LAB#6

2025-05-14

PSTAT 122 Lab 6

Part 1: Written Report

Method

The Question:

Can we influence how long you can hold your breath with visual clues or a lack thereof?

The goal of this lab was to see whether visual cues such as watching a stopwatch or closing your eyes would affect how long someone is able to hold their breath. There are three different conditions:

Control(move normally while breath-holding) Looking at a timer Eyes closed

This was done using a Generalized Randomized Block Design (GRBD), two students were holding the timer which service as the blocks. Each timer was in charge of recording the breath-holding times for the other students. Each student would be randomly assigned to a randomly three treatment conditions using the following R code:

```
sample(c("control", "look at stopwatch", "eyes closed"), 1)
```

This was run for every students by the timer, who then record the breath-holding time under that assigned condition. Then the time would be written on the blackboard.

There were also several technical issues during the experiment. Some students were feeling sick, which likely affected their breath-holding performance and introduced extrea variation into the data. Also, there was a confusion among students about when exactly to stop holding their breath. To fix this, we agreed as a group that everyone would stop at the point of breathing out to keep things consistent. Some timers were slightly delayed in starting or stopping their phone stopwatch. Since we couldn't talk while holding our breath, we all agreed to use a thumbs-up to signal when we were ready to begin.

For our sample size calculations and variance estimates, we used data collected from our own experiment as a pilot study. This let us estimate within-group variability under real experimental conditions.

Each observation was recorded with the breath-holding time, the assigned treatment, and the timer(block) who recorded it. These were entered into an excel with columns and down below is the data frame:

- Time
- Treatment
- Block

```
library(readxl)
# Input Excel file
breath_df <- read_excel("Lab6_Time.xlsx")
```

Results:

From Daily Check 11:

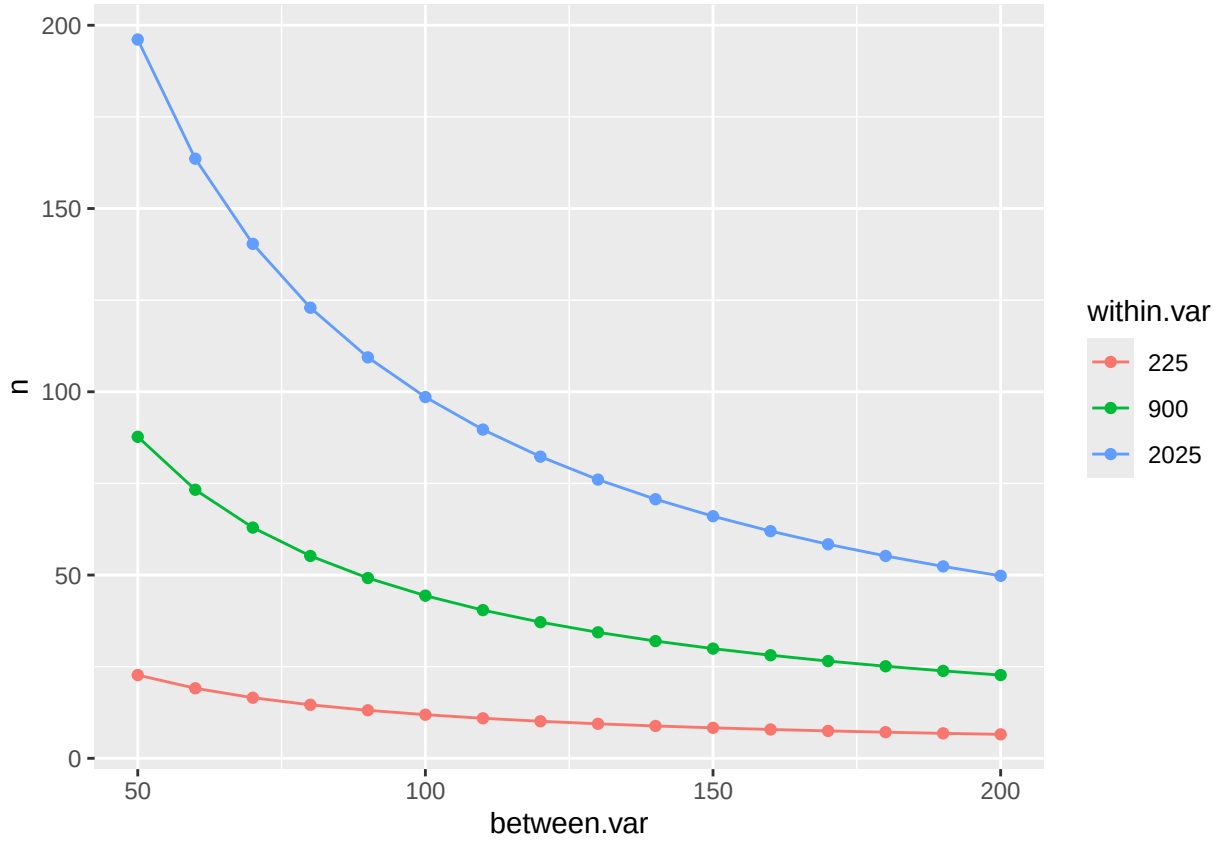
```r
# Define group means:
#control = 120s, look at stopwatch = 120 + 20 = 140s, eyes closed = 120+10 = 130s
groupmeans <- c(120, 140, 130)

# Calculate required sample size per group for the means, using the variance estimate
power.anova.test(groups = length(groupmeans),
                 between.var = var(groupmeans),
                 within.var = 900, # within variance 900 variance
                 #how much variability is there inherently in how long people can hold their breath
                 power = 0.8, sig.level = 0.05, n = NULL)
```

Above is the code for power analysis. The power analysis indicated that a sample size of approximately 45 participants per group is required to detect a moderate effect, assuming a within-group variance of 900 s^2 and a significance level of 0.05.

Sample size graph

```r
# Create a sequence between the groups
between.var <- seq(50, 200, by = 10)

# Run a loop which each of the value in it

# first column sample size value
n_var1 <- NA
for(i in 1: length(between.var)){
  n_var1[i] <- power.anova.test(groups = 3,# 3 treatment groups
                   between.var = between.var[i], # each individual between the var
                   within.var = 900,
                   power = 0.8, sig.level = 0.05, n = NULL)$n # $n extracts the required sample size
}

# second column between var value
n_var2 <- NA
for(i in 1: length(between.var)){
  n_var2[i] <- power.anova.test(groups = 3,
                   between.var = between.var[i],
                   within.var = 225,
                   power = 0.8, sig.level = 0.05, n = NULL)$n
}

# within var value
n_var3 <- NA
for(i in 1: length(between.var)){
  n_var3[i] <- power.anova.test(groups = 3,
                   between.var = between.var[i],
                   within.var = 2025,
                   power = 0.8, sig.level = 0.05, n = NULL)$n
}
```

```r
#data frame
sample_sizes <- data.frame(
  n = c(n_var1, n_var2, n_var3),
  between.var = rep(between.var, 3), # repeat give each n_var value
  within.var = c(rep("900", length(n_var1)),
                 rep("225", length(n_var2)),
                 rep("2025", length(n_var3)))
)

sample_sizes$within.var <- factor(sample_sizes$within.var,
                                  levels = c("225", "900", "2025"))

# Plot the graph
library(ggplot2)
ggplot(data = sample_sizes, mapping = aes(x = between.var,
                                          y = n,
                                          group = within.var,
                                          color = within.var)) +
  geom_point() + geom_line()
```



Commentary:

This sample size graph shows how the required sample size per group changes depanding on two main factors: the group variance and the effect size. For the within-group variance: 900 is the estimate taken from the research study (Parkes, 2006), with the green line in the graph. 225 is a possible lower variance scenario 2025 is the bad case scenario which has a much higher variance groups.

The x-axis which is the between-group variance represents the effect size that shows how different the group means are from each other. A larger effect size makes it easier to detect a difference, so the required sample size per group is smaller.

The between variance is determined by the group means. When the group means are more spread out, the between variance increases. If the group means are closer together, the between-group variance decreases, and a larger sample size is needed to detect a difference.

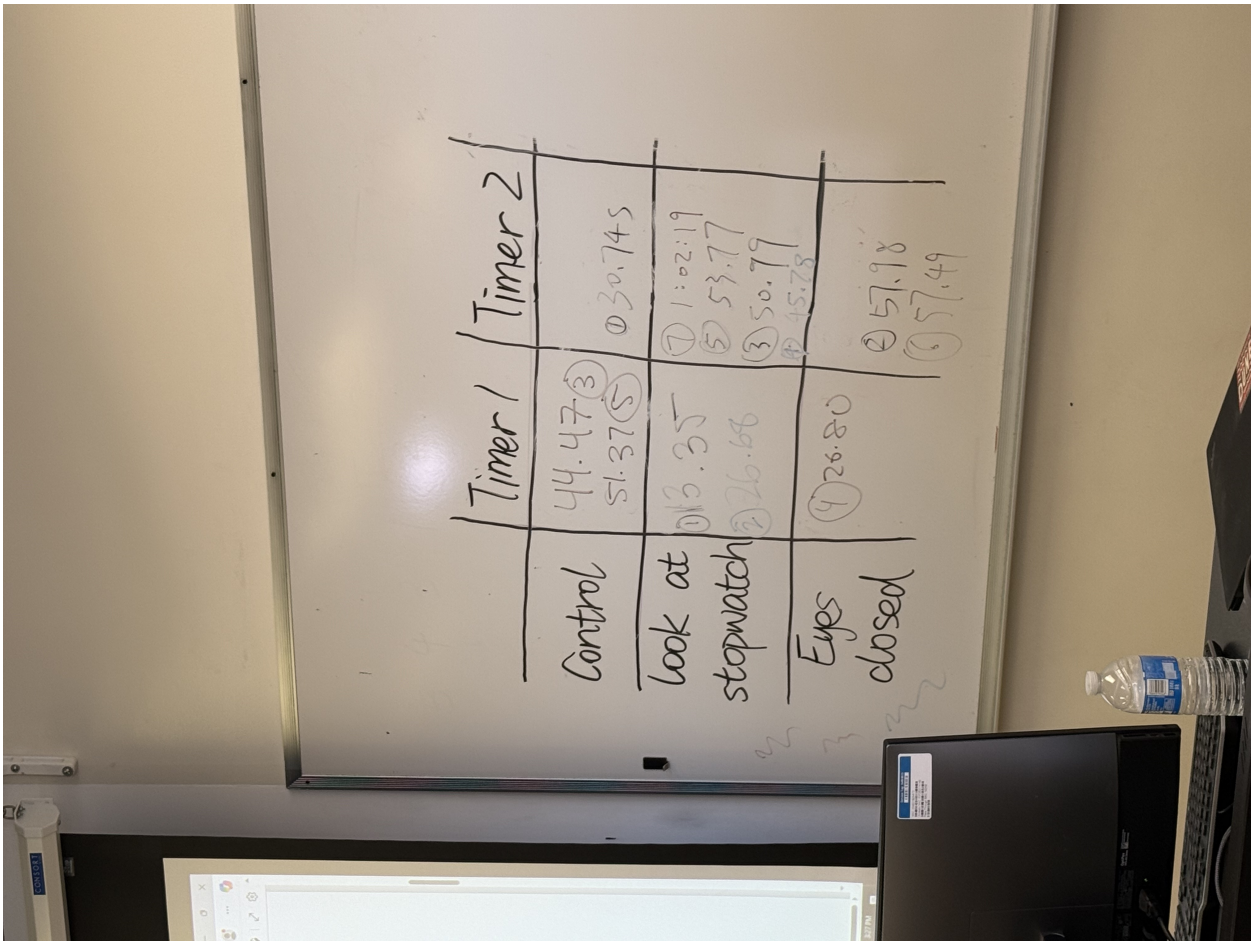
Based on the graph, if we use a within-group variance of 900 s^2 (from Parkes, 2006) and a moderate effect size, the required sample size per group is about 45.

We obtained a variance estimate of 900 s^2 from a previous study (Parkes, 2006).

Image:

This is the image of the whiteboard:

```
# Show image of the whiteboard
knitr::include_graphics("/Users/andy/Desktop/UCSBCourses/PSTAT122/LAB6/classLab6.jpeg")
```



Data Visualization for Breath-Holding Time by Treatment

```r
library(ggplot2)

# Increase the base text size for better readability in the plot
theme_update(text = element_text(size = 12))

# Set up the plot: x-axis is treatment, y-axis is Time, color points by Block
ggplot(data = breath_df, mapping = aes(x = Treatment, y = Time, color = Block)) +

  # Add jittered points to show all individual observations, with slight horizontal spread
  geom_jitter(width = 0.2, size = 4) +

  # Add a descriptive title to the plots
  ggtitle("Breath-Holding Time by Treatment")
```

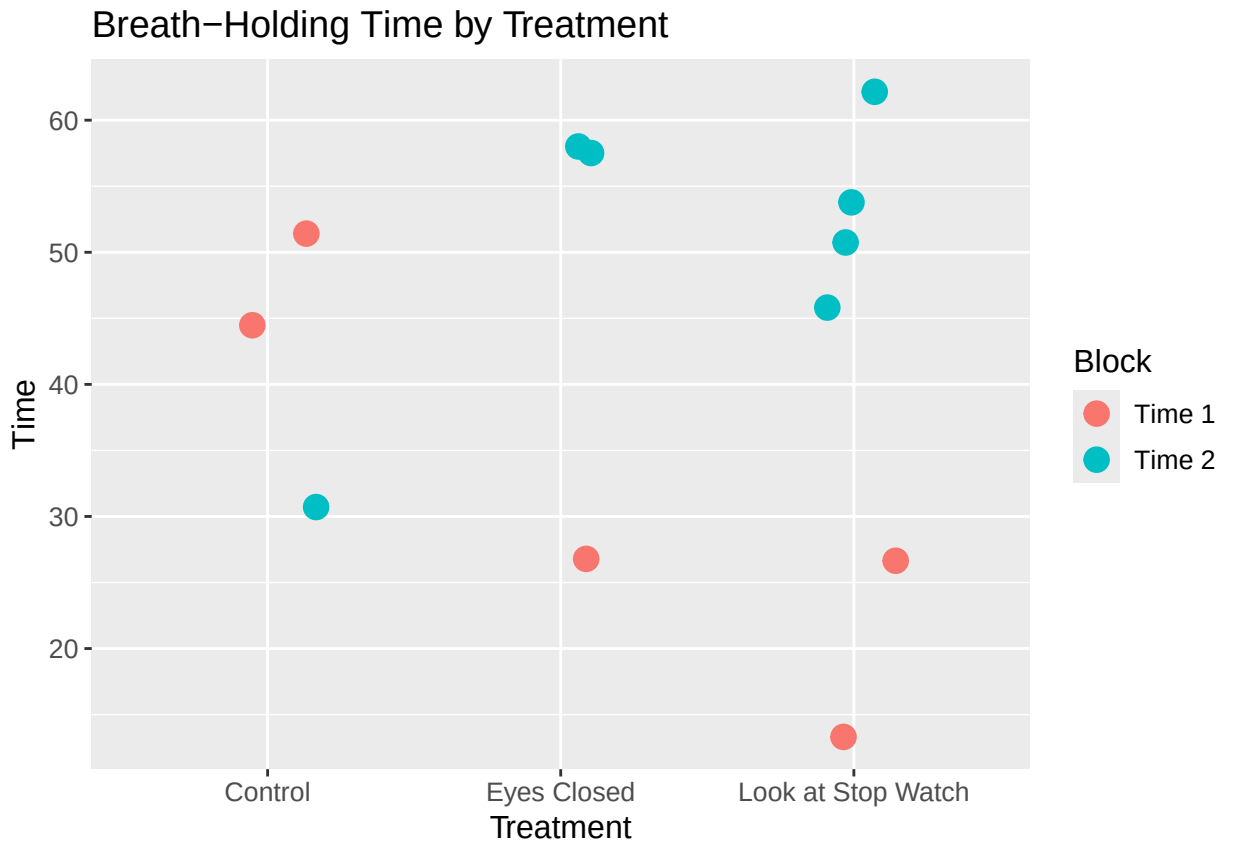


Figure:

Each point represents an individual breath-holding time under different treatments and blocks. We used a jitter plot instead of a box plot because each group/block has only a small number of data points, and plotting all individual values gives a more accurate representation of the data distribution.

Interpretation:

The plot suggests that the Eyes Closed treatment generally led to longer breath-holding times, especially in Block 2. We can see that the breath-holding times vary quite a bit, not just between the different treatments, but also from one block to another. This might due to differences between the students technical issues.

Report the Null Hypotheses

The null hypothesis is a default statement that there is no different or no association between groups or variables. It represents the idea that any observed differences are due to chance or random variation.

This is what we assume to be true unless you have strong evidence against it.

$$\mu_C = \text{the mean breath-holding time under the Control treatment}$$

$$\mu_E = \text{the mean breath-holding time under the Eyes Closed treatment}$$

$$\mu_S = \text{the mean breath-holding time under the Look at Stop Watch treatment}$$

$$H_0 : \text{the mean breath-holding time is the same for all treatments.}$$

$$H_0 : \mu_C = \mu_E = \mu_S$$

Report the Alternative Hypotheses

The alternative hypothesis is the statement that there is a difference, or an association. It represents what we are trying to find evidence for in we experiment or study.

The alternative hypothesis is what we hope to support with the data.

At least one of μ_C , μ_E, or μ_S is different from the others.

$$H_A : \text{At least one of } \mu_C, \mu_E, \text{ or } \mu_S \text{ is different from the others}$$

Hypothesis Test (ANOVA)

R code:

Summary:

We conducted ANOVA to examine whether the types of treatment had a significant effect on breath-holding time, accounting for block effects. At $\alpha = 0.05$, we fail to reject the null hypothesis. The p-value for treatment was 0.85, which we don't have statistically significant different in breath-holding time across the different visual conditions. In other words, the types of treatment does not appear to affect breath-holding time.

However, we found a statistically significant effect of block (p = 0.043), meaning that the breath-holding times differed between two timer blocks. This could be due to individual students differences and in timing accuracy by the timers, or other factors such as some students not feeling well or confusion about when stop holding their breath during the experiment.

Pairwise comparisons:

```
pairwise_results <- TukeyHSD(breath_model)
print(pairwise_results)
```

Pairwise comparisons between all treatment groups were performed using Tukey's HSD test. None of the differences between the groups were statistically significant. For example, the different in mean breath-holding time between the "Eyes Closed" and "Control" groups was about 5 seconds ($p = 0.89$), and between "Look at Stop Watch" and "Control" was nearly zero ($p = 1.00$). All p-values were much greater than 0.05, saying that there is no evidence of a significant difference between any of the treatment conditions.

The comparison between the two blocks showed a average times of about 18 seconds, with a p-value just above 0.05. This is possible but not statistically significant difference between timers.

Overall, the pairwise comparisons support the main ANOVA result. There is no significant different for visual cue or block had a meaningful impact on breath-holding time.

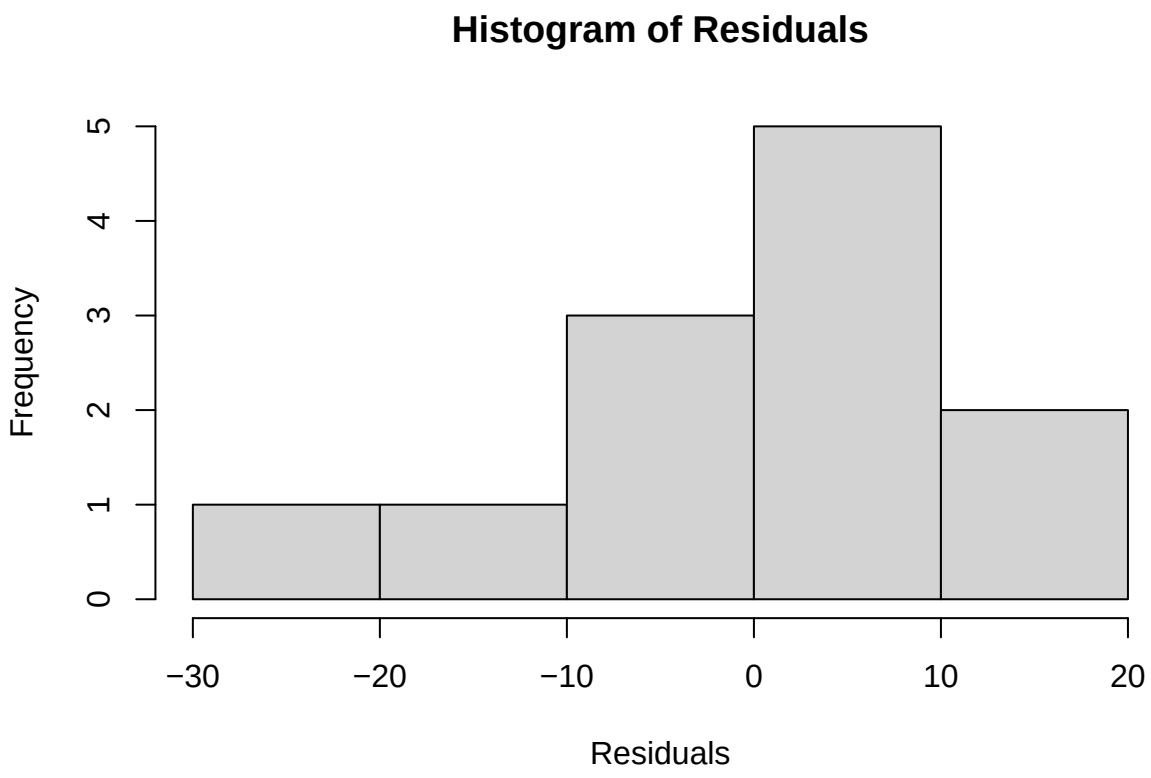
Model Checking

Model checking involves evaluating whether the assumptions required for ANOVA are met in our data. The assumptions we have are normality of the residuals, equal variances across groups, and no systematic structure in the data. Checking these assumptions ensures that the results and p-values from the ANOVA can be trusted.

Normality of the data: The residuals (differences between observed and predicted values) from the ANOVA model should be approximately normally distributed. This means that the errors are symmetrically distributed around zero, without extreme skewness or outliers. To check this assumption, we used a histogram of the residuals, a Q-Q plot, and the Shapiro-Wilk test for normality.

Histogram:

```
# Check normality
hist(breath_model$residuals, xlab="Residuals", main="Histogram of Residuals")
```

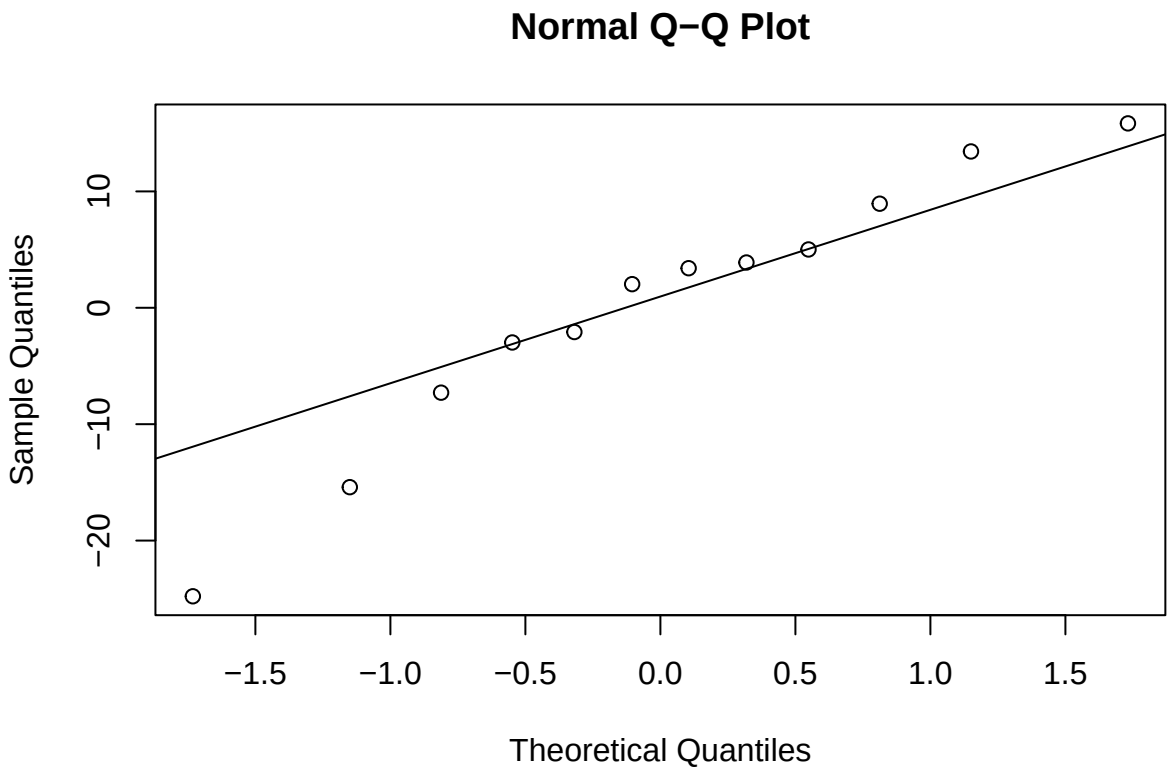


Interpretation:

This is the histogram of residuals appears to be right symmetric and centered around zero point. This shows that the normality assumption for the ANOVA model is reasonable since it is symmetric and most values are near zero.

Normal Q-Q plot:

```
qqnorm(breath_model$residuals)
qqline(breath_model$residuals)
```



Interpretation:

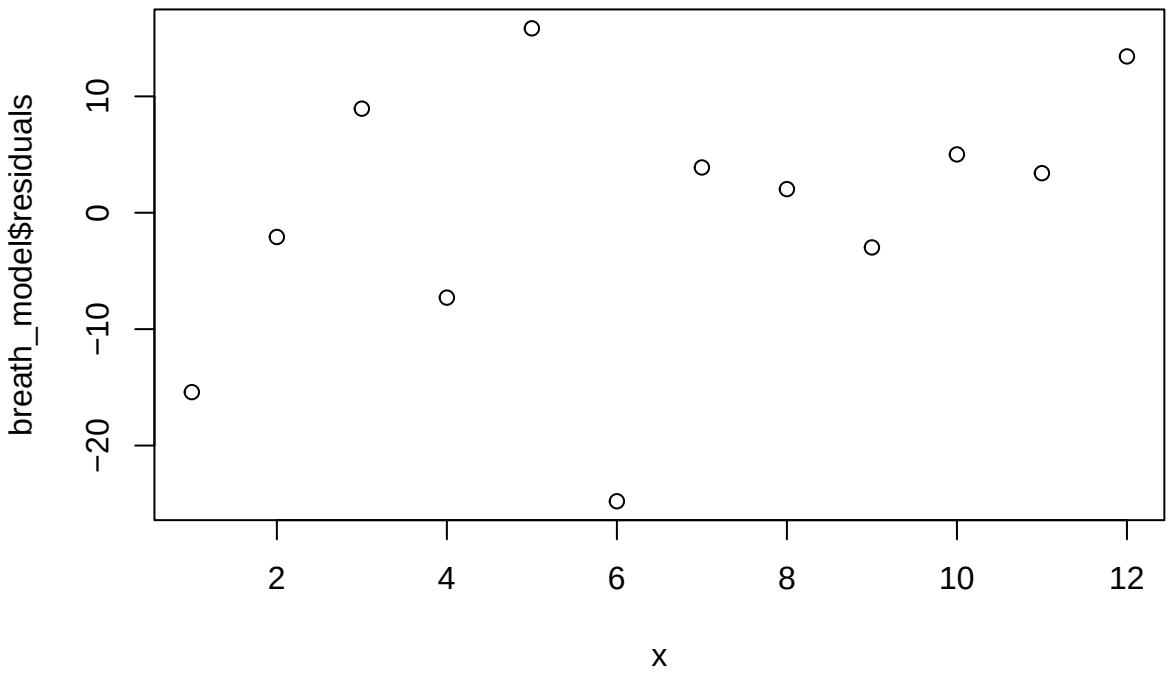
The Q-Q plot shows that the points closely follow the straight line, indicating that the residuals are approximately normally distributed.

The Shapiro-Wilk test gives a p-value of 0.5916, which much greater than 0.05. This means we fail to reject the null hypothesis and conclude that there is no significant evidence the residuals deviate from normality.

No Structure to the Data:

There should be no systematic pattern in the residuals related to the order in which the data were collected, or to any other variable not included in the model. In other words, the data should be independent and not show trends over time or sequence.

```
# Check for structure over order
x <- 1:length(breath_model$residuals)
plot(breath_model$residuals ~ x)
```



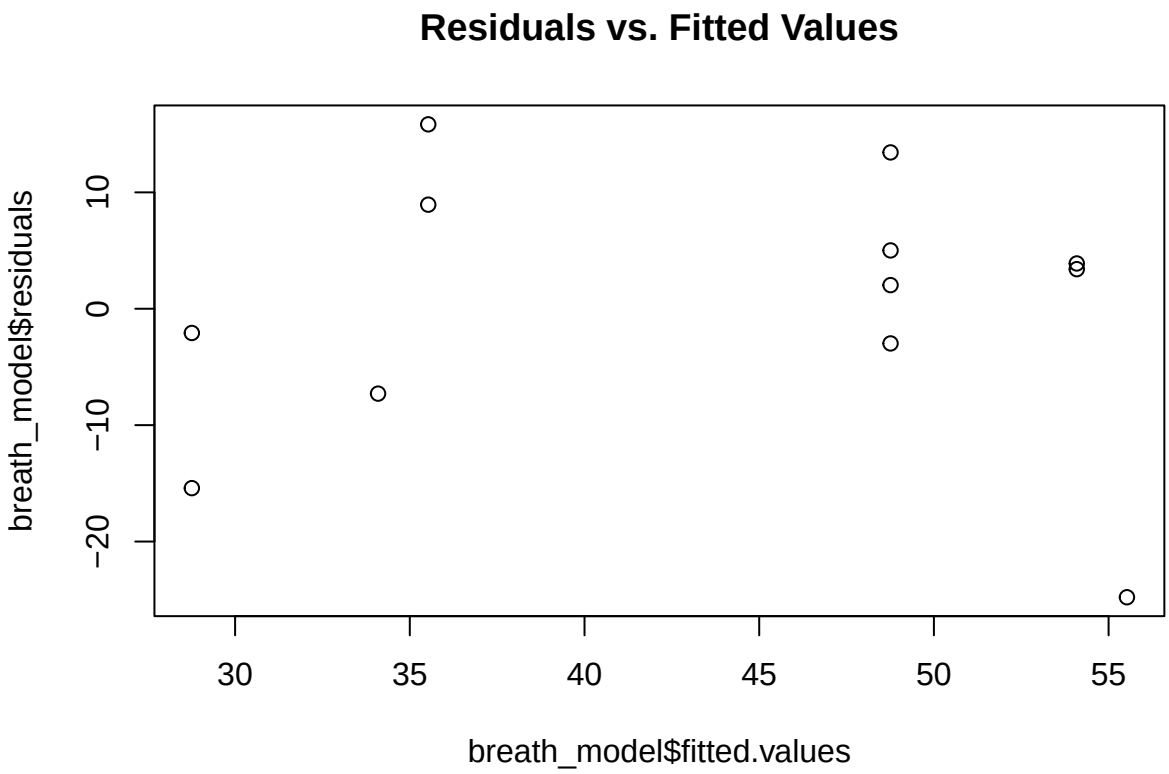
Interpretation:

The plot of residuals versus order of data collection does not show any clear trend or systematic pattern. This suggests that there is no evidence of structure or dependence related to the order in which the data were collected, supporting the independence assumption of ANOVA.

Equal Variances:

The residuals should have similar variability across all treatment groups and across the range of fitted values. This means that the amount of "scatter" in the residuals should be roughly the same for all groups—there should not be more spread in one group than another.

```
# 3. Check equal variances
plot(breath_model$residuals ~ breath_model$fitted.values, main="Residuals vs. Fitted Values")
```



Interpretation:

The plot of residuals versus fitted values shows that the spread of residuals is fairly consistent across all fitted values, with no clear pattern or systematic change in variability. This suggests that the equal variances assumption for ANOVA is reasonably met.

Discussion

In this experiment, we tested if visual cues such as watching a stopwatch or closing your eyes would affect how long people can hold their breath. Students were randomly assigned to two blocks(timers) and then randomly assigned to one of the three treatment groups which are control, looking at a timer, or eyes closed. After collecting data and running an ANOVA, we did not find statistically significant evidence that the treatment type influenced breath-holding time. However, we did observe a statistically significant block effect, suggesting that differences between timers affected the results.

There were several drawbacks to our study. While students were randomly assigned to blocks, the randomization within blocks was not always ideal. Some treatment assignments had to be adjusted or chosen by students due to the certain treatment (looking at the stopwatch) were overly picked by the systems. This might have reduced the randomness of assignment. Another major drawbacks was the sample size. Our sample had approximately 7 participants per group, whereas our power analysis indicated we would need about 45 participants per group to reliably detect differences, assuming a moderate effect size and within-group variance of 900 s^2 (based on Parkes, 2006). The small sample size means our study had low statistical power, increasing the risk of Type II error. Other factors like technical issues such as inconsistent timing, students feeling unwell, and miscommunication about when to stop holding their breath introduced variability into the measurements. These factors further reduced our ability to isolate the effects of the treatments.

If we do this experiment again, we'd sim to have a larger sample size, ensure random assignments for both

blocks and treatments. We would also work on making the timing and communication during the experiment more consistent. While we did not find statistically significant difference in breath-holding times between the different groups, the experience highlighted the importance of careful planning to minimize bias and variability in experimental design.

References

Parkes MJ. Breath-holding and its breakpoint. Experimental physiology. 2006 Jan;91(1):1-5.

Part 2: The 2^2 Factorial Design

Enter the data from excerise 6.12

```
# Enter data
med1_time12 <- c(21,22,23,28,20,26)
med2_time12 <- c(25,26,24,25,29,27)
med1_time18 <- c(37,39,38,38,35,36)
med2_time18 <- c(31,34,29,33,30,35)

virus_df <- data.frame(
  growth = c(med1_time12, med2_time12, med1_time18, med2_time18),
  medium = factor(
    c(rep("1", 6), rep("2", 6), rep("1", 6), rep("2", 6))
  ),
    time = factor(
      c(rep("12", 12), rep("18", 12))
    )
)

sample(virus_df)
```

```
##    medium time growth
## 1       1   12     21
## 2       1   12     22
## 3       1   12     23
## 4       1   12     28
## 5       1   12     20
## 6       1   12     26
## 7       2   12     25
## 8       2   12     26
## 9       2   12     24
## 10      2   12     25
## 11      2   12     29
## 12      2   12     27
## 13      1   18     37
## 14      1   18     39
## 15      1   18     38
## 16      1   18     38
## 17      1   18     35
## 18      1   18     36
## 19      2   18     31
## 20      2   18     34
```

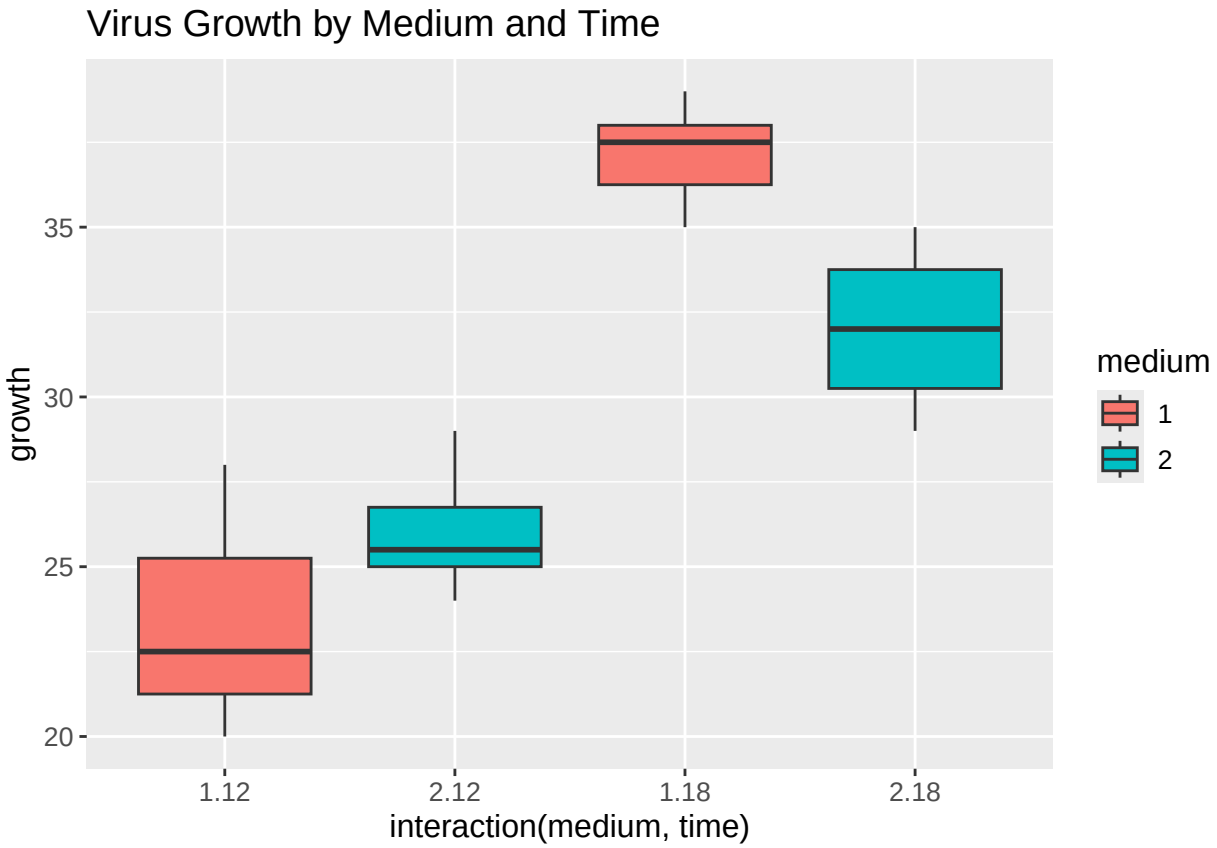
```
## 21       2    18      29
## 22       2    18      33
## 23       2    18      30
## 24       2    18      35
```

1) Create two appropriate graphical representations of the data, and briefly comment on what

you observe.

Boxplot of Growth by Group (Medium and Time)

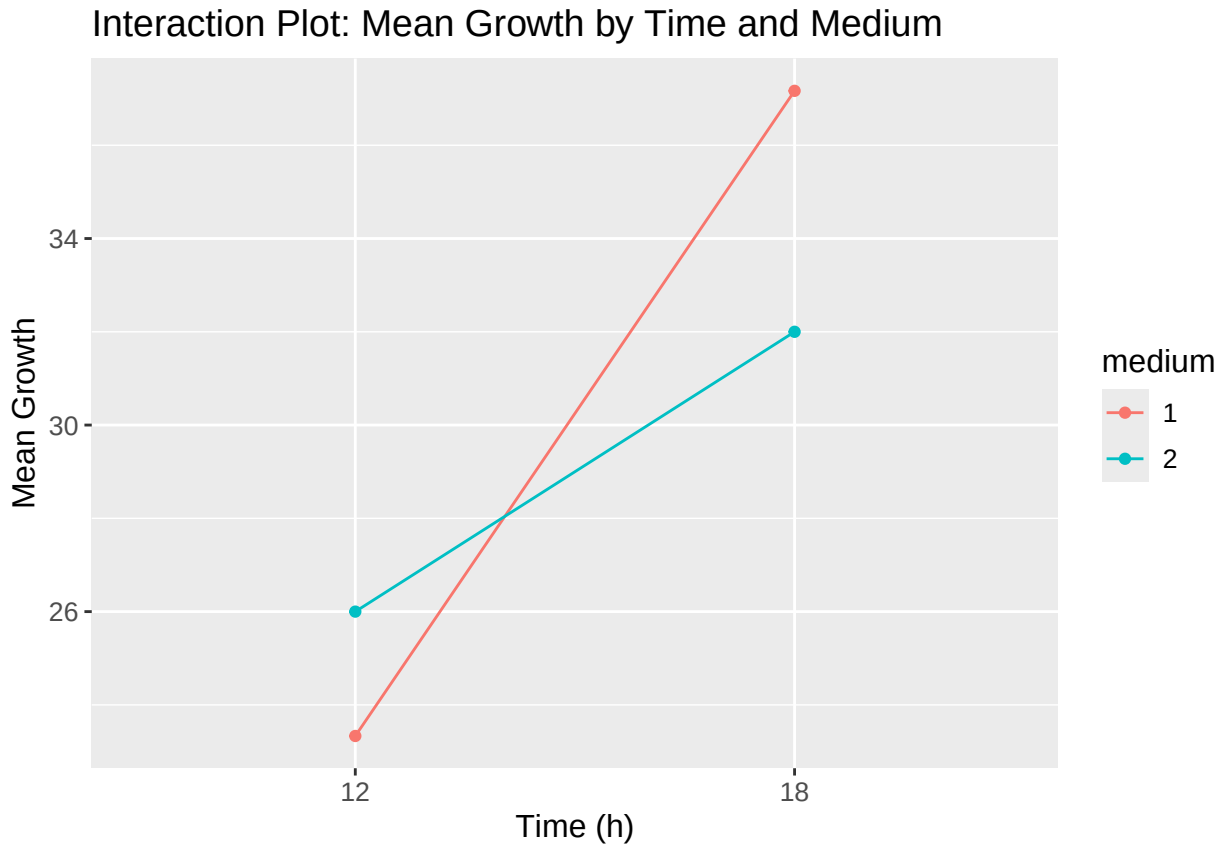
```
library(ggplot2)
theme_update(text = element_text(size = 12))
ggplot(data = virus_df, mapping=aes(x = interaction(medium, time), y = growth, fill = medium)) +
geom_boxplot() +
  ggtitle("Virus Growth by Medium and Time")
```



The boxplot shows that virus growth increases from 12 hours to 18 hours for both media. Medium 1 consistently results in higher virus growth than Medium 2 at both time points. The variability in growth appears slightly higher for Medium 2 at 18 hours, while growth for Medium 1 at 18 hours is not only higher but also more consistent which is less variable.

Interaction Plot (Mean Growth by Medium and Time)

```r
library(dplyr)
virus_tmp <- virus_df %>%
group_by(medium, time) %>%
summarize(mean_growth = mean(growth), .groups = "drop")

theme_update(text = element_text(size = 12))
ggplot(data = virus_tmp, aes(x = time, y = mean_growth, color = medium, group = medium)) +
geom_point() + geom_line() +
ggtitle("Interaction Plot: Mean Growth by Time and Medium") +
ylab("Mean Growth") + xlab("Time (h)")
```



The interaction plot shows that mean virus growth for Medium 1 increases sharply from 12 to 18 hours, while growth for Medium 2 also increases but to a lesser extent. This demonstrates that the effect of incubation time on virus growth is greater in Medium 1 than in Medium 2, indicating a significant interaction between medium and time.

2) Run an appropriate statistical analysis on the data.

Let μ_{ij} represent the mean virus growth for culture medium i (i = 1,2) at time j(j = 12,18) hours.

Null hypothesis(H_0): There is no effect of culture medium, time, or their interaction on mean virus growth.

$$H_0 : \text{All group means } (\mu_{ij}) \text{ are equal}$$

Alternative hypothesis (H_A): At least one group mean is different

H_A : At least one group mean is different

```r
# Two-way ANOVA: include main effects for medium and time, and their interaction
# The * operator in R means we include both main effects (medium, time)
anova_result <- aov(growth ~ medium * time, data = virus_df)
summary(anova_result)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## medium       1    9.4     9.4   1.835 0.190617
## time         1  590.0   590.0 115.506 9.29e-10 ***
## medium:time  1   92.0    92.0  18.018 0.000397 ***
## Residuals   20  102.2     5.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Conclusion:

At the $\alpha = 0.05$ significance level, there is strong evidence that both incubation time ($p < 0.001$) and the interaction between culture medium and time ($p < 0.001$) have significant effects on virus growth. There is not sufficient evidence to conclude that the main effect of culture medium alone is significant ($p = 0.019$).

In the context of this experiment, this means that the effect of increasing incubation time on virus growth depends on which medium is used, but the overall difference between the two media, regardless of time, is not statistically significant. The significant interaction indicates that one medium benefits more from increased incubation time than the other.

3) From the analysis, what is the estimate of the impact of the Culture Medium among observations that have undergone a 12 hour growth time?

The estimated impact is the difference in mean virus growth between Medium 2 and Medium 1 at 12 hours:

For Medium 1 (12h): (21 + 22 + 23 + 28 + 20 + 26) / 6 = 23.33

For Medium 2 (12h): (25 + 26 + 24 + 25 + 29 + 27) / 6 = 26.00

Difference = 26.00 - 23.33 = 2.67

4) Find the estimate of the interaction between Culture Medium and Time, and explain what it represents.

$$\big[(\text{Medium 2, 18h}) - (\text{Medium 2, 12h})\big] - \big[(\text{Medium 1, 18h}) - (\text{Medium 1, 12h})\big]$$

For Medium 1 (18h): (37 + 39 + 38 + 38 + 35 + 36) / 6 = 37.17

For Medium 2 (18h): (31 + 34 + 29 + 33 + 30 + 35) / 6 = 32.00

[(32.00) - (26.00)] - [(37.17) - (23.33)] = -7.84