Data Visualization Project on Lying and Perceived Ethicalness

Camille Proszanski

Vanessa Mpiana

Emily Skalkos

PSYC3031 A

York University

Data Visualization on Lying and Perceived Ethicalness

Humans can behave in ways that defy economic logic. Instead of acting to secure the greatest amount of resources (e.g., money, food) for themselves, they sometimes choose to help other individuals even when it results in a loss of their own resources. This tendency is referred to as altruism, a prosocial behaviour that describes the actions to help others without reward. Logically, this behaviour might not make much sense, but the evolutionary perspective suggests that prosocial behaviours were necessary to make positive group bonds required for the survival of early humans. These behaviours were passed down with increased frequency and are still present in the population today. Interviews conducted with individuals who faced incredible risks (e.g., imprisonment, death) to help save and hide Jewish people during World War II, a very altruistic act, claimed one motivation was due to strong moral beliefs (Fogelman & Wiener, 1985). Additional research has shown that having moral values as a defining feature of the self concept is associated with increased helpfulness (Aquino et al., 2009). These findings suggest a positive relationship between prosocial behaviours and morality.

The central idea of morality or ethics is the distinction of what is right and what is wrong. For example, children are often taught that lying is wrong and telling the truth is right. A study indicated that people try to avoid the immoral task of lying and prefer to select someone else to lie for them (Erat, 2013). However, lying is still a common occurrence and can sometimes be motivated by prosocial reasons (e.g. lie about liking someone’s cooking to maintain their feelings). Despite lies often being viewed as unethical, innate evolutionary pressure to engage and approve prosocial actions could cause a unique relationship to occur when viewing these concepts and their opposite (selfish honesty) together.

Both men and women engage in prosocial behaviours, but they do so in different situations. Through either past evolutionary pressure or social gender norms, women tend to engage in more communal behaviours that emphasize establishing and maintaining social bonds (Gneezy, 2005). Alternatively, men are more likely to engage in heroic acts and prosocial actions that are likely to be known about by others (Eagly & Crowley, 1986; Gneezy, 2005). These differing attitudes towards certain prosocial behaviours based on gender provide a unique avenue to explore and expand on.

**Description of Data**

The data our research team used for visualization was obtained from a study on lying and perceived ethicalness by Emma Levine and Maurice Schweitzer (2014). In the original study, “Are Liars Ethical? On the tension between benevolence and honesty,” the researchers examined moral judgements of telling prosocial lies and selfish truths. Prosocial lies are deceptions that are told for others' benefit while being detrimental to the liar and selfish truths are true statements that benefit the truth-teller at the cost of another.

In the study, the participants read about a situation where two individuals played a number game in which one was the “Sender” and the other the “Recipient.” The participants read about the Sender selecting to either tell the Recipient a prosocial lie or a selfish truth. In the prosocial lie condition, the Sender lied during the game resulting in both the Sender and Recipient receiving the same amount of money. In the selfish truth condition, the Sender would tell the truth and receive more money, but the Recipient would receive nothing. After reading about either of these scenarios, the participant was then asked to rate the Sender's ethicalness.

Initial ratings were assessed using a 7-point Likert scale, which was scored from 1 for “Not at all” to 7 for “Extremely” for a given statement. The ethical scale is the averaged value of five of these measures: whether the Sender was ethical, moral, a good person, that their decision was moral, and that their decision was ethical.

One hundred ninety-six participants were recruited for the study; however, ten failed the manipulation check and were excluded from the analysis. The final sample was 186 participants (61.29% female). Ages ranged from 18-39 years (*M* = 20.34, *SD* = 2.33).

For the visualization of the data, our team focused on three variables. These variables were gender, condition, and ethical scale. Gender is a categorical independent variable and is an indicator of the sex of the participant (female or male). Condition is a categorical independent variable that indicates what condition a particular participant is in (prosocial lie or selfish truth). The ethical scale is a continuous dependent variable and indicates the ethicalness rating assigned to the Sender after reading the short scenario.

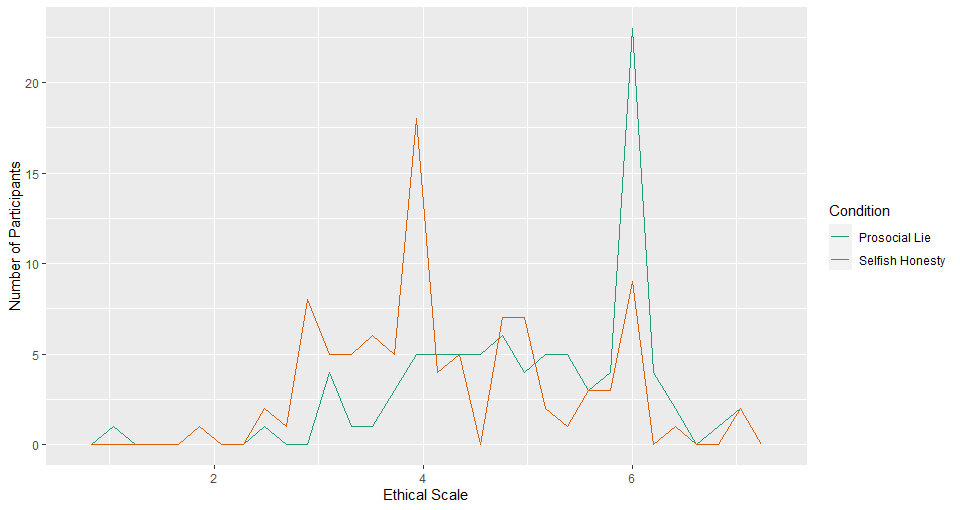
**Research Questions**

1. Does the nature of an individual's actions (prosocial lie or selfish honesty) alter the individual's ethical scale rating by a rater?
2. Is there a relationship between gender and ethical scale ratings of another individual after reading a brief description of their prosocial lie actions?
3. Does the ethical scale rating of a particular individual's actions (prosocial lie or selfish honesty) depend on the gender of the rater?

**Goals and Outcomes**

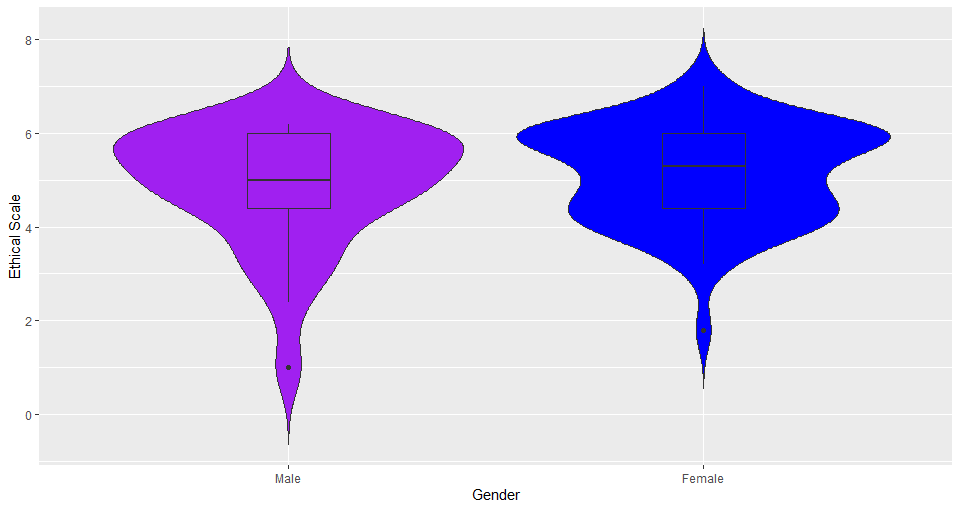
Our goal for visualization was to demonstrate how differences in participants’ rating of the Sender’s ethicalness were contingent upon the condition they were in, specifically, prosocial lie, and if this relationship was further influenced by the participants’ gender.

Participants were randomly assigned to either the prosocial lie (*n* = 91) or selfish honesty (*n* = 95) condition. Preliminary analysis revealed the participant who read about the prosocial lie rated the Sender higher on the ethical scale (*M* = 5.06, *SD* = 1.13) than those in the selfish honesty condition (*M* = 4.31, *SD* = 1.1). To visualize this data, our team chose to create a frequency polygon (See Figure 1).



*Figure 1.* Frequency polygon distribution of ethical scale rating as a function of condition (prosocial lie and selfish honesty).

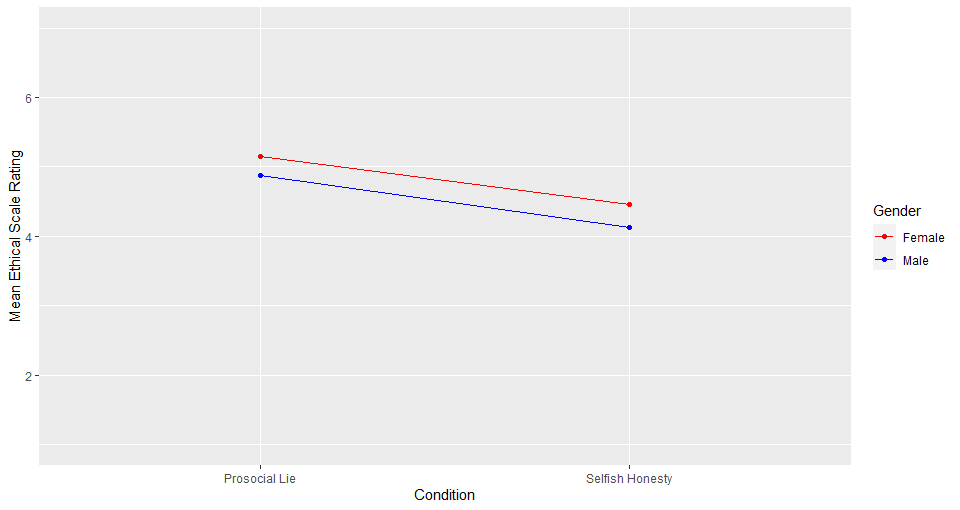
Gender differences in ethicalness ratings were further explored within the prosocial lie condition. There were 29 male and 62 female participants in this sample. Females rated the Sender’s ethicalness (*M* = 5.15, *SD* = 1.06) higher than males (*M* = 4.88, *SD* = 1.28). Only in the female condition was the highest possible rating (7) given, and only in the male condition was the lowest possible rating (1) given. To visualize this data, our team chose to create a violin plot (See Figure 2).



*Figure 2.* Distribution of ethical scale ratings as a function of gender in the prosocial lie condition.

Additional descriptive statistics were conducted to analyze for possible differences in ethical scale rating in the prosocial lie and selfish honesty groups by gender (see Table 1). Although there were mean differences between conditions and gender, there did not appear to be an interaction (See Figure 3).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 1.** The Descriptive Statistics for Each Condition by Gender | | | | | | |
|  | Prosocial lie | | | Selfish honesty | | |
|  | n | M | SD | n | M | SD |
| Male | 29 | 4.88 | 1.26 | 43 | 4.13 | 1.02 |
| Female | 62 | 5.15 | 1.06 | 52 | 4.47 | 1.14 |



*Figure 3.* Mean ethical scale ratings as a function of gender and condition.

**Limitations**

Of the 186 participants, there were more females (*n* = 114) in comparison to males

(*n* = 72). This gender disparity could pose a limitation to our data, especially when exploring gender differences. There might not be enough scores collected from male participants to represent the population adequately. In addition, there were more outliers found in the male group than in the female group, which had a greater impact on the mean due to the smaller representation of males within the sample. Overall, the participants' age range was very narrow, with 50% of participants being between age 19 to 21 years old. Having a narrow age range could present issues associated with generalizing findings to older populations. Lastly, our team had limited knowledge of the methods and practices used to collect the data. As such, it is assumed that data collection was performed in an unbiased way; however, it is not certain if bias impacted the collection of data.

References

Aquino, K., Freeman, D., Reed II, A., Lim, V.K., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: The interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology*, *97*(1), 123–141. https://doi.org/10.1037/a0015406

Eagly, A.H., & Crowley, M. (1986). Gender and helping behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, *100*(3), 283–308. https://doi.org/10.1037/0033-2909.100.3.283

Erat, S. (2013). Avoiding lying: The case of delegated deception. *Journal of Economic Behavior & Organization*, *93*, 273–278. https://doi.org/10.1016/j.jebo.2013.03.035

Fogelman, E., & Wiener, V.L. (1985). The few, the brave, the noble. *Psychology Today*, *19*(8), 60–65.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394. https://doi.org/10.1257/0002828053828662

Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? on the tension between benevolence and honesty. *Journal of Experimental Social Psychology, 53*, 107-117. doi:http://dx.doi.org.ezproxy.library.yorku.ca/10.1016/j.jesp.2014.03.005

Code

#Data Visualization Project

#There are no routine statistical questions, only questionable statistical routines- Sir David Cox

# Load packages to access necessary functions (e.g. ggplot(), mutate(), describe(), read\_excel()) to run analysis on the data

library(tidyverse)

library(psych)

library(readxl)

# Create pathway to data open data in R with read\_excel() function and assign data frame name Ethics\_of\_Lying\_Data via <-

Ethics\_of\_Lying\_Data <- read\_excel("data/Study 1 Ethics of Lying Data\_FINAL.xlsx")

# The View() function is used to open a data frame in another tab. This allows data to be checked to see if it was imported correctly and allows easy access to refer back to the data frame

View(Ethics\_of\_Lying\_Data)

# This code changes the name of column Male 1 to Gender so it has a more meaningful name

colnames(Ethics\_of\_Lying\_Data)[colnames(Ethics\_of\_Lying\_Data) == "Male1"] <- "Gender"

#Transform categorical data (gender and condition) into a factor in R with as.factor () and rename variable levels with meaningful names with fct\_recode(), all changes will transfer over into the data frame through code “Ethics\_of\_Lying\_Data <-” instead of just being an output

Ethics\_of\_Lying\_Data <- mutate(Ethics\_of\_Lying\_Data,

Gender = as.factor(Gender),

Gender = fct\_recode(Gender,

"Male"= "1",

"Female"= "2"),

Condition = as.factor(Condition),

Condition = fct\_recode(Condition,

"Prosocial Lie" = "ProsocialLie",

"Selfish Honesty" = "SelfishHonesty" ))

#This code indicates the variable type (e.g. numeric, factor, character) to make sure the variables are recognized properly in R

sapply(colnames(Ethics\_of\_Lying\_Data), function(x) class(Ethics\_of\_Lying\_Data[[x]]))

# The filter() function is used to remove participants that failed one or more manipulations checks (out of two) by selecting all the rows that answered 2 in the “NumberofMchecks\_correct” column and thus removing all the rows that did not answer 2 (e.g. 0 or 1), all changes will transfer over into the data frame through code “Ethics\_of\_Lying\_Data <-”

Ethics\_of\_Lying\_Data <- filter(Ethics\_of\_Lying\_Data,

NumberofMchecks\_correct == 2)

# Check for distinct ids with distinct () function, length of tibble output should align with length of data frame indicating distinct ideas for each participant --------------

Ethics\_of\_Lying\_Data%>%

distinct(ResponseID)

#Check Data for problems in the data --------------------------------

#Check for parsing problems with problems() function

problems(Ethics\_of\_Lying\_Data)

# Check for variable specifications with spec() function

spec(Ethics\_of\_Lying\_Data)

# ------------------------Check for ceiling effects greater than 15% which would indicate that questions used in the study had an upper limit that was too low to show difference in means between groups-----------------

# The filter() function is use to select all the scores the that had a 7 (the highest score) and nrow() function was used to find out the count of the selected scores and is then assigned an object name

Number\_Scores\_At\_Max <-nrow(filter(Ethics\_of\_Lying\_Data, EthicalScale == "7"))

#The count of total number of observations is then computed with nrows() and assigned an object name

Number\_Of\_Observations <- nrow(Ethics\_of\_Lying\_Data)

# The percentage is then calculated by taking the count of the number of ethical scale ratings that were 7 and dividing it by the total number of scores taken

Number\_Scores\_At\_Max/Number\_Of\_Observations\*100

#General descriptives of the data--------------------------------------------------

#The select() function is used to select variables of interest (condition, gender, age and ethical scale) and then with summary() function to get the count of categorical variables, and mean and quartile range of continuous variables

select(Ethics\_of\_Lying\_Data, Condition, Gender, Age, EthicalScale)%>% summary()

#The variable age is subset from the data frame with this code “Ethics\_of\_Lying\_Data$Age”. The describe() function is then used to get mean, standard deviation, max and min scores

Ethics\_of\_Lying\_Data$Age%>%describe()

#Descriptive statistics of relationship between condition and ethical scale -------------------------------------------

# The describeBy () function output descriptive statistics (e.g. mean, standard deviation, max, min, skewness) broken down by condition (prosocial lie and selfish honesty)

describeBy(x = Ethics\_of\_Lying\_Data$EthicalScale, group = Ethics\_of\_Lying\_Data$Condition)

#------------Visualization plots used to help see distribution of scores and to get a feel for the data-----------

# Creating a Histogram-------------------------------------------

#ggplot() function is used to make a plot, geom\_histogram() function is used to make histogram, facet\_wrap function is used to break graphs down by condition

ggplot(data = Ethics\_of\_Lying\_Data, aes(x = EthicalScale)) +

geom\_histogram(color = "gold", fill = "black") +

facet\_wrap(~ Condition)

# Creating a violin plot with a boxplot inside---------------------------------------------------

# ggplot() function is used to make a plot, geom\_violin() function is used to make violin plot, geom\_boxplot(width = 0.2) is used to make a boxplot that is small enough to fit in the violin plot, labs() function is used to add labels to the graph, theme\_dark() at a dark background to the graph (mostly just messing around with r code to see how cool a graph can look)

ggplot(Ethics\_of\_Lying\_Data, aes(x = Condition, y = EthicalScale, fill = Condition)) +

geom\_violin(trim = FALSE) +

geom\_boxplot(width = 0.2) +

labs( x="Condition", y="Ethical Scale") +

theme\_dark()

#Creating a box plot-------------------------------------------

# ggplot() function is used to make a plot, geom\_boxplot() function is used to make a boxplot

ggplot(Ethics\_of\_Lying\_Data, aes(x = Condition, y = EthicalScale)) +

geom\_boxplot()

#Creating a qq plot -------------------------------------------

#ggplot() function is used to make a plot, geom\_qq() function is used to make qq plot, facet\_wrap function is used to break graphs down by condition

ggplot(Ethics\_of\_Lying\_Data, aes(sample = EthicalScale)) +

geom\_qq() +

geom\_qq\_line() +

facet\_grid(~Condition)

# Creating frequency polygon -------------------------------

# ggplot() function is used to make a plot, geom\_freqploy() function is used to make frequency polygon, scale\_color\_brewer(palette="Dark2") code is used to colour graph lines orange and teal, the favourite colours of a team member, by using the palette “Dark2” ,labs() function is used to add labels to the graph

ggplot(Ethics\_of\_Lying\_Data, aes(x = EthicalScale))+

geom\_freqpoly(aes(color = Condition))+

scale\_color\_brewer(palette="Dark2")+

labs( y="Number of Participants", x="Ethical Scale")

#--- Descriptive statistics of relationship between gender and ethical scale in the prosocial lie condition --

# The filter() function is used to select all data rows that had "Prosocial Lie" in the condition column indicating that they were in that condition, a new data frame was created via “Ethics\_of\_Lying\_Data\_Prosocial\_Lie <-” code

Ethics\_of\_Lying\_Data\_Prosocial\_Lie <-filter(Ethics\_of\_Lying\_Data, Condition == "Prosocial Lie")

#The describeBy () function output descriptive statistics (e.g. mean, standard deviation, max, min, skewness) broken down by gender (male and female), from the data set “Ethics\_of\_Lying\_Data\_Prosocial\_Lie” so all scores are from the prosocial condition

describeBy(x = Ethics\_of\_Lying\_Data\_Prosocial\_Lie$EthicalScale, group = Ethics\_of\_Lying\_Data\_Prosocial\_Lie$Gender)

#------------Visualization plots used to help see distribution of score and to get a feel for the data------------

# The data used for all graphs in this section is “Ethics\_of\_Lying\_Data\_Prosocial\_Lie” as gender differences are only assessed in the prosocial condition

# Create a Histogram----------------------------------

#ggplot() function is used to make a plot, geom\_histogram() function is used to make histogram, facet\_wrap function is used to break graphs down by gender

ggplot(data = Ethics\_of\_Lying\_Data\_Prosocial\_Lie, aes(x = EthicalScale)) +

geom\_histogram(color = "blue") +

facet\_wrap(~ Gender)

# Creating a violin plot---------------------------------

#ggplot() function is used to make a plot, geom\_violin() function is used to make violin plot, geom\_boxplot(width = 0.2) is used to make boxplot that is small enough to fit in the violin plot, labs() function is used to add labels to the graph, scale\_fill\_manual(values=c("purple", "blue")) code is used to fill the violin plots with the favorite colours of a group member, theme(legend.position = "none") code was used to remove the legend as it was unnecessary and took up space

ggplot(Ethics\_of\_Lying\_Data\_Prosocial\_Lie, aes(x = Gender, y = EthicalScale, fill = Gender)) +

geom\_violin(trim = FALSE) +

geom\_boxplot(width = 0.2) +

labs (x="Gender", y="Ethical Scale") +

scale\_fill\_manual(values=c("purple", "blue"))+

theme(legend.position = "none")

#Creating a box plot-------------------------------------------

# ggplot() function is used to make a plot, geom\_boxplot() function is used to make a boxplot

ggplot(Ethics\_of\_Lying\_Data\_Prosocial\_Lie, aes(x = Gender, y = EthicalScale)) +

geom\_boxplot()

#Creating qq plot---------------------------------------------

#ggplot() function is used to make a plot, geom\_qq() function is used to make qq plot, facet\_wrap function is used to break graphs down by Gender

ggplot(Ethics\_of\_Lying\_Data\_Prosocial\_Lie, aes(sample = EthicalScale)) +

geom\_qq() +geom\_qq\_line() +

facet\_grid(~Gender)

# -Descriptive statistics of relationship between both condition and gender on ethical scale (interaction)-

#The describeBy () function output descriptive statistics (e.g. mean, standard deviation, max, min, skewness) broken down by gender (male and female) and condition (prosocial lie and selfish honesty), from the data set

describeBy(x = Ethics\_of\_Lying\_Data$EthicalScale, group = Ethics\_of\_Lying\_Data$Condition : Ethics\_of\_Lying\_Data$Gender)

#------------Visualization plots used to help see distribution of scores and to get a feel for the data------------

#Creating a box plot-------------------------------------------

# ggplot() function is used to make a plot, geom\_boxplot() function is used to make a boxplot, facet\_wrap function is used to break graphs down by gender

ggplot(Ethics\_of\_Lying\_Data, aes(y = EthicalScale,

x = Condition)) +

geom\_boxplot()+

facet\_wrap(~Gender)

#Creating a qq plot -------------------------------------------

#ggplot() function is used to make a plot, geom\_qq() function is used to make qq plot, facet\_wrap function is used to break graphs down by condition and gender

ggplot(Ethics\_of\_Lying\_Data, aes(sample = EthicalScale))+

geom\_qq() +

geom\_qq\_line() +

facet\_grid(Condition~Gender)

#Creating interaction plot-------------------------------------------------------------------------------------------------------

#-------------Compute group means for interaction plot

# Group the data by gender and condition with group\_by()function and save it to new object “Ethics\_of\_Lying\_Data\_Grouped\_By\_Condition\_and\_Gender” as not to get confused or have to ungroup data later

Ethics\_of\_Lying\_Data\_Grouped\_By\_Condition\_and\_Gender<-group\_by(Ethics\_of\_Lying\_Data,

Condition,Gender)

# The summarise () function is used to create a new data frame of the means of the ethical scale, grouped by gender and condition from pervious code

Ethics\_of\_Lying\_Summary <- summarise(Ethics\_of\_Lying\_Data\_Grouped\_By\_Condition\_and\_Gender,

mean = mean (EthicalScale))

# ggplot() function is used to make a plot, geom\_point() function is used to make a point graph, geom\_line() is used to connect the points, labs() function is used to add labels to the graph, scale\_color\_manual(breaks = c("Female", "Male"), values=c("red", "blue")) code is used to reorganize the order of the legend so female occurs first in the legend and to designate each line with a colour, ylim(c(1,7)) codes so that the graph does not zoom and make the data look more significant than it is, it causes the y-axis limit to range from 1 to 7 (the range of the variable)

ggplot(Ethics\_of\_Lying\_Summary, aes( x = Condition,

y = mean,

color = Gender,

group = Gender))+

geom\_point() +

geom\_line()+ ylim(c(1,7))+

labs(y = "Mean Ethical Scale Rating")+

scale\_color\_manual(breaks = c("Female", "Male"),

values=c("red", "blue"))

#----------------------------------------------------------------------------------------------------------------------------------

# This section below was the original code we used for looking at gender; however, that data was bimodal and we realised that because the gender was taken from individuals in both conditions there was a third variable impacting analysis. As such we decided to explore gender differences in one condition(see code above). We decided to include the code to express the process we went through

#---------------------------------------------------------------------------------------------------------------------------------

#The describeBy () function output descriptive statistics (e.g. mean, standard deviation, max, min, skewness) broken down by gender (male and female)

describeBy(x = Ethics\_of\_Lying\_Data$EthicalScale, group = Ethics\_of\_Lying\_Data$Gender)

# Create a Histogram----------------------------------

#ggplot() function is used to make a plot, geom\_histogram() function is used to make histogram, facet\_wrap function is used to break graphs down by gender

ggplot(data = Ethics\_of\_Lying\_Data, aes(x = EthicalScale)) +

geom\_histogram(color = "blue") +

facet\_wrap(~ Gender) # divide the histogram, based on Gender

# Creating a violin plot---------------------------------

#ggplot() function is used to make a plot, geom\_violin() function is used to make violin plot, geom\_boxplot(width = 0.2) is used to make boxplot that is small enough to fit in the violin plot, labs() function is used to add labels to the graph

ggplot(Ethics\_of\_Lying\_Data, aes(x = Gender, y = EthicalScale, fill = Gender)) +

geom\_violin(trim = FALSE) +

geom\_boxplot(width = 0.2) +

labs(title="GGPLOT VIOLIN PLOT", x="Gender", y="Ethical Scale")

#Creating box plot-------------------------------------------------

# ggplot() function is used to make a plot, geom\_boxplot() function is used to make a boxplot, facet\_wrap function is used to break graphs down by gender

ggplot(Ethics\_of\_Lying\_Data, aes(x = Gender, y = EthicalScale)) +

geom\_boxplot()

#Creating qq plot---------------------------------------------

#ggplot() function is used to make a plot, geom\_qq() function is used to make qq plot, facet\_wrap function is used to break graphs down by gender

ggplot(Ethics\_of\_Lying\_Data, aes(sample = EthicalScale)) +

geom\_qq() +

geom\_qq\_line() +

facet\_grid(~Gender)