual effects. Those factors may also be important predictors of the results. Based on the lack of randomization and important influencers, the sample of 5 students could not be representative to all the statistic students and the result could not be generalized to a wilder population.

Further studies could be done to explore which specific type of slow-paced music, such as classic or pop music, has the most significant effects on students' performance in math problems. Also, the lyrics of music in our experiment were all in English, which was our second language. Future experiments could investigate whether music that has lyrics in participants' first-language could be effective in improving or diminishing the correctness in students' math problems solving.

Reference

Lehmann, Janina A M, and Tina Seufert. “The Influence of Background Music on Learning in the Light of Different Theoretical Perspectives and the Role of Working Memory Capacity.” Frontiers in psychology vol. 8 1902. 31 Oct. 2017, doi:10.3389/fpsyg.2017.01902.

Stat 424 Project

Rufan Yao, Zimo Wan, Yuqi Niu, Jing Duan, Yanxu Guo

12/14/2020

```r
library("dplyr")
library("ggplot2")
library("readr")
library("reshape2")
library("tufte")
library("ggpubr")
```

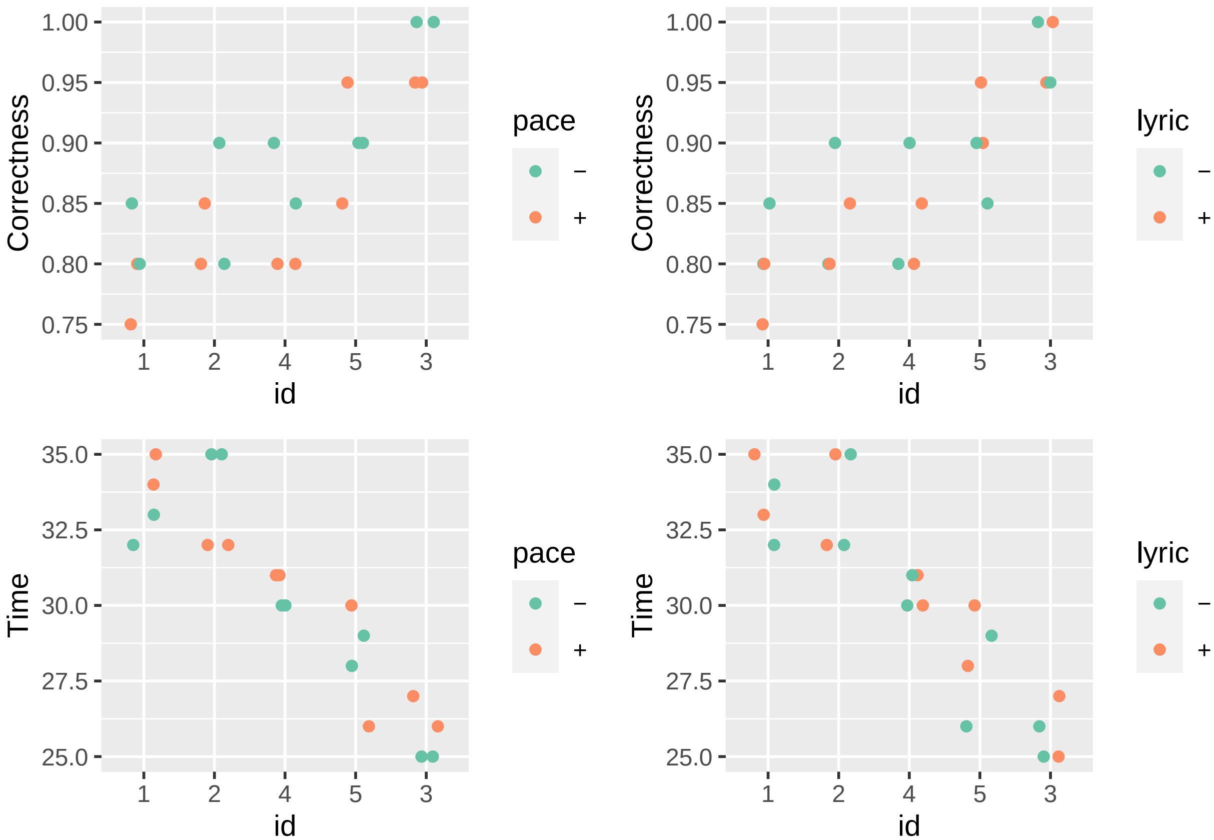
```r
our.data <- data.frame(
   id=rep(1:5,each=4),
   pace=c("+","+","-","-","+","+","-","-","+","+",
          "-","-","+","+","-","-","+","+","-","-"),
   lyric=c("+","-","+","-","+","-","+","-","+","-",
           "+","-","+","-","+","-","+","-","+","-"),
   time=c(35,34,33,32,32,32,35,35,27,26,
          25,25,31,31,30,30,30,26,28,29),
   correctness=c(0.75,0.8,0.8,0.85,0.85,0.8,0.8,0.9,0.95,0.95,
                 1,1,0.8,0.8,0.85,0.9,0.95,0.85,0.9,0.9)
)

fcn = function(x)ifelse(x == "-", -1, 1)
our.data.coded <- our.data %>%
  mutate_at(vars(pace, lyric),fcn) %>%
  mutate(id = as.factor(id))
```

```r
A = ggplot(our.data) +
  geom_point(aes(x = reorder(id, correctness, mean),
                 y = correctness, col = pace),
             position = position_jitter(w = 0.2, h = 0))+
  scale_color_brewer(palette = "Set2") +
  xlab("id") + ylab("Correctness")
B = ggplot(our.data) +
  geom_point(aes(x = reorder(id, correctness, mean),
                 y = correctness, col = lyric),
             position = position_jitter(w = 0.2, h = 0))+
  scale_color_brewer(palette = "Set2") +
  xlab("id") + ylab("Correctness")
C = ggplot(our.data) +
  geom_point(aes(x = reorder(id, correctness, mean),
                 y = time, col = pace),
             position = position_jitter(w = 0.2, h = 0))+
  scale_color_brewer(palette = "Set2") +
```

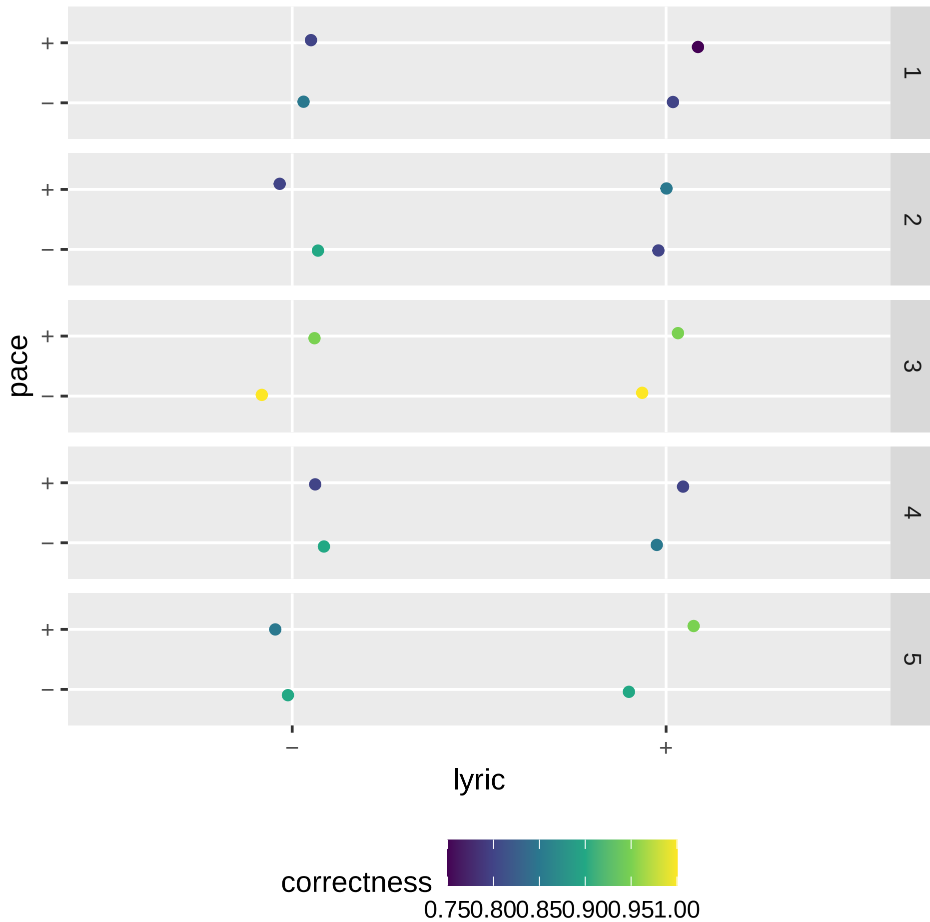
```
  xlab("id") + ylab("Time")
D = ggplot(our.data) +
  geom_point(aes(x = reorder(id, correctness, mean),
                 y = time, col = lyric),
             position = position_jitter(w = 0.2, h = 0))+
  scale_color_brewer(palette = "Set2") +
  xlab("id") + ylab("Time")

ggarrange(A, B, C, D, ncol = 2, nrow = 2)
```




```
#observations:
#When there is no lyric, listening to the lower-pace music resulted in
#a higher correctness cross all individuals.
#When there is lyric, the pace of music does not seem to have the same
#effects on different individuals.
#Thus, there is an interaction effect between the lyrics and
#the pace of music (which is also verified in our model analysis)
```

```
ggplot(our.data) +
  geom_point(aes(x = lyric, y = pace, col = correctness),
             position = position_jitter(w = 0.1, h = 0.1))+
  facet_grid(id ~ .) +
  scale_color_viridis_c() +
  theme(legend.position = "bottom")
```




```r
model1<-lm(time~id+pace*lyric,our.data.coded)
model2<-lm(correctness~id+pace*lyric,our.data.coded)

summary(model1)
```

```
##
## Call:
## lm(formula = time ~ id + pace * lyric, data = our.data.coded)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2000 -0.6500 -0.1750  0.8875  1.6000
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.350e+01  6.877e-01  48.714 3.67e-15 ***
## id2          9.930e-16  9.725e-01   0.000 1.000000
## id3         -7.750e+00  9.725e-01  -7.969 3.91e-06 ***
## id4         -3.000e+00  9.725e-01  -3.085 0.009455 **
## id5         -5.250e+00  9.725e-01  -5.398 0.000161 ***
## pace         1.000e-01  3.075e-01   0.325 0.750662
## lyric        3.000e-01  3.075e-01   0.975 0.348582
## pace:lyric   3.000e-01  3.075e-01   0.975 0.348582
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.375 on 12 degrees of freedom
## Multiple R-squared:  0.891,  Adjusted R-squared:  0.8274
## F-statistic: 14.01 on 7 and 12 DF,  p-value: 6.556e-05
```

```
summary(model2)
```

```
##
## Call:
## lm(formula = correctness ~ id + pace * lyric, data = our.data.coded)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.04000 -0.01625  0.00000  0.02250  0.06000
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.80000    0.01662  48.151 4.21e-15 ***
## id2          0.03750    0.02350   1.596  0.13648
## id3          0.17500    0.02350   7.448 7.76e-06 ***
## id4          0.03750    0.02350   1.596  0.13648
## id5          0.10000    0.02350   4.256  0.00112 **
## pace        -0.02000    0.00743  -2.692  0.01961 *
## lyric       -0.00500    0.00743  -0.673  0.51374
## pace:lyric   0.01500    0.00743   2.019  0.06643 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03323 on 12 degrees of freedom
## Multiple R-squared:  0.8701, Adjusted R-squared:  0.7943
## F-statistic: 11.48 on 7 and 12 DF,  p-value: 0.0001784
```

```
anova(model1)
```

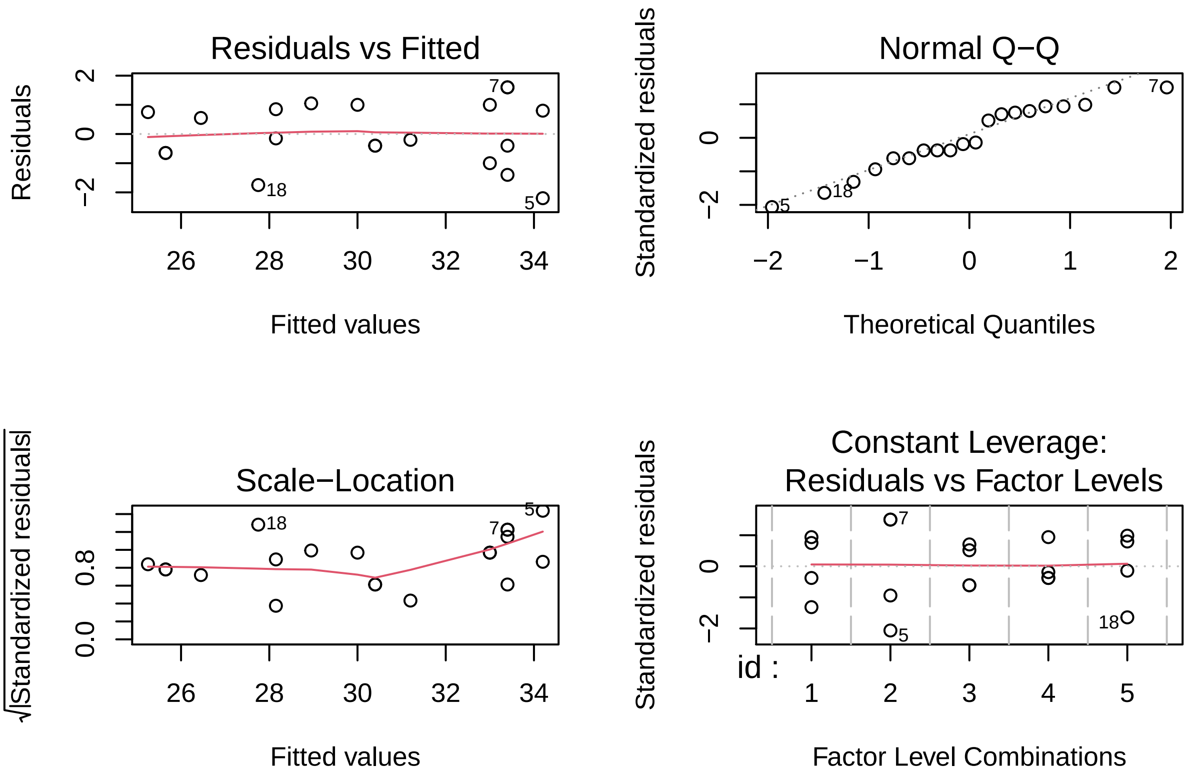
```
## Analysis of Variance Table
##
## Response: time
##            Df Sum Sq Mean Sq F value    Pr(>F)
## id          4  181.7  45.425 24.0132 1.188e-05 ***
## pace        1    0.2   0.200  0.1057    0.7507
## lyric       1    1.8   1.800  0.9515    0.3486
## pace:lyric  1    1.8   1.800  0.9515    0.3486
## Residuals  12   22.7   1.892
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(model2)
```

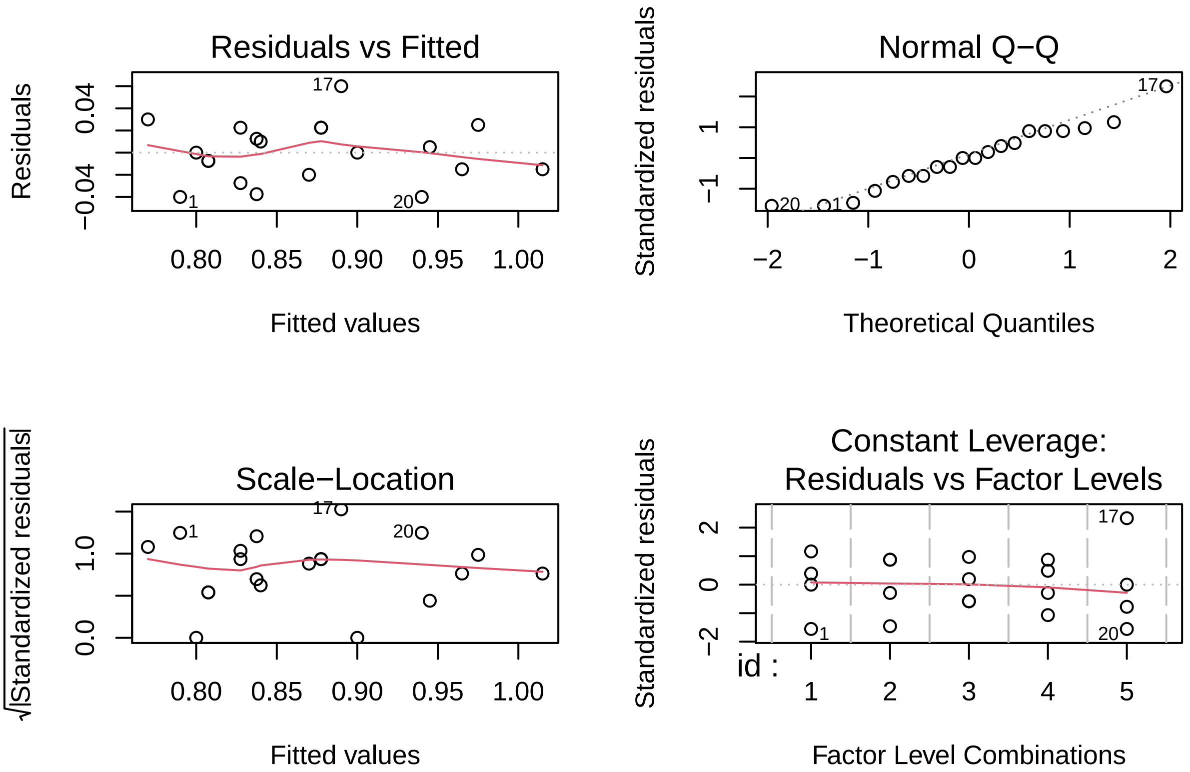
```
## Analysis of Variance Table
##
## Response: correctness
```

```
##              Df  Sum Sq   Mean Sq F value    Pr(>F)
## id            4 0.07575 0.0189375 17.1509 6.649e-05 ***
## pace          1 0.00800 0.0080000  7.2453   0.01961 *
## lyric         1 0.00050 0.0005000  0.4528   0.51374
## pace:lyric    1 0.00450 0.0045000  4.0755   0.06643 .
## Residuals    12 0.01325 0.0011042
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
par(mfrow=c(2,2))
plot(model1)
```




```r
par(mfrow=c(2,2))
plot(model2)
```


Residuals vs Fitted
Residuals
Fitted values
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Fitted values
Constant Leverage:
Residuals vs Factor Levels
Standardized residuals
id :
Factor Level Combinations