BACS HW9

109062710

4/22/2021

Before answering the questions, let's discuss what Type I Error and Type II Error first.

A Type I Error is known as a false positive and occurs when a researcher incorrectly rejects a true null hypothesis. This means that your report that your findings are significant when in fact they have occurred by chance.

A Type II Error is known as a false negative and occurs when a researcher fails to reject a null hypothesis which is really false. Here a researcher concludes there is not a significant effect, when actually there really is.

Question 1

a. You discover that your colleague wanted to target the general population of Taiwanese users of the product. However, he only collected data from a pool of young consumers, and missed many older customers who you suspect might use the product much less every day.

- Would this scenario create systematic or random error (or both or neither)?

In this scenario, it will be neither systematic nor random error. However, the data that he has are not representing the actual population being measured.

- Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?

Considering that he has a small data set, he has to gather more data. Once the number of data is increased, it will affect $diff$ and sd.

- Will it increase or decrease our power to reject the null hypothesis?

Since Taiwan's population is considered an aging population, my colleague might have missed a great number of older users who are fond of the new product. In addition, the older users might be the ones who are willing to afford the product and use it more often. If the older users were included in the data set, it would increase our power to reject the null hypothesis.

In this case, where the users are mostly young people, our power to reject the null hypothesis is weak.

- Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Since they are many busy young people these days, my colleague might conclude that the null hypothesis should be accepted. Clearly, the null hypothesis should be rejected if the older users were also included. Thus, Type I error is more likely to happen.

b. You find that 20 of the respondents are reporting data from the wrong wearable device, so they should be removed from the data. These 20 people are just like the others in every other respect.

- Would this scenario create systematic or random error (or both or neither)?

Since my colleague has less and less data points in the data set, it might bring the mean value closer to null hypothesis' mean value rather than the true mean value of a complete population. Thus, random error is more likely to happen since the number of data points is small.

- Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?

According to a paper published at NCBI, random error and systematic error can affect *alpha* or P-value.

Clearly, $diff$, sd, n, and $alpha$ will change.

- Will it increase or decrease our power to reject the null hypothesis?

Our power to reject the null hypothesis is getting weaker because there aren't enough data points.

- Which kind of error (Type I or Type II) becomes more likely because of this scenario?

It sounds similar to scenario (a), thus type I error is more likely.

c. A very annoying professor visiting your company has criticized your colleague's "95% confidence" criteria, and has suggested relaxing it to just 90%.

- Would this scenario create systematic or random error (or both or neither)?

Neither of these two errors will happen.

- Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?

Alpha value will change from 0.05 to 0.1.

- Will it increase or decrease our power to reject the null hypothesis?

It will increase the power to reject the null hypothesis because the size of the purple area increases.

- Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Since the numyber of data points is small, Random Error is more likely to happen and Random Error is believed to favor the null hypothesis over the alternative one, stated in this research paper. Also, by relaxing the p-value to 90

d. Your colleague has measured usage times on five weekdays and taken a daily average. But you feel this will under report usage for younger people who are very active on weekends, whereas it over-reports usage of older users.

- Would this scenario create systematic or random error (or both or neither)?

In this scenario, both errors will occur. The reason that systematic error occurs is instead of observing users activity everyday, my colleague chose to observe them 5 days in a week. While the random error occurs because the number of data points might decrease significantly by leaving young users unobserved.

- Which part of the t-statistic or significance (diff, sd, n, alpha) would be affected?

Since the older users are over reported and younger users are under reported, it increases the mean value. Thus, diff, sd, and alpha will increase.

- Will it increase or decrease our power to reject the null hypothesis?

Since the number of over reported old users is high, it will increase the power to reject the null hypothesis.

- Which kind of error (Type I or Type II) becomes more likely because of this scenario?

Since the number of over reported old users is high, it convinces my colleague to reject the null hypothesis. In reality, if he observed old and young users everyday in a week, the null hypothesis should be accepted. In this scenario, my colleague is incorrectly rejects the null hypothesis. Thus, Type I Error is more likely to happen.

Question 2

Load packages

```
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(reshape2)
```

Load data

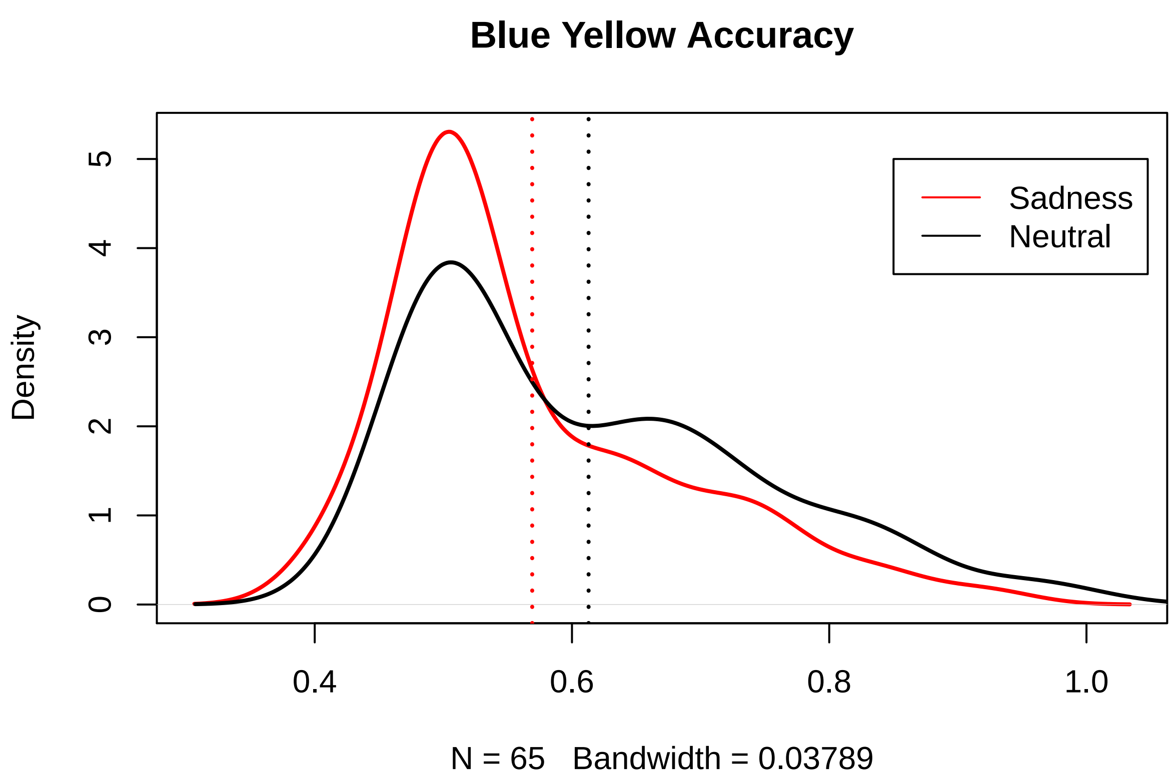
```r
emotion <- read_csv("/home/johnbjohn/Documents/git_repos/bacs-hw/hw9/study2Data.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   Subject = col_double(),
##   Emotion_Condition = col_character(),
##   RG_ACC = col_double(),
##   BY_ACC = col_double(),
##   SAD_ESRI = col_double()
## )
```

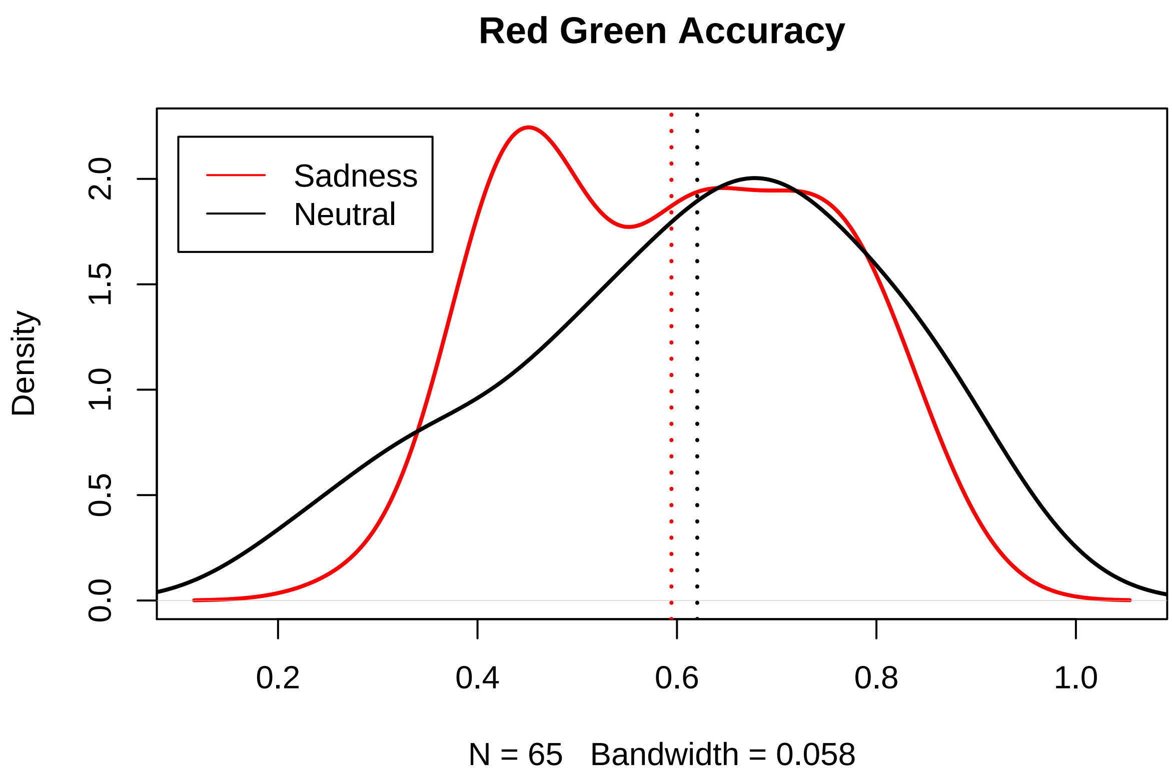
```r
sadness <- emotion[emotion$Emotion_Condition == "Sadness",]
neutral <- emotion[emotion$Emotion_Condition == "Neutral",]
```

a. Visualize the differences between blue-yellow accuracy (BY_ACC) and red-green accuracy (RG_ACC) for both the sad and neutral viewers (Emotion_Condition). You are free to choose any visualization method you wish, but only report the most useful visualizations and any first impressions.

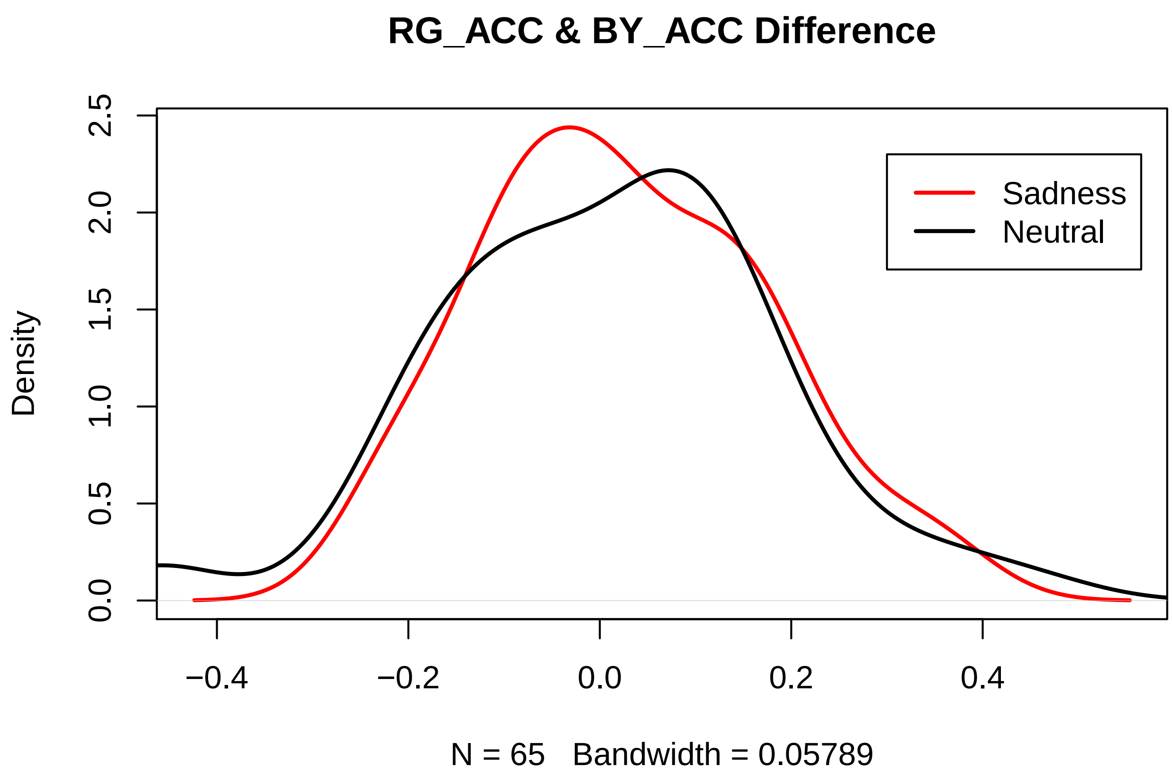
```r
plot(density(sadness$BY_ACC), main="Blue Yellow Accuracy", col="red", lwd = 2)
lines(density(neutral$BY_ACC), col="black", lwd = 2)

abline(v=mean(sadness$BY_ACC), col="red", lwd = 2, lty="dotted")
abline(v=mean(neutral$BY_ACC), col="black", lwd = 2, lty="dotted")

legend(
  0.85,
  5,
  c("Sadness", "Neutral"),
  lty = c("solid", "solid"),
  col = c("red", "black"),
)
```



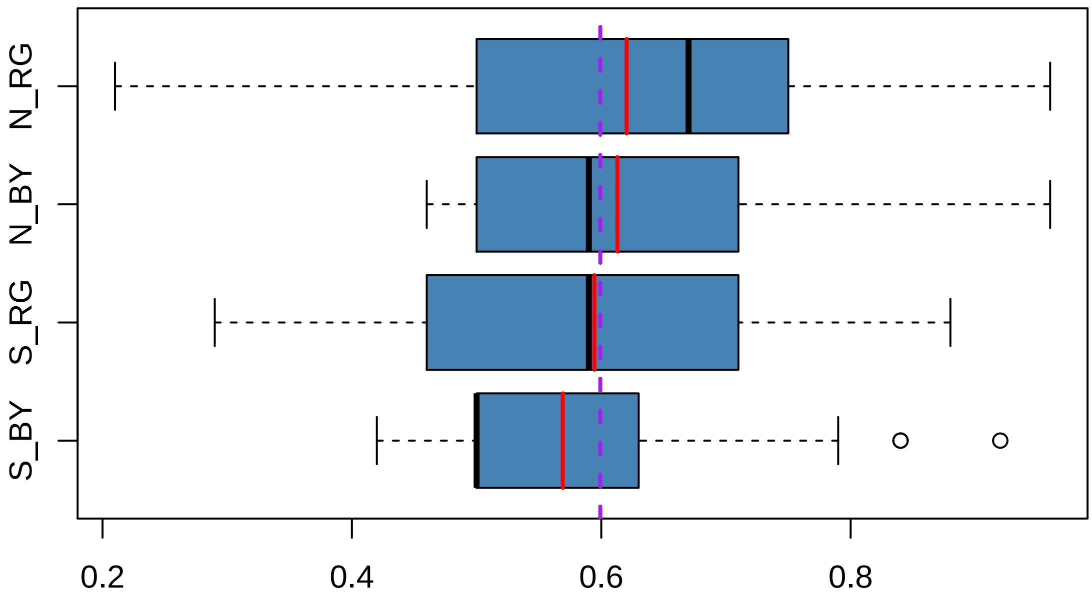

```
plot(density(sadness$RG_ACC), main="Red Green Accuracy", col="red", lwd = 2)
lines(density(neutral$RG_ACC), col="black", lwd = 2)

abline(v=mean(sadness$RG_ACC), col="red", lwd = 2, lty="dotted")
abline(v=mean(neutral$RG_ACC), col="black", lwd = 2, lty="dotted")

legend(
  0.1,
  2.2,
  c("Sadness", "Neutral"),
  lty = c("solid", "solid"),
  col = c("red", "black"),
)
```




```
sadness_diff <- sadness$RG_ACC - sadness$BY_ACC
neutral_diff <- neutral$RG_ACC - neutral$BY_ACC

plot(
  density(sadness_diff),
  lwd = 2,
  col="red",
  main = "RG_ACC & BY_ACC Difference"
)

lines(density(neutral_diff), lwd = 2, col="black")
legend(
  0.3,
  2.3,
  c("Sadness", "Neutral"),
  lty = c("solid", "solid"),
  lwd = c(2,2),
  col = c("red", "black")
)
```




```
accuracies <-
  c(sadness$BY_ACC, sadness$RG_ACC, neutral$BY_ACC, neutral$RG_ACC)

BY_RG_ACC <-
  data.frame(
    S_BY = sadness$BY_ACC,
    S_RG = sadness$RG_ACC,
    N_BY = neutral$BY_ACC,
    N_RG = neutral$RG_ACC
  )

boxplot(
  BY_RG_ACC,
  horizontal = TRUE,
  col = "steelblue"
)

y0s <- 1:4 - 0.4
y1s <- 1:4 + 0.4
means <- c(
    mean(sadness$BY_ACC),
    mean(sadness$RG_ACC),
    mean(neutral$BY_ACC),
    mean(neutral$RG_ACC)
  )

segments(x0 = means, y0 = y0s, y1 = y1s, col="red", lwd = 2)

abline(v = mean(accuracies), lwd = 2, col="purple", lty="dashed")
```



From the graph above, we can see that the difference and the mean of each treatment don't differ that much.

b. Run a t-test (traditional) to check if there is a significant difference in blue-yellow accuracy between sad and neutral participants at 95% confidence.

```
t.test(sadness$BY_ACC, neutral$BY_ACC)
```

```
##
##  Welch Two Sample t-test
##
## data:  sadness$BY_ACC and neutral$BY_ACC
## t = -2.0435, df = 125.61, p-value = 0.04309
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.086308149 -0.001384159
## sample estimates:
## mean of x mean of y
## 0.5690769 0.6129231
```

Since the p-value is less than `0.05`, we can safely reject the null hypothesis and accept the alternative hypothesis. Thus, since the p-value is less than `0.05`, we can say that the result is statistically significant.

c. Run a t-test (traditional) to check if there is a significant difference in red-green accuracy between sad and neutral participants at 95% confidence.

```
t.test(sadness$RG_ACC, neutral$RG_ACC)
```

```
##
##  Welch Two Sample t-test
##
```

```
## data:  sadness$RG_ACC and neutral$RG_ACC
## t = -0.87491, df = 121.98, p-value = 0.3833
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.08432635  0.03263405
## sample estimates:
## mean of x mean of y
## 0.5944615 0.6203077
```

Since the p-value is greater than 0.05, we can not reject the null hypothesis and, thus, the result is not statistically significant.

d. (not graded) Do the above t-tests support a claim that there is an interaction between emotion and color axis? (i.e., does people's accuracy of color perception along different color-axes depend on their emotion? Here, accuracy is an outcome variable, while color-axis and emotion are independent factors)

No answer.

e. Run a factorial design ANOVA where color perception accuracy is determined by (sad vs. neutral), color-axis (RG vs. BY), and the interaction of emotion and color-axis.

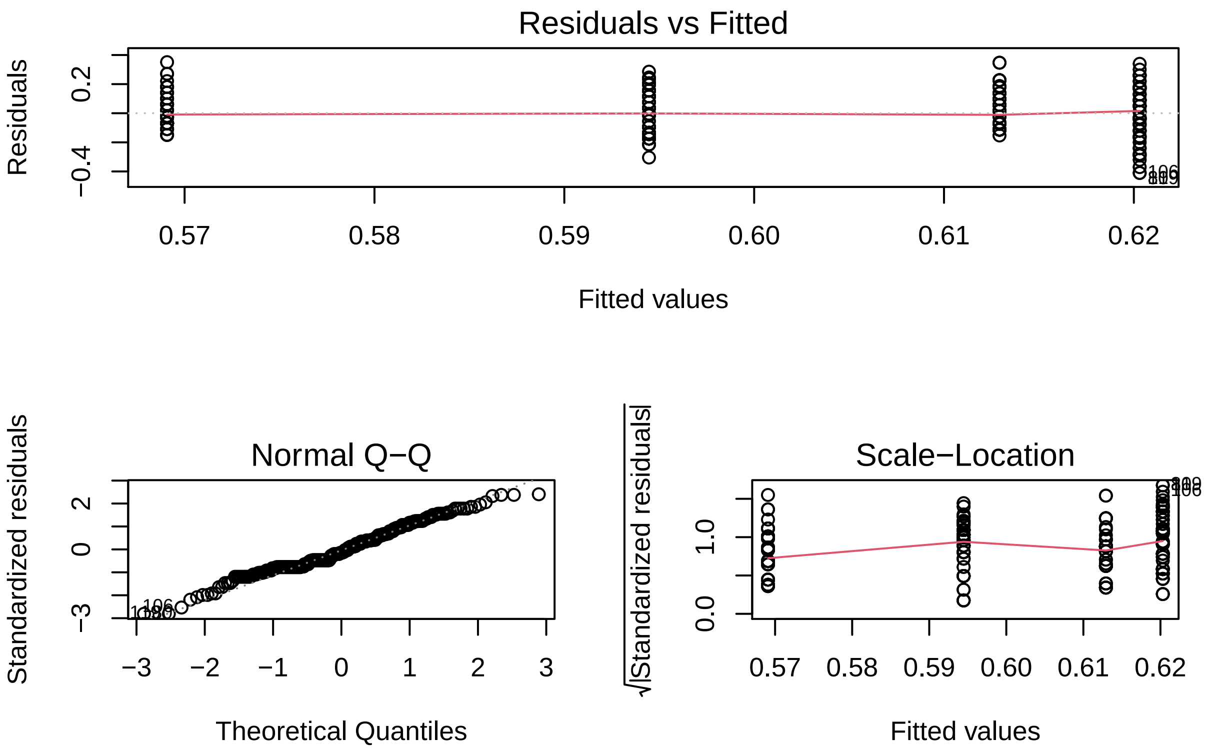
```
sadness_neutral <-
  rbind(
    sadness %>% select(Emotion_Condition, RG_ACC, BY_ACC),
    neutral %>% select(Emotion_Condition, RG_ACC, BY_ACC)
  ) %>% melt()
```

```
## Using Emotion_Condition as id variables
```

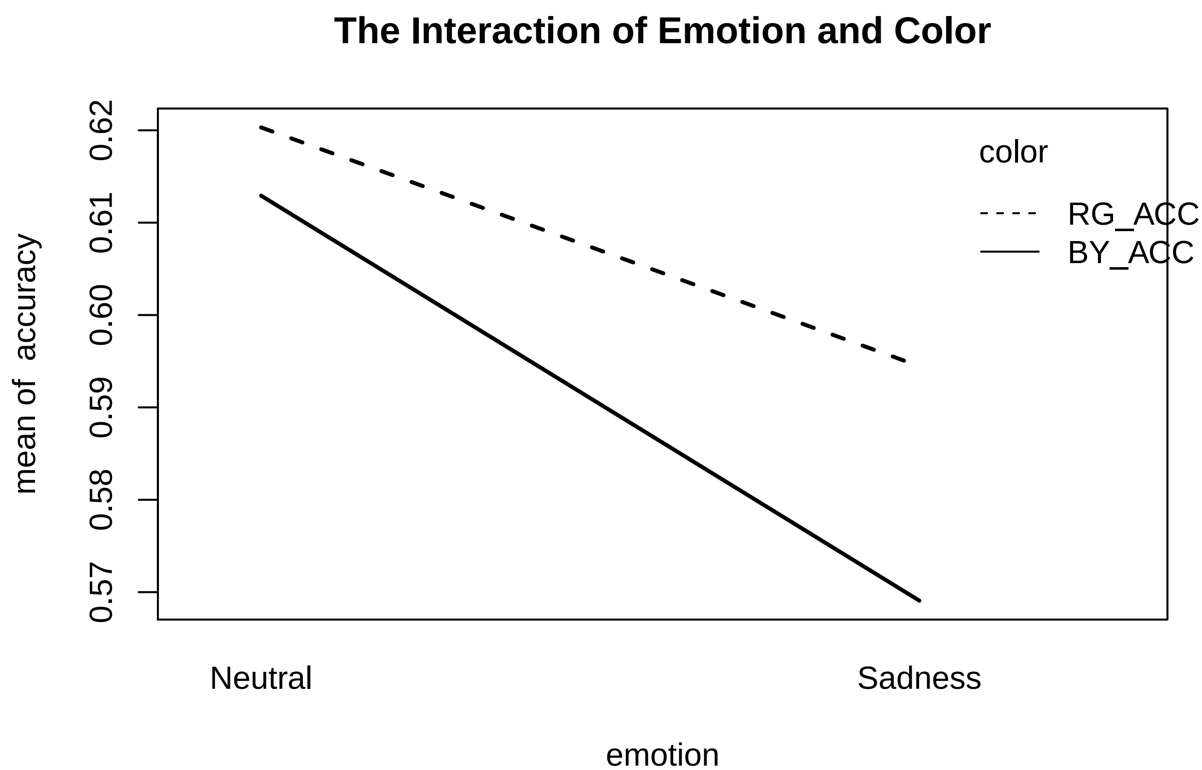
```
colnames(sadness_neutral) <- c("emotion", "color", "accuracy")

anova <-
  aov(
    formula = accuracy ~ emotion + color + emotion:color,
    data = sadness_neutral
  )

summary(anova)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)
## emotion         1  0.079 0.07893   3.644 0.0574 .
## color           1  0.017 0.01745   0.806 0.3703
## emotion:color   1  0.005 0.00527   0.243 0.6224
## Residuals     256  5.545 0.02166
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
layout(matrix(c(1,1,2,3),2,2, byrow=TRUE))
plot(anova, 1: 3)
```




```r
require(graphics)
with(
  sadness_neutral,
  interaction.plot(
    x.factor = emotion,
    trace.factor = color,
    response = accuracy,
    lwd = 2,
    main = "The Interaction of Emotion and Color",
    legend=TRUE
  )
)
```



By looking at the graph above, since they differ only by a little fraction, we can say that these three factors don't influence color perception accuracy.