Survey Two - Analysis

Hari VS

CLEAN DATA OBTAINED FROM CLEANING.RMD

THIS FILE CONTAINS ONLY THE CODE FOR ANALYSIS

Breakdown of analysis

1. Each person is an observation - Testing accuracy

```
1a. Overview of the variables through plots
1b. Paired t-tests on *NO_AI vs AI* conditions
1c. Two sample t-tests on *NO_BAR vs BAR* conditions (uncertainty)
1d. Simple linear regressions
1e. Mulitple linear regressions
```

2. Each person is an observation - Testing confidence

```
same procedure followed
```

Importing datasets

```
# 1. person - both plants and animals included. All stimulus
#Dataset - averages - each observation is a person
person <- read_csv("Datasets/Data_cleaned_person.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   X1 = col_double(),
##   ResponseId = col_character(),
##   age = col_double(),
##   college = col_double(),
##   male_num = col_double(),
##   AI_trust_num = col_double(),
##   Task_diff_num = col_double(),
##   Dmn_know_a_num = col_double(),
##   Dmn_know_p_num = col_double(),
```

```
##   atn_ch = col_double(),
##   accuracy = col_double(),
##   confidence = col_double(),
##   time_taken = col_double(),
##   AI_use = col_double(),
##   AI = col_double(),
##   bar = col_double()
## )
```

```
#Converting the dataframe to a tibble
person <- as_tibble(person)


# 2. plants_person - only plant stimulus.
#Dataset - averages - each observation is a person
plants_person <- read_csv("Datasets/Data_cleaned_plants_person.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   X1 = col_double(),
##   ResponseId = col_character(),
##   age = col_double(),
##   college = col_double(),
##   male_num = col_double(),
##   AI_trust_num = col_double(),
##   Task_diff_num = col_double(),
##   Dmn_know_a_num = col_double(),
##   Dmn_know_p_num = col_double(),
##   atn_ch = col_double(),
##   accuracy = col_double(),
##   confidence = col_double(),
##   time_taken = col_double(),
##   AI_use = col_double(),
##   AI = col_double(),
##   bar = col_double()
## )
```

```
#Converting the dataframe to a tibble
plants_person <- as_tibble(plants_person)


# 3. animals_person - only animal stimulus
#Dataset - averages - each observation is a person
animals_person <- read_csv("Datasets/Data_cleaned_animals_person.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   X1 = col_double(),
##   ResponseId = col_character(),
##   age = col_double(),
##   college = col_double(),
##   male_num = col_double(),
##   AI_trust_num = col_double(),
##   Task_diff_num = col_double(),
```

```
##   Dmn_know_a_num = col_double(),
##   Dmn_know_p_num = col_double(),
##   atn_ch = col_double(),
##   accuracy = col_double(),
##   confidence = col_double(),
##   time_taken = col_double(),
##   AI_use = col_double(),
##   AI = col_double(),
##   bar = col_double()
## )
```

```
#Converting the dataframe to a tibble
animals_person <- as_tibble(animals_person)
```

Overview Statistics - Mean, Median, Mode, and Interquantile Ranges

```
#Summary Statistic for the overall study
summary(person)
```

```
##        X1          ResponseId             age           college
##  Min.   :  1.0   Length:402         Min.   :18.00   Min.   :0.0000
##  1st Qu.:101.2   Class :character   1st Qu.:23.00   1st Qu.:0.0000
##  Median :201.5   Mode  :character   Median :29.00   Median :0.0000
##  Mean   :201.5                      Mean   :31.37   Mean   :0.4826
##  3rd Qu.:301.8                      3rd Qu.:37.00   3rd Qu.:1.0000
##  Max.   :402.0                      Max.   :64.00   Max.   :1.0000
##
##     male_num       AI_trust_num  Task_diff_num  Dmn_know_a_num
##  Min.   :0.0000   Min.   :1.00   Min.   :1.00   Min.   :0.0000
##  1st Qu.:0.0000   1st Qu.:3.00   1st Qu.:2.00   1st Qu.:0.4000
##  Median :0.0000   Median :4.00   Median :3.50   Median :0.4000
##  Mean   :0.4876   Mean   :3.48   Mean   :3.17   Mean   :0.4806
##  3rd Qu.:1.0000   3rd Qu.:4.00   3rd Qu.:4.00   3rd Qu.:0.6000
##  Max.   :1.0000   Max.   :5.00   Max.   :5.00   Max.   :1.0000
##                   NA's   :2      NA's   :2
##  Dmn_know_p_num      atn_ch          accuracy        confidence
##  Min.   :0.000   Min.   :0.0000   Min.   :0.1562   Min.   :0.08125
##  1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:0.3438   1st Qu.:0.36875
##  Median :0.200   Median :1.0000   Median :0.4375   Median :0.50000
##  Mean   :0.191   Mean   :0.7413   Mean   :0.4502   Mean   :0.50044
##  3rd Qu.:0.400   3rd Qu.:1.0000   3rd Qu.:0.5625   3rd Qu.:0.62344
##  Max.   :1.000   Max.   :1.0000   Max.   :0.6875   Max.   :0.91250
##
##    time_taken          AI_use             AI           bar
##  Min.   :  3.901   Min.   :0.0000   Min.   :0.0   Min.   :0.0000
##  1st Qu.:  7.887   1st Qu.:0.4625   1st Qu.:0.0   1st Qu.:0.0000
##  Median : 10.419   Median :0.5563   Median :0.5   Median :0.0000
##  Mean   : 13.159   Mean   :0.5609   Mean   :0.5   Mean   :0.2537
##  3rd Qu.: 14.305   3rd Qu.:0.6687   3rd Qu.:1.0   3rd Qu.:1.0000
##  Max.   :258.516   Max.   :0.9750   Max.   :1.0   Max.   :1.0000
##                    NA's   :201
```

```
#creating a new column - overconfidence, which will also be used in the models
#as a response variable.
```

```r
person$over_conf <- person$confidence - person$accuracy

animals_person$over_conf <- animals_person$confidence - animals_person$accuracy

plants_person$over_conf <- animals_person$confidence - animals_person$accuracy
```

```r
#Subsetting to multiple tibbles for ease of use
#contains only reponses measured for no-AI condition
person_noAI <- subset(person, AI == 0)

#Responses measured for all AI-conditions
person_AI <- subset(person, AI == 1)

#responses of participants who were not provided uncertainty information
person_nobar <- subset(person, AI == 1) %>%
  filter(bar == 0)

#responses of participants who received uncertainty information
person_bar <- subset(person, AI ==1) %>%
  filter(bar == 1)
```

Univariate Plots

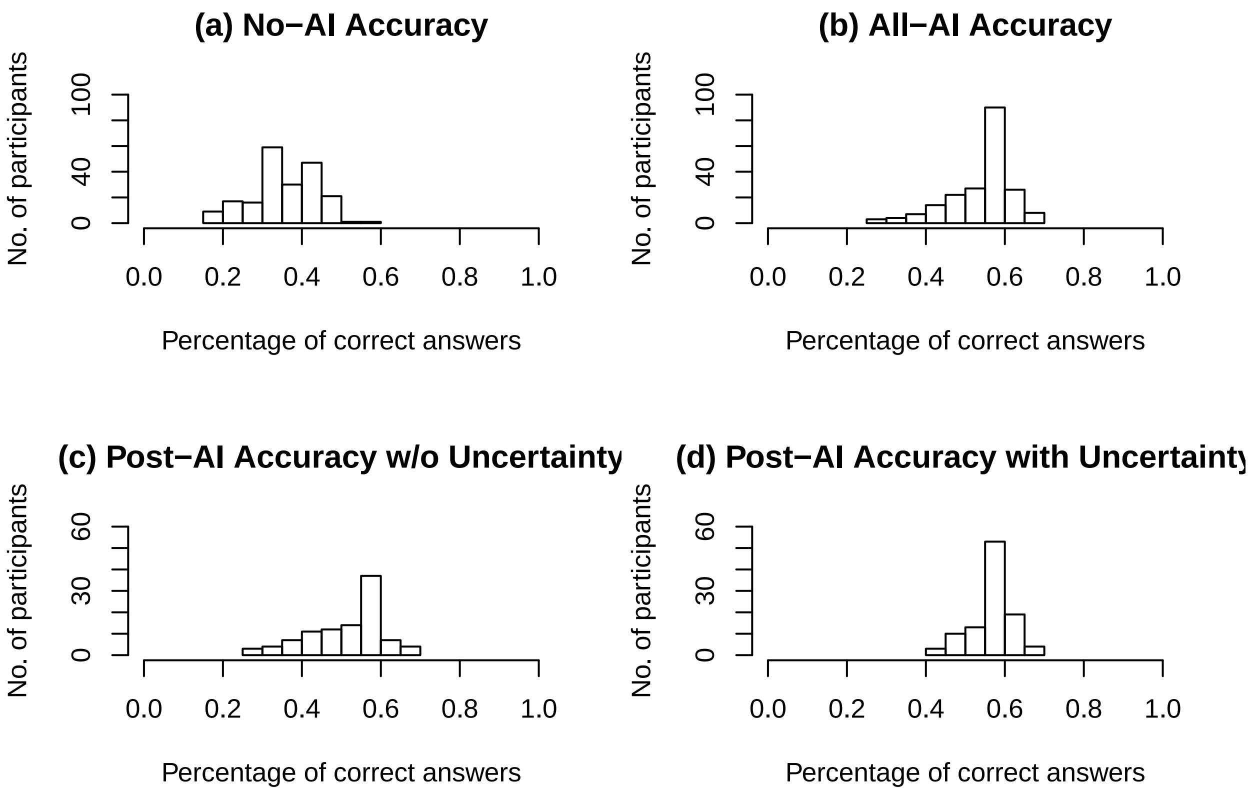
Accuracy Plots

The average accuracy of the participants was 0.36 (SD = 0.08), when no AI recommendations were provided. In comparison, when AI recommendations were provided, the average accuracy was 0.54(SD = 0.08). The means and the plots clearly indicate the positive relationship between AI recommendations and accuracy of the participants.

The average accuracy of the participants was 0.51 (SD = 0.09), when uncertainty information was not provided. In comparison, when the uncertainty information were provided, the average accuracy was 0.57 (SD = 0.05). The means and the plots show that accuracy of the participants increase slightly when uncertainty information is provided.

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

#accuracy percentage for baseline - No AI condition
hist(person_noAI$accuracy, #choosing column in a dataset
     main = "(a) No-AI Accuracy", #main plot label
     xlab = "Percentage of correct answers", #x-axis label
     ylab = "No. of participants", #y-axis label
     ylim = c(0,100), xlim = c(0,1)) #limits for x- & y-axis in the plot

#accuracy percentage in all AI Condition
hist(person_AI$accuracy, main = "(b) All-AI Accuracy",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,100), xlim = c(0,1))

#accuracy percentage in AI Condition without bars
hist(person_nobar$accuracy,
     main = "(c) Post-AI Accuracy w/o Uncertainty",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,60), xlim = c(0,1))
```

```r
#accuracy percentage in AI with Uncertianty Information (AI_bars)
hist(person_bar$accuracy, main ="(d) Post-AI Accuracy with Uncertainty",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,60), xlim = c(0,1))
```

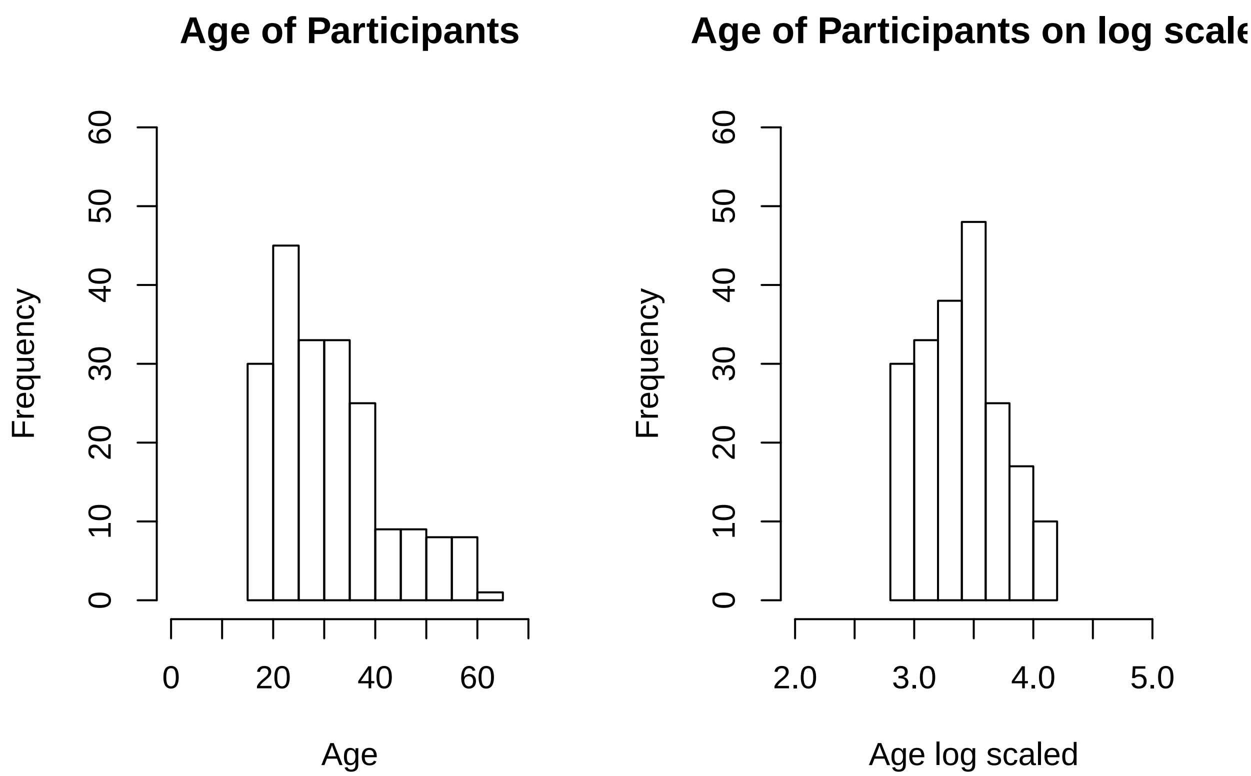



```r
par(mfrow=c(1,1))
```

Age

The average age of the participants was 31.37 years with SD = 10.66. The youngest person to take the survey was 18 years old. The oldest person to take the survey was 64 years old. The age of the participants was not distributed evenly as it can be seen in the figure below (left). Hence, age is log scaled after which the distribution is relatively better. On all linear regression models, the age will be introduced as a variable in a log scale.

```r
#Spread of AGE
par(mfrow=c(1,2))

hist(person_noAI$age, #choosing column in a dataset to plot
     main = "Age of Participants", #main plot title
     xlab = "Age", #x-axis title
     ylim = c(0,60), xlim = c(0,70)) #both axes limits.

hist(log(person_noAI$age), main = "Age of Participants on log scale",
     xlab = "Age log scaled", ylim = c(0,60), xlim = c(2,5))
```



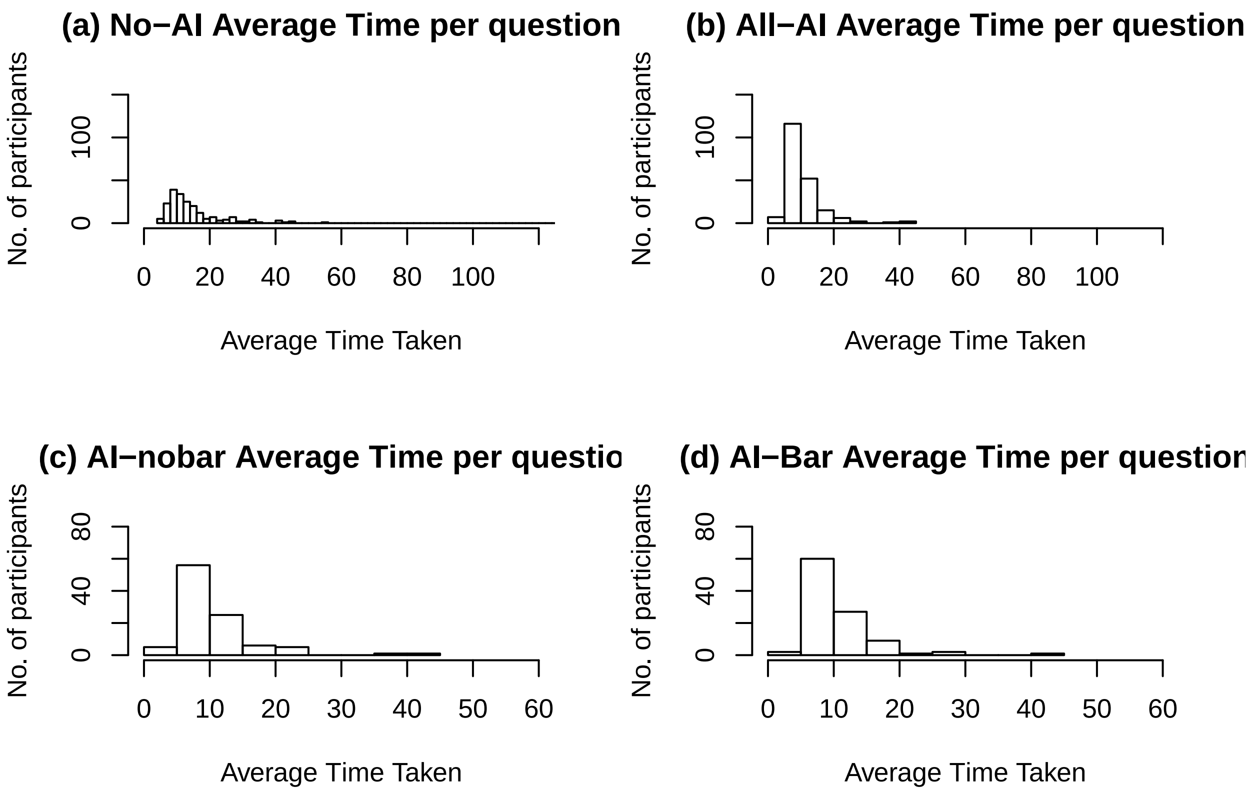
Time taken

The average time spent per question by the participants was 15.88 (SD = 19.23), when no AI recommendations were provided. In comparison, when AI recommendations were provided, the average time spent per question was 10.44 (SD = 5.68). The participants took less time to identify the stimulus when AI recommendations were provided. The standard deviation also reduces significantly when AI recommendations are provided.

The average time spent per question by the participants was 10.45 (SD = 6.02), when uncertainty information was not provided. In comparison, when the uncertainty information were provided, the average time spent per question was 10.43 (SD = 5.36). SImilar to confidence ratings, the effect of uncertainty information on time taken is unclear.

```
par(mfrow=c(2,2))

#Average Time spent per question for baseline - No AI condition
hist(person_noAI$time_taken,main="(a) No-AI Average Time per question",
     xlab = "Average Time Taken", ylab = "No. of participants",
     ylim = c(0,150), xlim = c(0,120), breaks = "FD")

#Average Time spent per question in all AI Condition
hist(person_AI$time_taken, main ="(b) All-AI Average Time per question",
     xlab = "Average Time Taken", ylab = "No. of participants",
     ylim = c(0,150), xlim = c(0,120))

#Average Time spent per question in AI Condition without bars
hist(person_nobar$time_taken,
     main = "(c) AI-nobar Average Time per question",
     xlab = "Average Time Taken", ylab = "No. of participants",
     ylim = c(0,80), xlim = c(0,60))
```

```r
#Average Time spent per question in AI with Uncertianty Information (AI_bars)
hist(person_bar$time_taken, main ="(d) AI-Bar Average Time per question",
     xlab = "Average Time Taken", ylab = "No. of participants",
     ylim = c(0,80), xlim = c(0,60))
```




```r
par(mfrow=c(1,1))
```

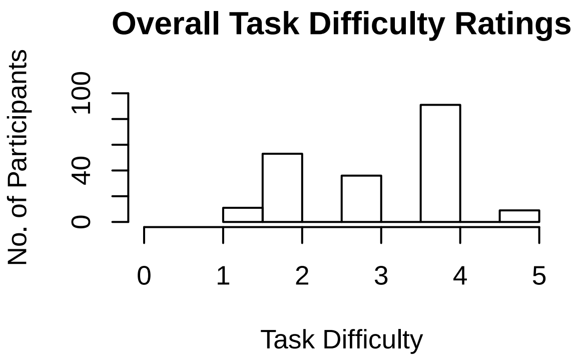
Task Difficulty

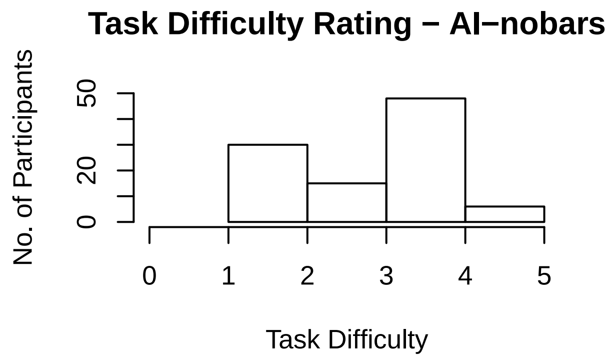
The overall average task difficulty rating in the experiment was 3.17 (SD = 1.05). The average task difficulty rating for participants that did not receive uncertainty information was 3.24 (SD = 1.08) compared to participants who received the uncertainty information; mean = 3.1 (SD = 1.01).

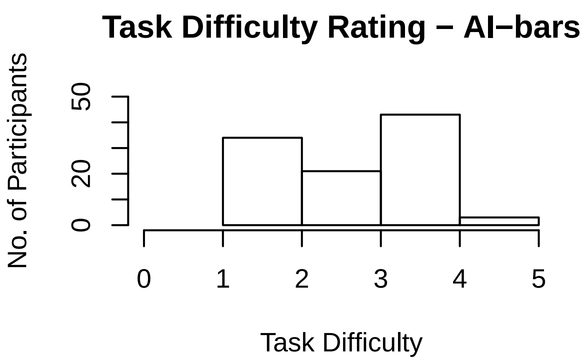
Perceived task difficulty ratings are not normaly distributed. Log transformations will not help either due to irrational spread of the responses.

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

#plotting task difficulty AI vs No-AI
hist(person_noAI$Task_diff_num,
     main = "Overall Task Difficulty Ratings",
     xlab = "Task Difficulty", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,100), xlim = c(0,5))

#plotting task difficulty AI-nobars
hist(person_nobar$Task_diff_num,
     main = "Task Difficulty Rating - AI-nobars",
     xlab = "Task Difficulty", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,50), xlim = c(0,5))
```

```r
#plotting task difficulty AI-bars
hist(person_bar$Task_diff_num,
     main = "Task Difficulty Rating - AI-bars",
     xlab = "Task Difficulty", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,50), xlim = c(0,5))
```





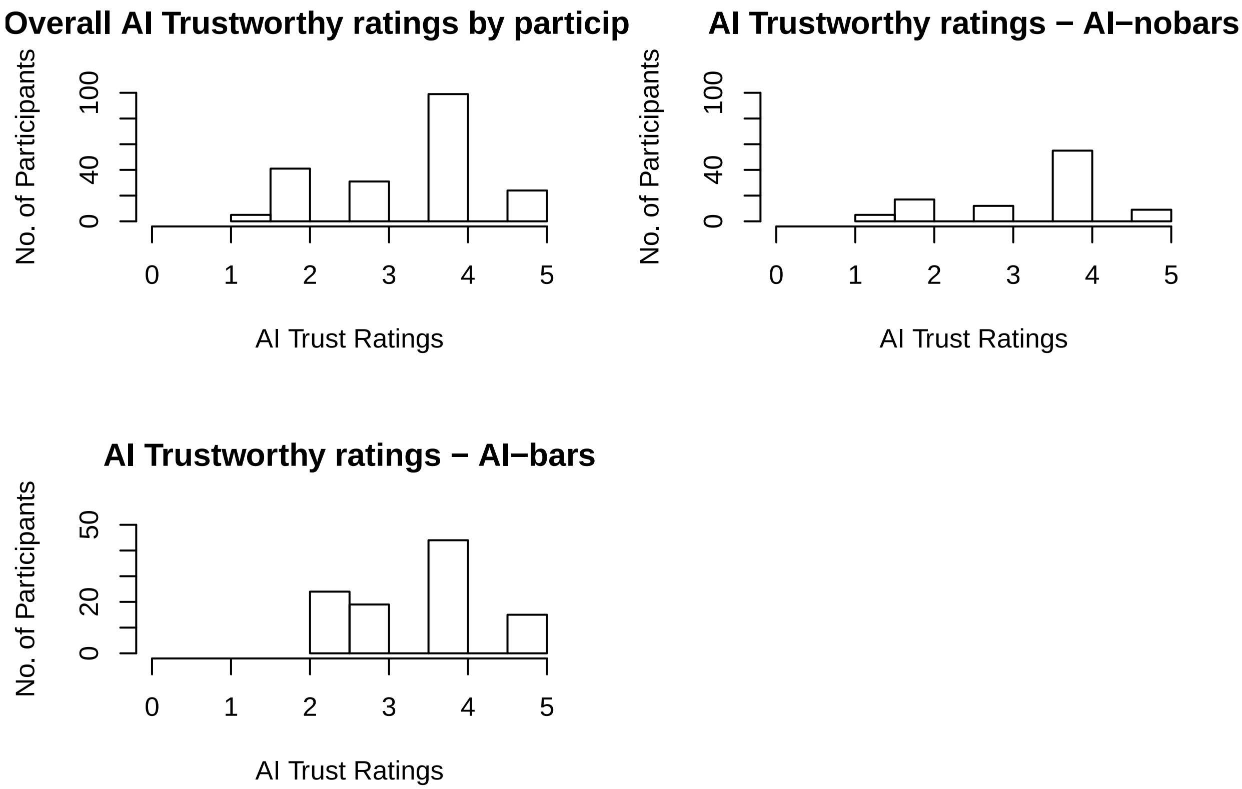


AI Trustworthiness

The overall average AI trust ratings in the experiment was 3.48 (SD = 1.03). The average AI trust rating for participants that did not receive uncertainty information was 3.47 (SD = 1.05) compared to participants who received the uncertainty information; mean = 3.49 (SD = 1.01).
Similiar to task difficulty, AI trustworthyness is also distributed without a pattern. A log transformation was performed but not included as it did not help.

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

#plotting AI Trust Ratings for AI vs No-AI
hist(person_noAI$AI_trust_num,
     main = "Overall AI Trustworthy ratings by participants",
     xlab = "AI Trust Ratings", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,100), xlim = c(0,5))

#plotting AI Trust Ratings for AI-nobars
hist(person_nobar$AI_trust_num,
     main = "AI Trustworthy ratings - AI-nobars",
     xlab = "AI Trust Ratings", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,100), xlim = c(0,5))

#plotting AI Trust Ratings for AI-nobars
```

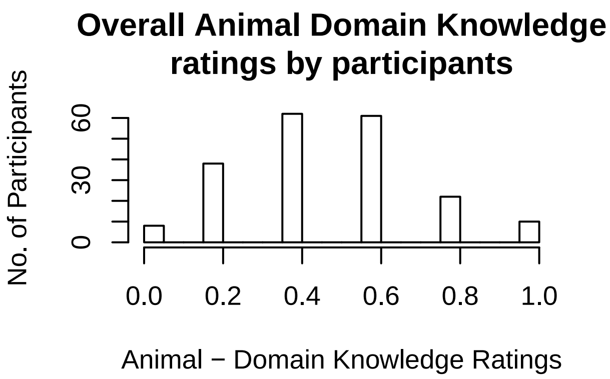
```r
hist(person_bar$AI_trust_num,
     main = "AI Trustworthy ratings - AI-bars",
     xlab = "AI Trust Ratings", ylab = "No. of Participants", breaks = "FD",
     ylim = c(0,50), xlim = c(0,5))
```

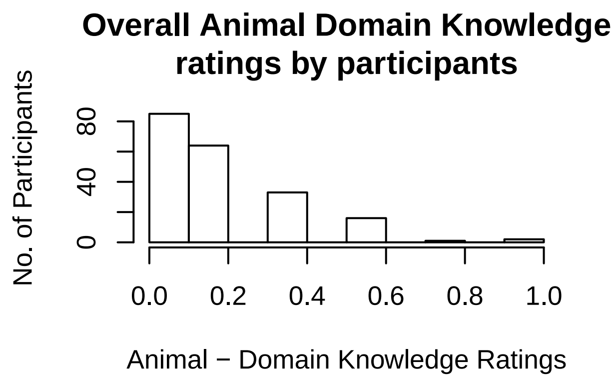


Domain Knowledge

Responses to Animal domain knowledge question is normally distributed whereas, responses to Plant domain knowledge question is not. A log transformation was performed which did not provide the satisfactory distribution. Therefore, plant domain knowledge will be represented in binary terms where participants who rated their knowledge as "not well at all" will be given 0, and the rest will be denoted with 1.

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

#plotting ANIMAL Domain Knowledge ratings
hist(person_noAI$Dmn_know_a_num,
     main = "Overall Animal Domain Knowledge\nratings by participants",
     xlab = "Animal - Domain Knowledge Ratings", ylab = "No. of Participants",
     breaks = "FD")

#plotting PLANT Domain Knowledge ratings
hist(person_noAI$Dmn_know_p_num,
     main = "Overall Animal Domain Knowledge\nratings by participants",
     xlab = "Animal - Domain Knowledge Ratings", ylab = "No. of Participants",
     breaks = "FD")
```





Bi-Variate Plots - Plotted against Accuracy (response variable)

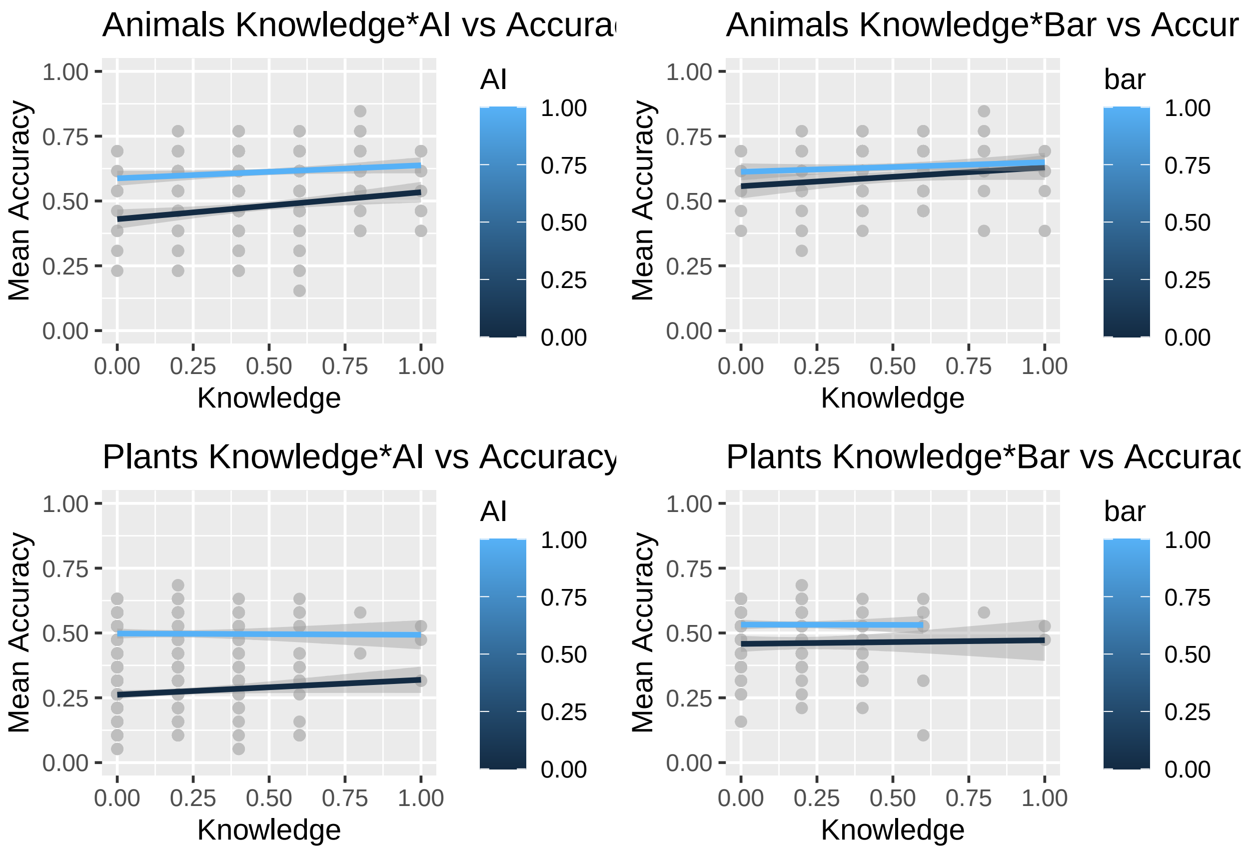
Domain Knowledge interaction with AI and Bar vs. Accuracy

Domain Knoelwedge interaction with AI and Bar for animals

```
#Filtering for AI vs No-AI
animals_person_noAI <- filter(animals_person, AI == 0)
animals_person_AI <- filter(animals_person, AI == 1)

#Filtering for Bar vs No-bar
animals_person_bar <- filter(animals_person_AI, bar == 1)
animals_person_nobar <- filter(animals_person_AI, bar == 0)

animal_dmn_AI_plot <- ggplot(animals_person) +
  aes(x = Dmn_know_a_num, y = accuracy, color = AI) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = animals_person_noAI) +
  geom_smooth(method = "lm", data = animals_person_AI) +
  xlab("Knowledge") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Animals Knowledge*AI vs Accuracy") #main plot title

animal_dmn_bar_plot <- ggplot(animals_person_AI) +
  aes(x = Dmn_know_a_num, y = accuracy, color = bar) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = animals_person_nobar) +
  geom_smooth(method = "lm", data = animals_person_bar) +
  xlab("Knowledge") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Animals Knowledge*Bar vs Accuracy") #main plot title

#Filtering for AI vs No-AI
plants_person_noAI <- filter(plants_person, AI == 0)
plants_person_AI <- filter(plants_person, AI == 1)

#Filtering for Bar vs No-bar
plants_person_bar <- filter(plants_person_AI, bar == 1)
plants_person_nobar <- filter(plants_person_AI, bar == 0)
```

```r
plant_dmn_AI_plot <- ggplot(plants_person) +
  aes(x = Dmn_know_p_num, y = accuracy, color = AI) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = plants_person_noAI) +
  geom_smooth(method = "lm", data = plants_person_AI) +
  xlab("Knowledge") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Plants Knowledge*AI vs Accuracy") #main plot title

plant_dmn_bar_plot <- ggplot(plants_person_AI) +
  aes(x = Dmn_know_p_num, y = accuracy, color = bar) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = plants_person_nobar) +
  geom_smooth(method = "lm", data = plants_person_bar) +
  xlab("Knowledge") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Plants Knowledge*Bar vs Accuracy") #main plot title

cowplot::plot_grid(animal_dmn_AI_plot, animal_dmn_bar_plot, plant_dmn_AI_plot, plant_dmn_bar_plot)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

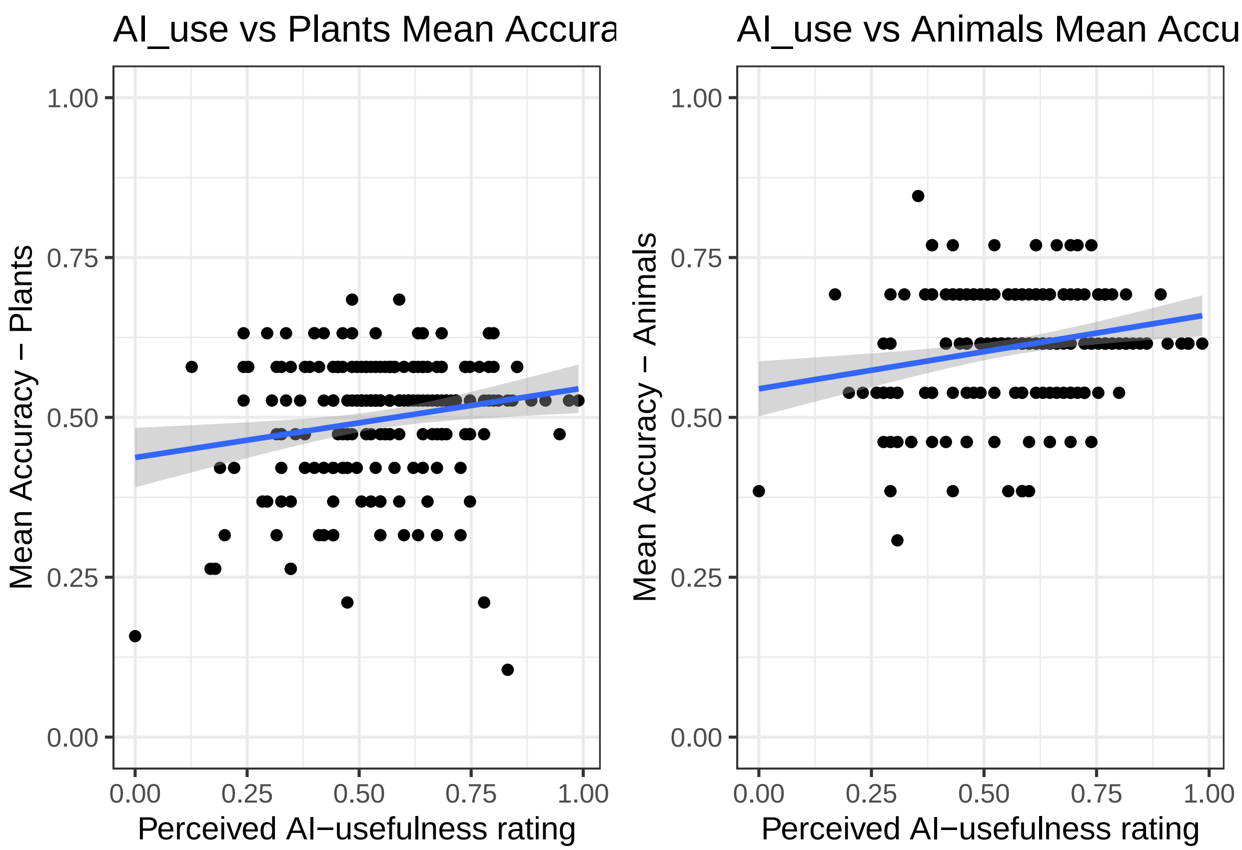


AI-usefulness rating vs Accuracy

The mean accuracy of the partcipants vs perceived AI-usefulness rating does not show a linear relationship. The fitted line is almost horizontal. The data points also do not indicate any relationship. But the analysis is split between plants and animals images. When looking at separate datasets, it is clear that there exists a linear relationship between AI usefulness ratings and accuracy.

```
# AI-use vs. Accuracy - all AI- conditions
AI_use_plot.1 <- filter(plants_person, AI ==1) %>%
  ggplot(aes(x=AI_use, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI-usefulness rating") +
  ylab("Mean Accuracy - Plants") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI_use vs Plants Mean Accuracy") #main plot title

# AI-use vs. Accuracy - all AI- conditions
AI_use_plot.2 <- filter(animals_person, AI ==1) %>%
  ggplot(aes(x=AI_use, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI-usefulness rating") +
  ylab("Mean Accuracy - Animals") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
```

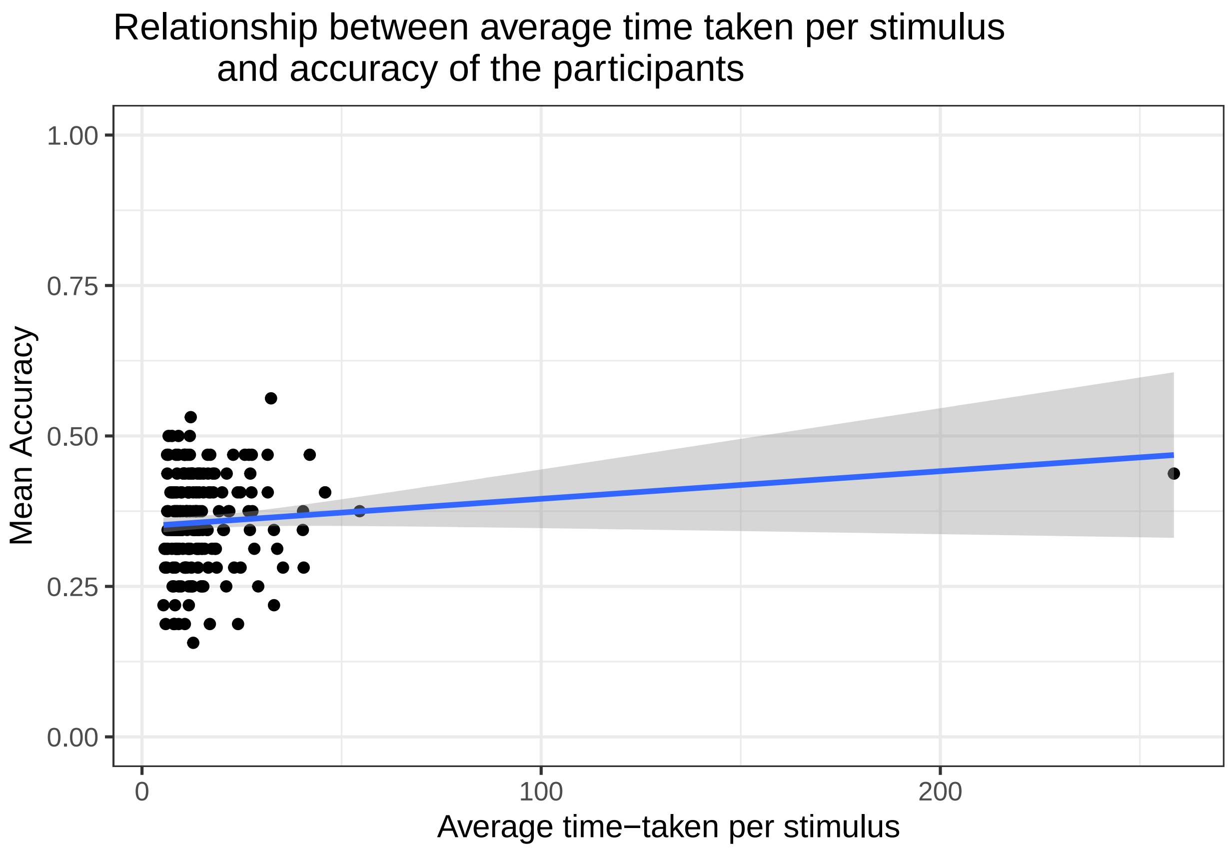
```r
  ggtitle("AI_use vs Animals Mean Accuracy") #main plot title


cowplot::plot_grid(AI_use_plot.1, AI_use_plot.2)
```



Time taken vs accuracy

There is a clear outlier that is affecting the fit of the data. To find a relationship between these two variables, the outlier will be removed. During the analysis, linear models will be fit with and without the outlier to see its effect on the results.

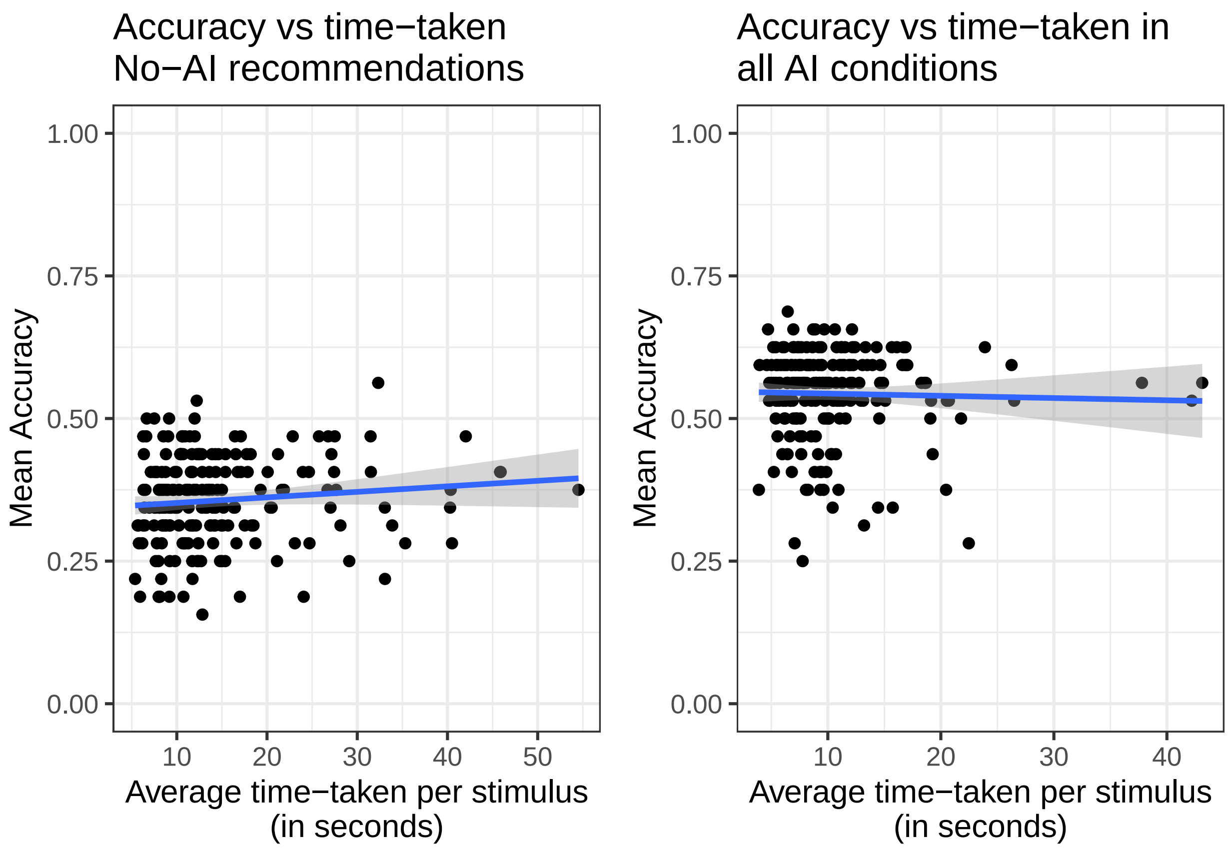
```r
#Time taken vs accuracy
time_taken_plot1 <- ggplot(person_noAI, aes(x=time_taken, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Relationship between average time taken per stimulus
          and accuracy of the participants") #main plot title

time_taken_plot1
```



There is no clear relationship between time taken and accuracy in AI or No-AI conditions. Even though an extreme outlier was removed, new outliers will come up since the range in average time taken in wide. Graphs supports that as well. Based on the plots generated, time taken does not have any effect on the accuracy of the participants.

```
#replotting the same graph without the outlier
person_noAI2 <- person_noAI[-201,]

#Time taken in no-AI condition
time_taken_plot_1 <- ggplot(person_noAI2, aes(x=time_taken, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus\n(in seconds)") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Accuracy vs time-taken\nNo-AI recommendations") #main plot title

#Time-taken in AI-condition
time_taken_plot_2 <- ggplot(person_AI, aes(x=time_taken, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus\n(in seconds)") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Accuracy vs time-taken in\nall AI conditions") #main plot title
```

```
cowplot::plot_grid(time_taken_plot_1, time_taken_plot_2)
```




```
#Not needed anymore. Just used to check without observation 201.
remove(person_noAI2)
```

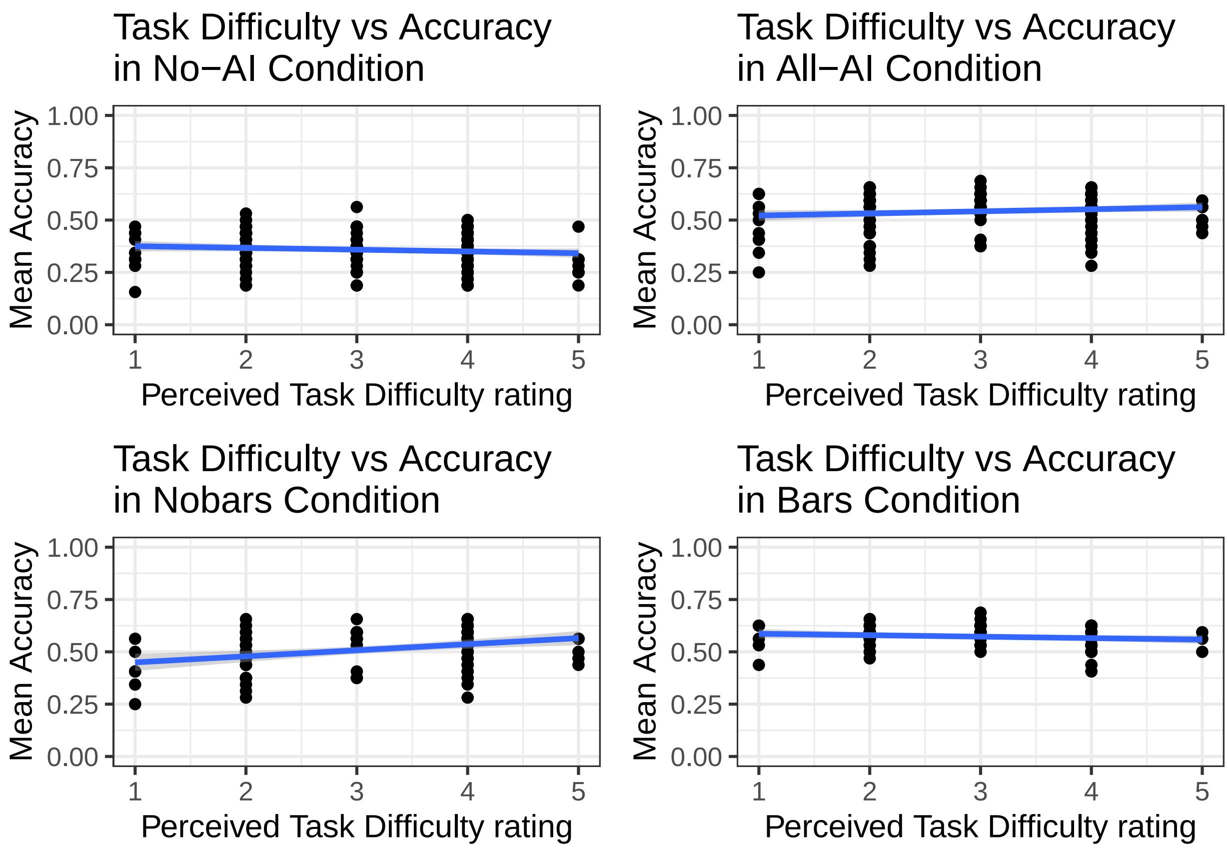
Task Difficulty vs accuracy

Compared to No-AI condition, the accuracy should improve in all AI conditions when the perceived task difficulty rating increases. Accuracy for participants placed in AI-nobar condition increases as participants task difficulty rating increases. However, there does not seem to be any relationship between their perceived ratings and accuracy for participants placed in AI-bars condition.

```
Task_diff_plot_1 <- ggplot(person_noAI, aes(x=Task_diff_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs Accuracy\nin No-AI Condition") #main plot title

Task_diff_plot_2 <- ggplot(person_AI, aes(x=Task_diff_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
```

```r
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs Accuracy\nin All-AI Condition") #main plot title

Task_diff_plot_3 <- ggplot(person_nobar, aes(x=Task_diff_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs Accuracy\nin Nobars Condition") #main plot title

Task_diff_plot_4 <- ggplot(person_bar, aes(x=Task_diff_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs Accuracy\nin Bars Condition") #main plot title

cowplot::plot_grid(Task_diff_plot_1, Task_diff_plot_2, Task_diff_plot_3, Task_diff_plot_4)
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```



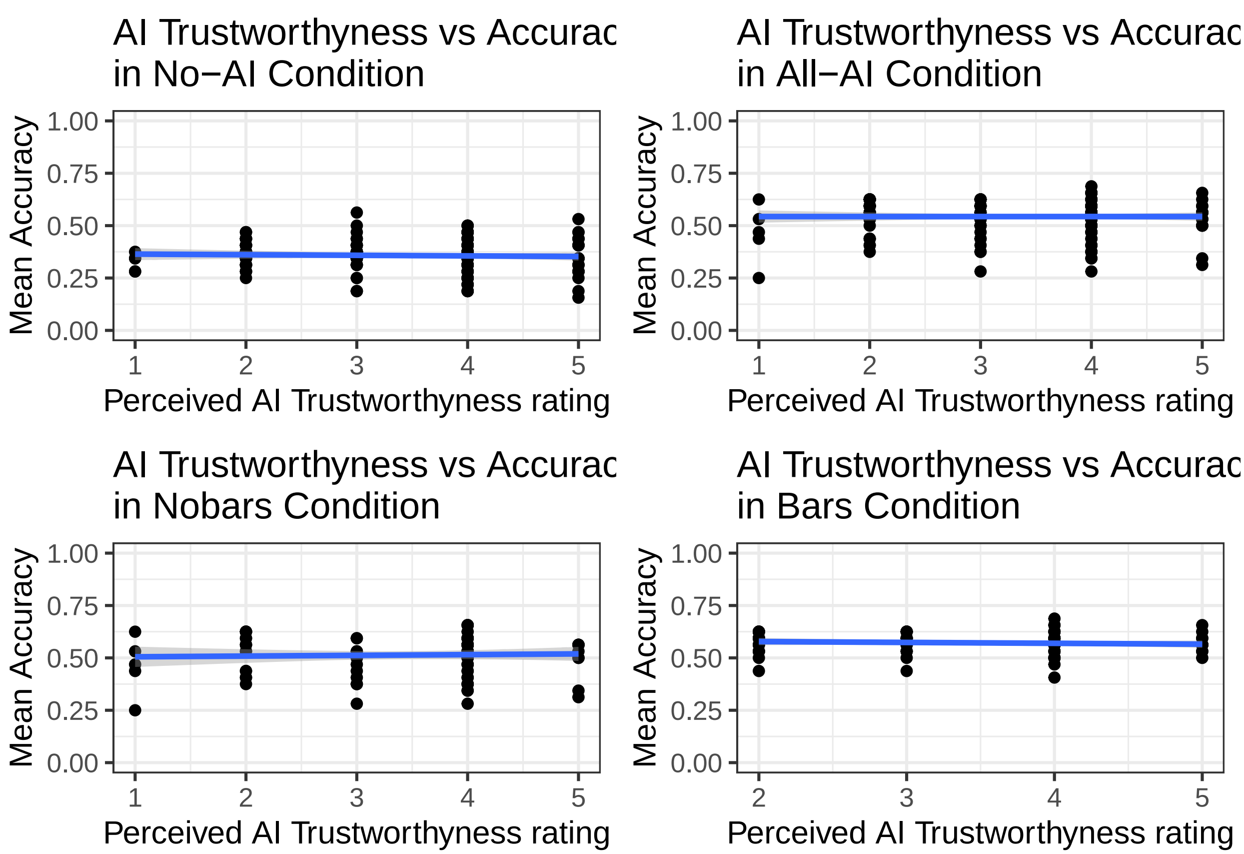
Perceived AI trustworthyness ratings vs Accuracy

There is no apparent relationship between AI trustworthyness and accuracy in any and all conditions.

```
AI_trust_plot_1 <- ggplot(person_noAI, aes(x=AI_trust_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs Accuracy\nin No-AI Condition") #main plot title

AI_trust_plot_2 <- ggplot(person_AI, aes(x=AI_trust_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs Accuracy\nin All-AI Condition") #main plot title

AI_trust_plot_3 <- ggplot(person_nobar, aes(x=AI_trust_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
```

```
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs Accuracy\nin Nobars Condition") #main plot title

AI_trust_plot_4 <- ggplot(person_bar, aes(x=AI_trust_num, y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs Accuracy\nin Bars Condition") #main plot title

cowplot::plot_grid(AI_trust_plot_1, AI_trust_plot_2, AI_trust_plot_3, AI_trust_plot_4)
```

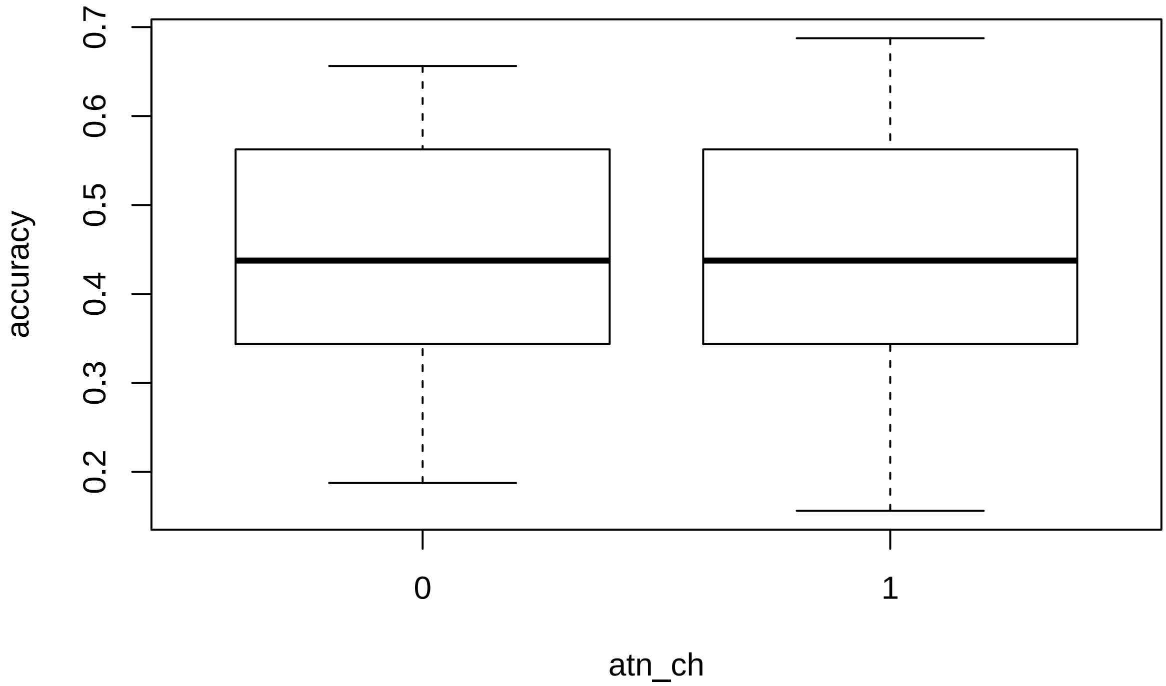
```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```



Attention Check vs Accuracy

There is no evidence suggesting accuracy will be affected by participants performance on attention checks.

```
boxplot(accuracy ~ atn_ch, data = person)
```




```
#Average accuracy in no-bar vs bar
t.test(accuracy ~ atn_ch, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  accuracy by atn_ch
## t = 0.12935, df = 171.27, p-value = 0.8972
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02642902  0.03013562
## sample estimates:
## mean in group 0 mean in group 1
##       0.4516226       0.4497693
```

```
#Effect Size of the t.test
cohen.d(accuracy ~ atn_ch, data = person)
```

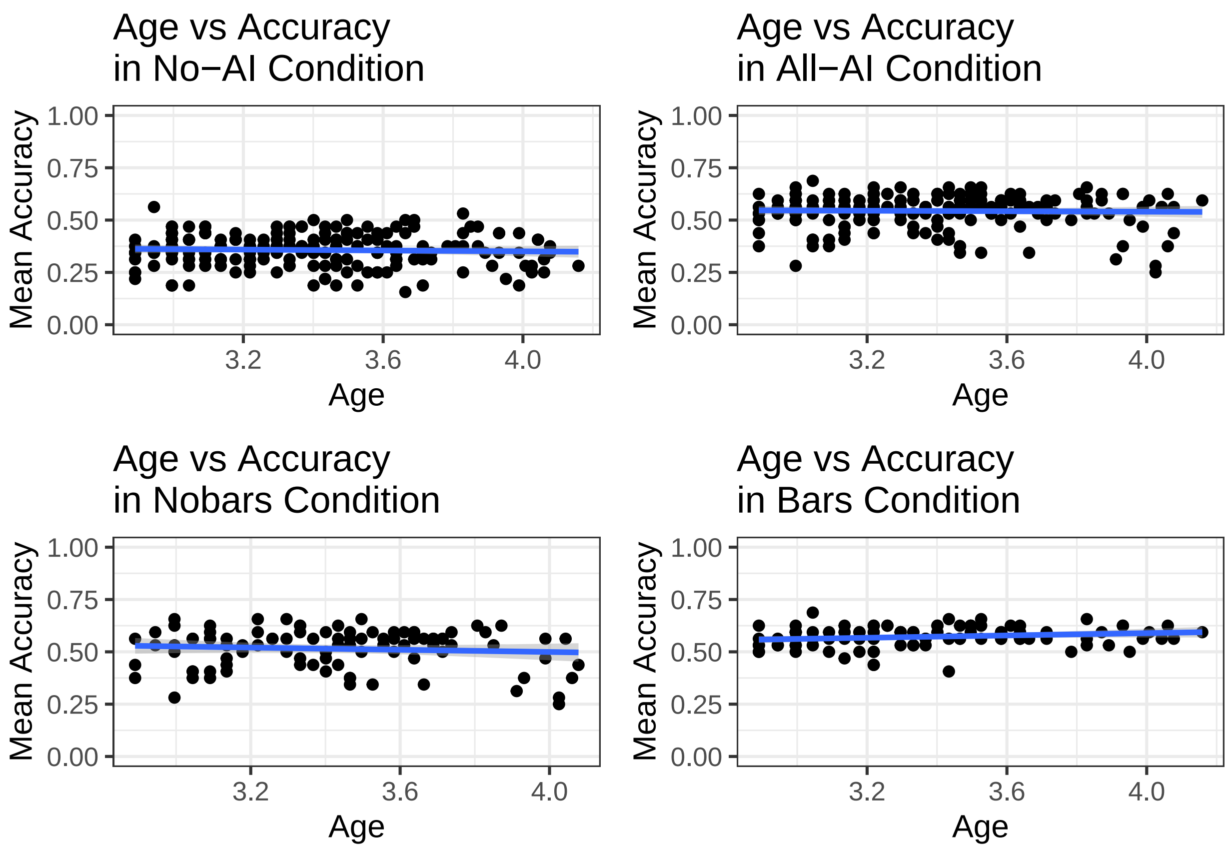
```
## Warning in cohen.d.formula(accuracy ~ atn_ch, data = person): Cohercing rhs of
## formula to factor
```

```
##
## Cohen's d
##
## d estimate: 0.01514839 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.2087532  0.2390500
```

Age vs Accuracy

There is no evidence suggesting accuracy will be affected by participants age.

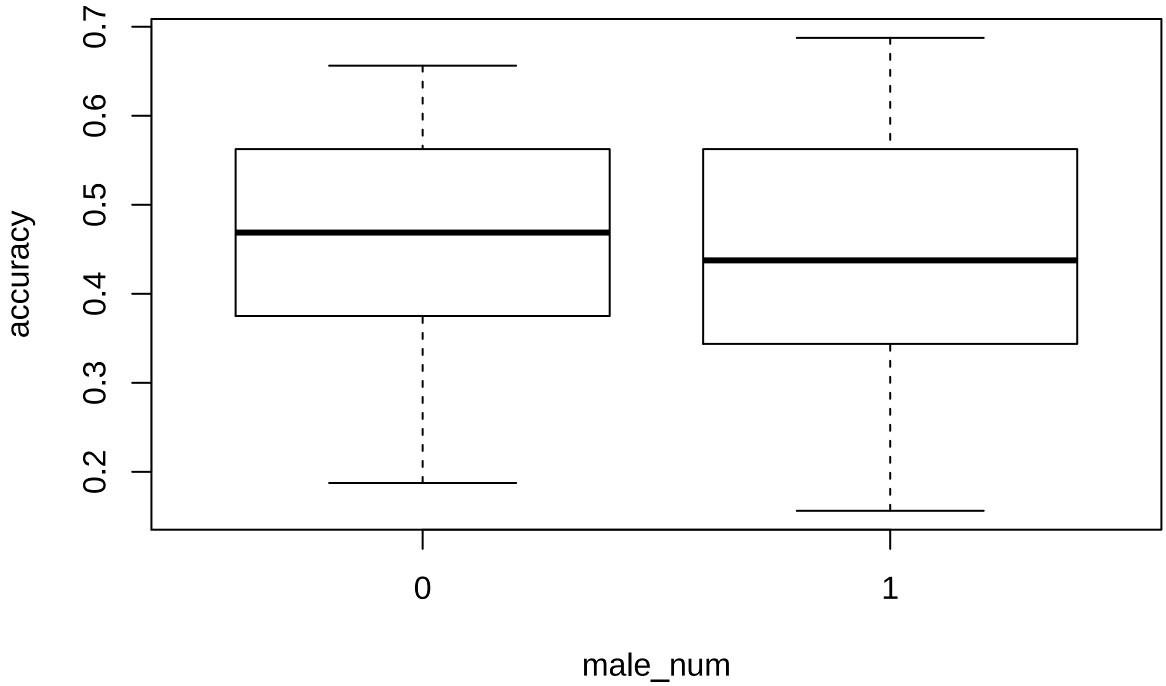
```r
age_plot_1 <- ggplot(person_noAI, aes(x=log(age), y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs Accuracy\nin No-AI Condition") #main plot title

age_plot_2 <- ggplot(person_AI, aes(x=log(age), y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs Accuracy\nin All-AI Condition") #main plot title

age_plot_3 <- ggplot(person_nobar, aes(x=log(age), y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs Accuracy\nin Nobars Condition") #main plot title

age_plot_4 <- ggplot(person_bar, aes(x=log(age), y = accuracy)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean Accuracy") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs Accuracy\nin Bars Condition") #main plot title

cowplot::plot_grid(age_plot_1, age_plot_2, age_plot_3, age_plot_4)
```



Gender vs Accuracy

Although there is a slight difference, it is not significant and the effect size is negligible. Gender should not affect accuracy.

```
boxplot(accuracy ~ male_num, data = person)
```




```r
t.test(accuracy ~ male_num, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  accuracy by male_num
## t = 0.96674, df = 393.64, p-value = 0.3343
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01220819  0.03582989
## sample estimates:
## mean in group 0 mean in group 1
##       0.4560073       0.4441964
```

```r
cohen.d(accuracy ~ male_num, data = person)
```

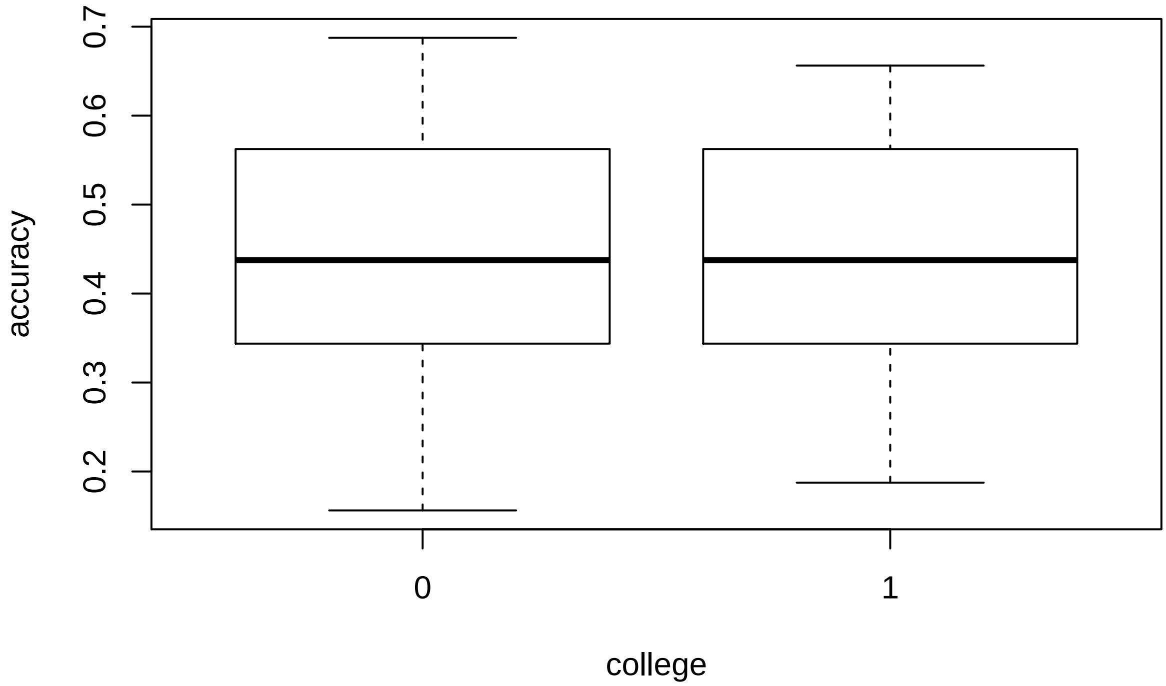
```
## Warning in cohen.d.formula(accuracy ~ male_num, data = person): Cohercing rhs of
## formula to factor
```

```
##
## Cohen's d
##
## d estimate: 0.0966498 (negligible)
## 95 percent confidence interval:
##       lower       upper
## -0.09962691  0.29292651
```

College vs Accuracy

Participants education level should not affect accuracy.

```r
boxplot(accuracy ~ college, data = person)
```




```r
t.test(accuracy ~ college, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  accuracy by college
## t = -0.42957, df = 397.02, p-value = 0.6677
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02926347  0.01876833
## sample estimates:
## mean in group 0 mean in group 1
##       0.4477163       0.4529639
```

```r
cohen.d(accuracy ~ college, data = person)
```

```
## Warning in cohen.d.formula(accuracy ~ college, data = person): Cohercing rhs of
## formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.04290122 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.2391444  0.1533419
```

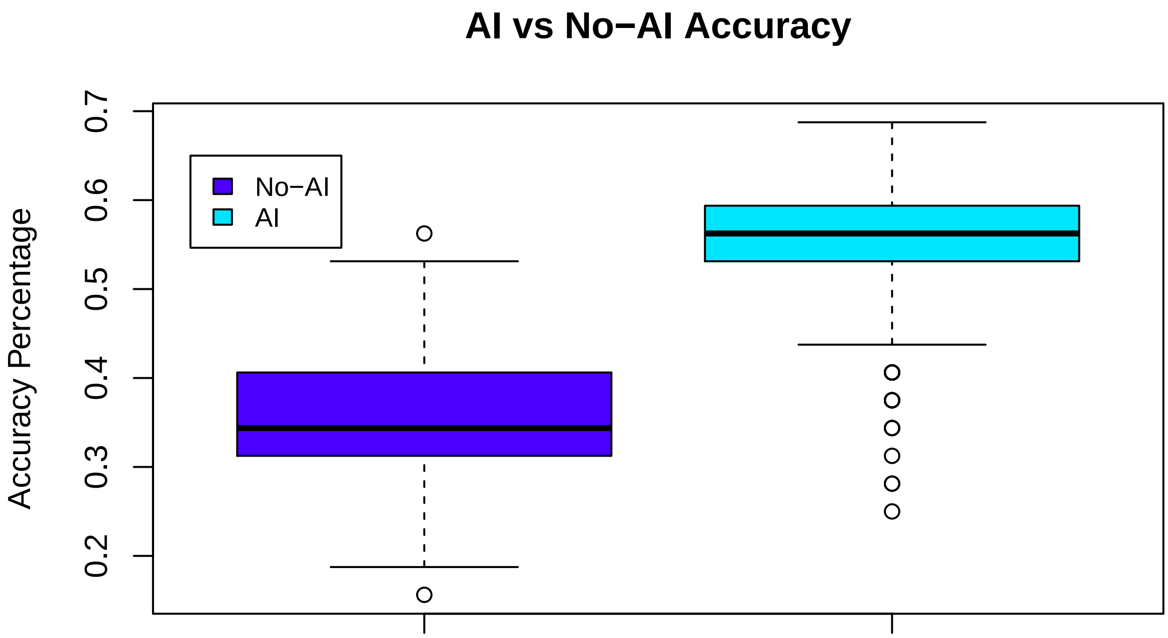
END OF EXPLORATORY PLOTS FOR DATASET WHERE EACH OBSERVATION WAS A PERSON

Box-Plots and T-test

Comparing Accuracy Pre-AI and Post-AI

The box plot clearly indicates a positive change in accuracy when participants were provided AI recommendations. The t-test results along with the box plots shows that AI recommendations will improve accuracy. The repeated measures paired t-test indicates that accuracy was significantly higher when AI information was provided (M = 0.187). The significant t-test with t(200) = -30.426, $p < 0.0001$. An effect size d = -2.36 proves the change in acccuracy is most likely not due to chance.

```
#Boxplot
boxplot(person_noAI$accuracy,#box plot. Comparing two sets of data
        person_AI$accuracy, #comparing no-AI vs AI accuracy
        col= topo.colors(2), #mentioning colors of the boxes
        main = "AI vs No-AI Accuracy", #main title
        ylab = "Accuracy Percentage") #y-axis title
legend(.5,.65, #legend, mentioning where in the graph to place
       inset = 0.2, c("No-AI","AI"), #mentioning text for the legend
       fill = topo.colors(2), cex=0.8) #providing color info for the boxes.
```



```r
#Average accuracy in AI vs no_AI
t.test(person_noAI$accuracy, person_AI$accuracy, #t-test to compare.
       paired = TRUE) #paired is true as this is a within-subjects comparison
```

```
##
##  Paired t-test
##
## data:  person_noAI$accuracy and person_AI$accuracy
## t = -30.426, df = 200, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1986587 -0.1744756
## sample estimates:
## mean of the differences
##              -0.1865672
```

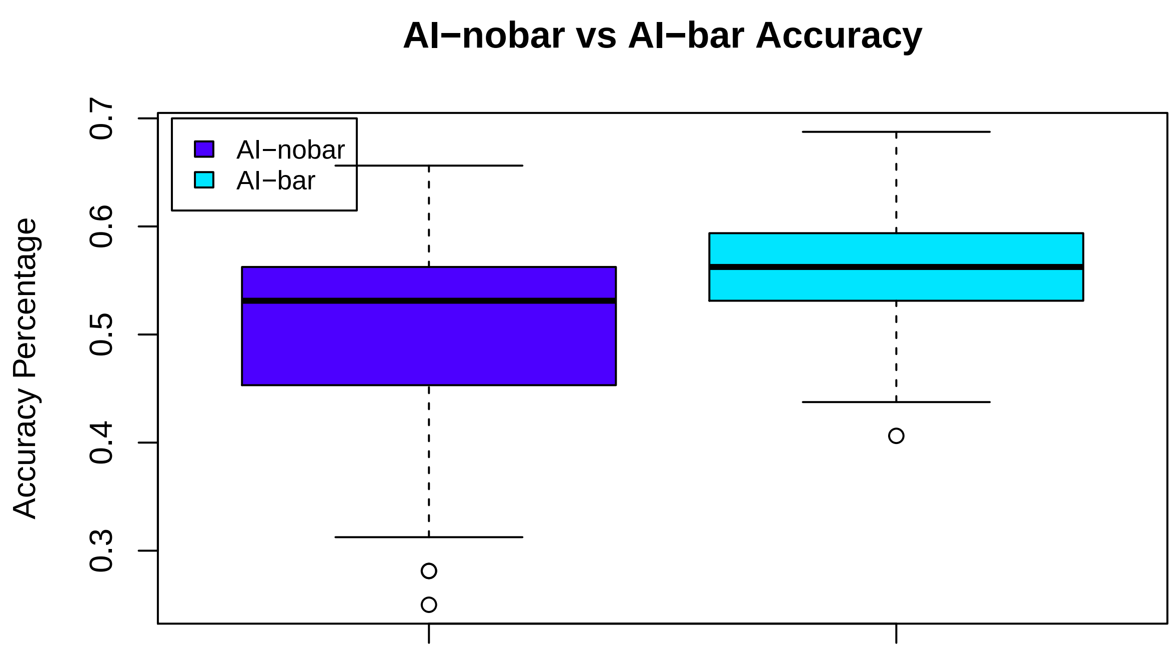
```r
#Effect Size of the t.test
cohen.d(person_noAI$accuracy, person_AI$accuracy, #effect size of the comparison
        paired = TRUE)
```

```
##
## Cohen's d
##
## d estimate: -2.364885 (large)
## 95 percent confidence interval:
##     lower     upper
## -2.662612 -2.067157
```

Comparing Accuracy in AI-nobar and AI-bar conditions

By looking at the box plot, it is clear accuracy of the participants who received uncertainty information performed better compared to participants who did not. The average accuracy changed significantly when comparing AI vs No-AI performance but the change, although significant, is less pronounced, when comparing between participants receiving uncertainty information vs participants not receiving uncertainty information. The Two sample t-test indicates that accuracy was significantly higher when uncertainty information was provided (M = 0.572) compared to when uncertinaty information was not provided (M = 0.515). The significant t-test with $t(147.6) = -5.40$, $p < 0.0001$. An effect size $d = -0.77$ proves the change in acccuracy is most likely not due to chance.

```r
#Boxplot
boxplot(person_nobar$accuracy, person_bar$accuracy,
        col= topo.colors(2),
        main = "AI-nobar vs AI-bar Accuracy",
        ylab = "Accuracy Percentage")
legend(.45,.70, inset = 0.2, c("AI-nobar","AI-bar"), fill = topo.colors(2), cex=0.8)
```




```r
#Average accuracy in no-bar vs bar
t.test(person_nobar$accuracy, person_bar$accuracy)
```

```
##
##  Welch Two Sample t-test
##
## data:  person_nobar$accuracy and person_bar$accuracy
## t = -5.401, df = 147.6, p-value = 2.594e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.07808919 -0.03625276
## sample estimates:
## mean of x mean of y
## 0.5145202 0.5716912
```

```r
#Effect Size of the t.test
cohen.d(person_nobar$accuracy, person_bar$accuracy)
```

```
##
## Cohen's d
##
## d estimate: -0.7684853 (medium)
## 95 percent confidence interval:
##      lower      upper
## -1.0567824 -0.4801881
```

Examining the correlations between predictor variables. None of the correlations are beyond 0.58. Some correlated variables: 1. AI information (AI) & Uncertainty (Bar) 0.58, $p < 0.001$

Uncertainty information was provided only when AI information was shown. Nevertheless, AI and Bar will be examined in separate regressions. So, we need not worry about multicollinearity.

2. Animal Domain Knowledge & Plant Domain Knowledge 0.40, $p < 0.001$

Plants and Animals will be examined in seperate regressions when domain knowledge is included as a predicted variable. So, we need not worry about multicollinearity

3. AI Trustworthy rating (AI Trsut num) & AI usefulness rating (AI use) 0.41, $p < 0.001$

Both these variables will be used together as predictor variables in some of the models. Needs discussion!!!!

4. Accuracy and Confidence 0.37, $p < 0.001$ These are two response variables modeled separately. No need to worry about multicollinearity
5. Confidence and AI-use 0.62, $p < 0.001$
6. All other variables had a correlation value less than 0.50

```r
# Checking for strong correlations between predictor variables
rcorr(as.matrix(person[,c("AI","bar","Dmn_know_a_num",
                          "Dmn_know_p_num", "AI_use","time_taken",
                          "Task_diff_num", "AI_trust_num", "atn_ch","age",
                          "male_num","college", "accuracy", "confidence",
                          "over_conf")]))
```

```
##                   AI   bar Dmn_know_a_num Dmn_know_p_num AI_use time_taken
## AI              1.00  0.58           0.00           0.00    NaN      -0.19
## bar             0.58  1.00          -0.03          -0.05   0.07      -0.11
## Dmn_know_a_num  0.00 -0.03           1.00           0.40   0.04       0.07
## Dmn_know_p_num  0.00 -0.05           0.40           1.00   0.25       0.11
## AI_use           NaN  0.07           0.04           0.25   1.00       0.03
## time_taken     -0.19 -0.11           0.07           0.11   0.03       1.00
## Task_diff_num   0.00 -0.04          -0.12          -0.24  -0.11      -0.10
## AI_trust_num    0.00  0.01          -0.02           0.15   0.41       0.09
## atn_ch          0.00 -0.14          -0.17          -0.18  -0.08      -0.02
## age             0.00 -0.04          -0.03           0.09   0.00       0.18
## male_num        0.00  0.04          -0.02          -0.01   0.14      -0.06
## college         0.00 -0.07          -0.05          -0.12   0.01       0.03
## accuracy        0.76  0.58           0.06           0.06   0.24      -0.10
## confidence      0.43  0.29           0.32           0.31   0.62       0.07
## over_conf      -0.13 -0.13           0.28           0.27   0.46       0.14
##                Task_diff_num AI_trust_num atn_ch   age male_num college
## AI                      0.00         0.00   0.00  0.00     0.00    0.00
## bar                    -0.04         0.01  -0.14 -0.04     0.04   -0.07
## Dmn_know_a_num         -0.12        -0.02  -0.17 -0.03    -0.02   -0.05
## Dmn_know_p_num         -0.24         0.15  -0.18  0.09    -0.01   -0.12
## AI_use                 -0.11         0.41  -0.08  0.00     0.14    0.01
## time_taken             -0.10         0.09  -0.02  0.18    -0.06    0.03
## Task_diff_num           1.00        -0.11   0.06  0.05    -0.04    0.26
## AI_trust_num           -0.11         1.00   0.00  0.05     0.16   -0.04
## atn_ch                  0.06         0.00   1.00 -0.14    -0.01    0.05
## age                     0.05         0.05  -0.14  1.00    -0.06    0.28
## male_num               -0.04         0.16  -0.01 -0.06     1.00   -0.03
## college                 0.26        -0.04   0.05  0.28    -0.03    1.00
## accuracy                0.01        -0.01  -0.01 -0.04    -0.05    0.02
## confidence             -0.17         0.10  -0.14 -0.03     0.06   -0.02
## over_conf              -0.17         0.11  -0.13 -0.01     0.10   -0.03
##                accuracy confidence over_conf
## AI                 0.76       0.43     -0.13
## bar                0.58       0.29     -0.13
## Dmn_know_a_num     0.06       0.32      0.28
## Dmn_know_p_num     0.06       0.31      0.27
## AI_use             0.24       0.62      0.46
## time_taken        -0.10       0.07      0.14
```

```
## Task_diff_num      0.01      -0.17     -0.17
## AI_trust_num      -0.01       0.10      0.11
## atn_ch            -0.01      -0.14     -0.13
## age               -0.04      -0.03     -0.01
## male_num          -0.05       0.06      0.10
## college            0.02      -0.02     -0.03
## accuracy           1.00       0.37     -0.36
## confidence         0.37       1.00      0.73
## over_conf         -0.36       0.73      1.00
##
## n
##                 AI bar Dmn_know_a_num Dmn_know_p_num AI_use time_taken
## AI             402 402            402            402    201        402
## bar            402 402            402            402    201        402
## Dmn_know_a_num 402 402            402            402    201        402
## Dmn_know_p_num 402 402            402            402    201        402
## AI_use         201 201            201            201    201        201
## time_taken     402 402            402            402    201        402
## Task_diff_num  400 400            400            400    200        400
## AI_trust_num   400 400            400            400    200        400
## atn_ch         402 402            402            402    201        402
## age            402 402            402            402    201        402
## male_num       402 402            402            402    201        402
## college        402 402            402            402    201        402
## accuracy       402 402            402            402    201        402
## confidence     402 402            402            402    201        402
## over_conf      402 402            402            402    201        402
##                Task_diff_num AI_trust_num atn_ch age male_num college accuracy
## AI                       400          400    402 402      402     402      402
## bar                      400          400    402 402      402     402      402
## Dmn_know_a_num           400          400    402 402      402     402      402
## Dmn_know_p_num           400          400    402 402      402     402      402
## AI_use                   200          200    201 201      201     201      201
## time_taken               400          400    402 402      402     402      402
## Task_diff_num            400          398    400 400      400     400      400
## AI_trust_num             398          400    400 400      400     400      400
## atn_ch                   400          400    402 402      402     402      402
## age                      400          400    402 402      402     402      402
## male_num                 400          400    402 402      402     402      402
## college                  400          400    402 402      402     402      402
## accuracy                 400          400    402 402      402     402      402
## confidence               400          400    402 402      402     402      402
## over_conf                400          400    402 402      402     402      402
##                confidence over_conf
## AI                    402       402
## bar                   402       402
## Dmn_know_a_num        402       402
## Dmn_know_p_num        402       402
## AI_use                201       201
## time_taken            402       402
## Task_diff_num         400       400
## AI_trust_num          400       400
## atn_ch                402       402
## age                   402       402
```

```
## male_num             402       402
## college              402       402
## accuracy             402       402
## confidence           402       402
## over_conf            402       402
##
## P
##               AI     bar    Dmn_know_a_num Dmn_know_p_num AI_use time_taken
## AI                   0.0000 1.0000         1.0000                0.0001
## bar           0.0000        0.4886         0.3596         0.3111 0.0265
## Dmn_know_a_num 1.0000 0.4886                0.0000         0.5378 0.1550
## Dmn_know_p_num 1.0000 0.3596 0.0000                        0.0004 0.0251
## AI_use               0.3111 0.5378         0.0004                0.6345
## time_taken    0.0001 0.0265 0.1550         0.0251         0.6345
## Task_diff_num 1.0000 0.4308 0.0126         0.0000         0.1152 0.0408
## AI_trust_num  1.0000 0.9076 0.7155         0.0020         0.0000 0.0737
## atn_ch        1.0000 0.0054 0.0009         0.0004         0.2320 0.6624
## age           1.0000 0.3719 0.5017         0.0774         0.9793 0.0002
## male_num      1.0000 0.4548 0.6723         0.9079         0.0500 0.2016
## college       1.0000 0.1542 0.3011         0.0163         0.8544 0.5858
## accuracy      0.0000 0.0000 0.2681         0.2533         0.0006 0.0450
## confidence    0.0000 0.0000 0.0000         0.0000         0.0000 0.1713
## over_conf     0.0087 0.0070 0.0000         0.0000         0.0000 0.0043
##               Task_diff_num AI_trust_num atn_ch age    male_num college
## AI            1.0000        1.0000       1.0000 1.0000 1.0000   1.0000
## bar           0.4308        0.9076       0.0054 0.3719 0.4548   0.1542
## Dmn_know_a_num 0.0126       0.7155       0.0009 0.5017 0.6723   0.3011
## Dmn_know_p_num 0.0000       0.0020       0.0004 0.0774 0.9079   0.0163
## AI_use        0.1152        0.0000       0.2320 0.9793 0.0500   0.8544
## time_taken    0.0408        0.0737       0.6624 0.0002 0.2016   0.5858
## Task_diff_num               0.0228       0.2034 0.3057 0.3910   0.0000
## AI_trust_num  0.0228                     0.9929 0.2996 0.0018   0.4266
## atn_ch        0.2034        0.9929              0.0047 0.7689   0.3409
## age           0.3057        0.2996       0.0047        0.2647   0.0000
## male_num      0.3910        0.0018       0.7689 0.2647          0.6065
## college       0.0000        0.4266       0.3409 0.0000 0.6065
## accuracy      0.8919        0.8152       0.8943 0.4700 0.3333   0.6676
## confidence    0.0008        0.0389       0.0066 0.5052 0.2054   0.7329
## over_conf     0.0005        0.0252       0.0086 0.8901 0.0472   0.5105
##               accuracy confidence over_conf
## AI            0.0000   0.0000     0.0087
## bar           0.0000   0.0000     0.0070
## Dmn_know_a_num 0.2681  0.0000     0.0000
## Dmn_know_p_num 0.2533  0.0000     0.0000
## AI_use        0.0006   0.0000     0.0000
## time_taken    0.0450   0.1713     0.0043
## Task_diff_num 0.8919   0.0008     0.0005
## AI_trust_num  0.8152   0.0389     0.0252
## atn_ch        0.8943   0.0066     0.0086
## age           0.4700   0.5052     0.8901
## male_num      0.3333   0.2054     0.0472
## college       0.6676   0.7329     0.5105
## accuracy               0.0000     0.0000
## confidence    0.0000              0.0000
```

```
## over_conf      0.0000   0.0000
```

LINEAR MODELS ON ACCURACY

Effect of AI on accuracy

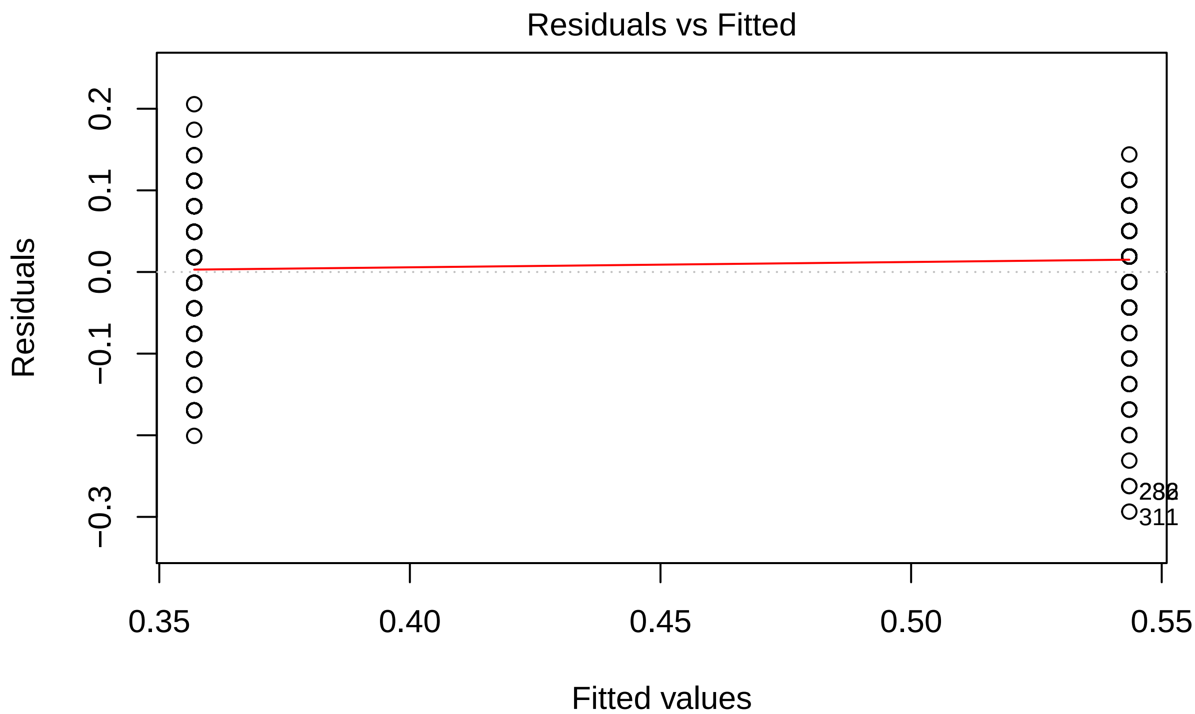
Results of the simple linear regression indicate a positive significant relationship between AI recommendations and accuracy (F(1,400) = 562.1, p <0.001, R2 = 0.58).

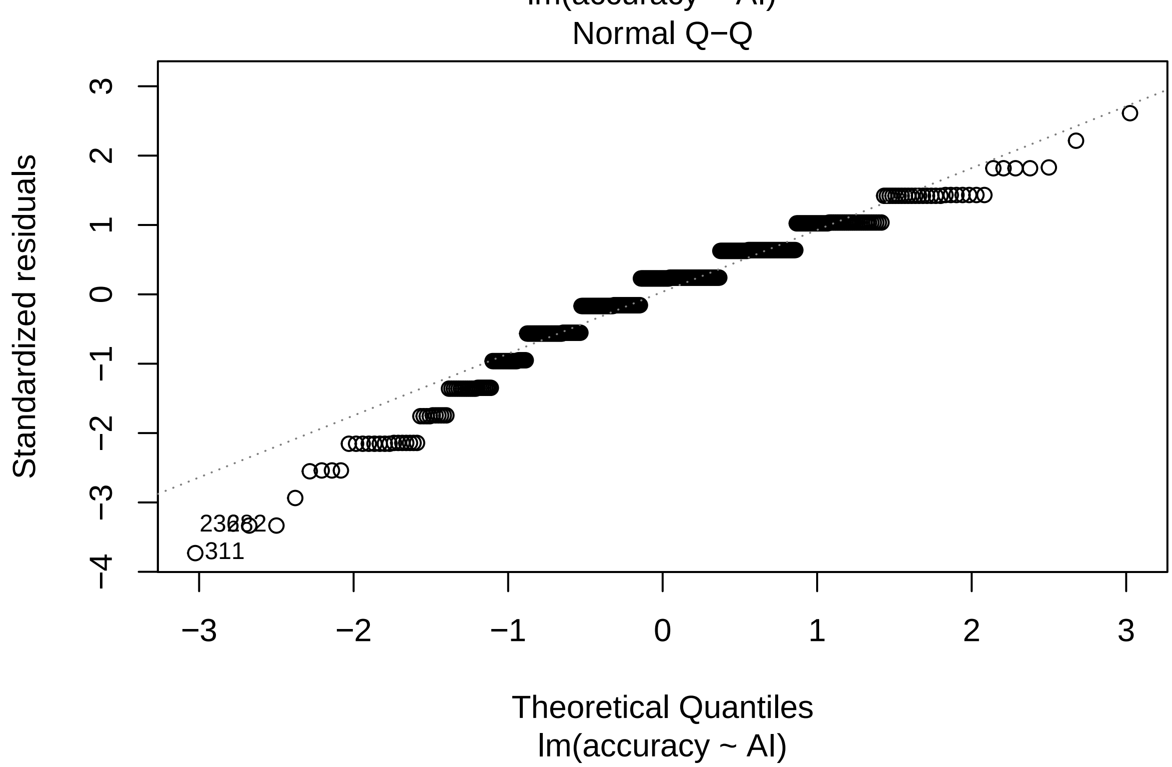
```r
lm.1.acc <- lm(accuracy ~ AI, data = person) #linear model
summary(lm.1.acc)
```

```
##
## Call:
## lm(formula = accuracy ~ AI, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.29353 -0.04446  0.01803  0.05022  0.20553
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.356965   0.005564   64.15   <2e-16 ***
## AI          0.186567   0.007869   23.71   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07889 on 400 degrees of freedom
## Multiple R-squared:  0.5842, Adjusted R-squared:  0.5832
## F-statistic: 562.1 on 1 and 400 DF,  p-value: < 2.2e-16
```

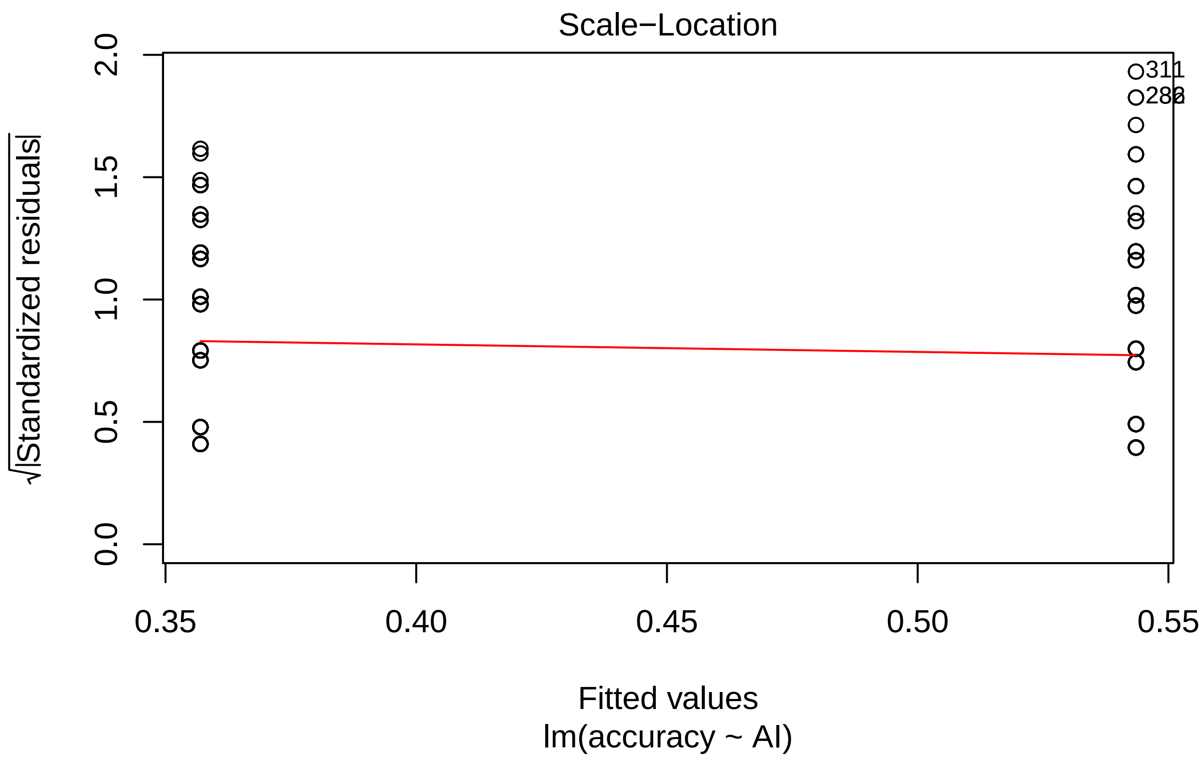
Given the predictor variable is binary, we see a pattern in the residuals vs fitted plot. Otherwise, the model is acceptable. The patter in the Q-Q plot is acceptable given the binary predictor variable, but it does deviate from the line at the edges.

```r
plot(lm.1.acc)
```






```
## hat values (leverages) are all = 0.004975124
##  and there are no factor predictors; no plot no. 5
```



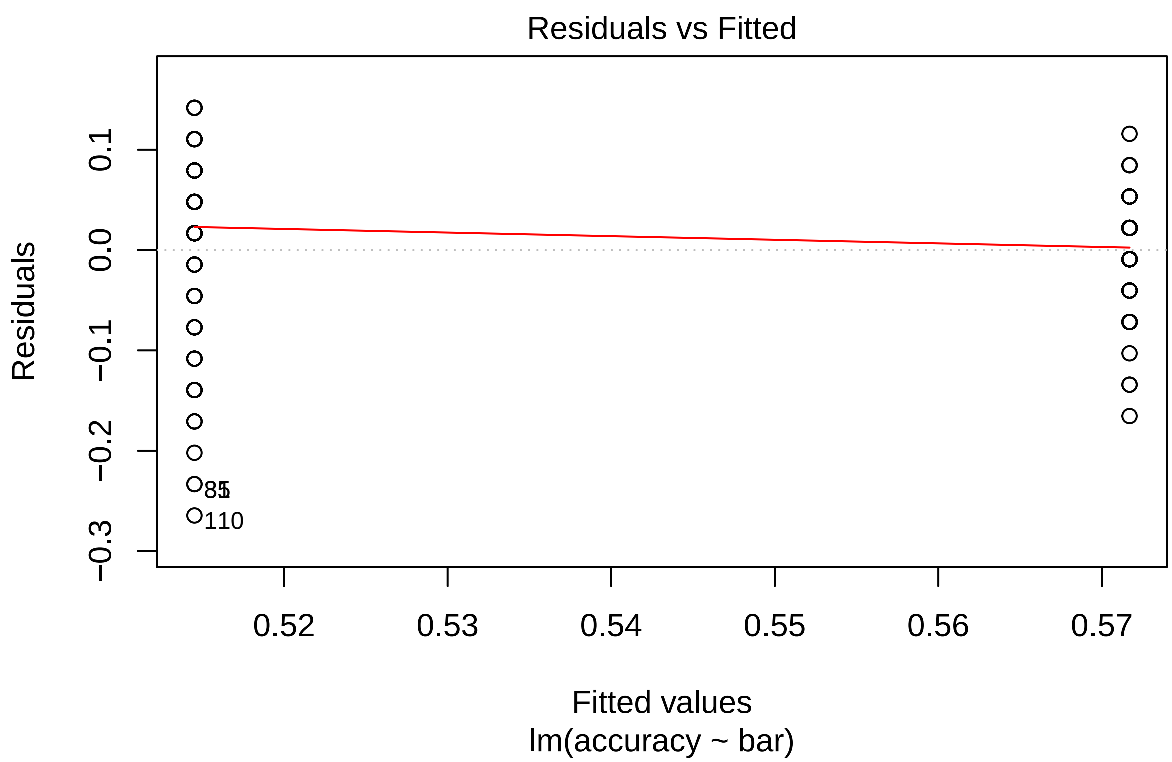
Effect of Uncertainty Information on accuracy

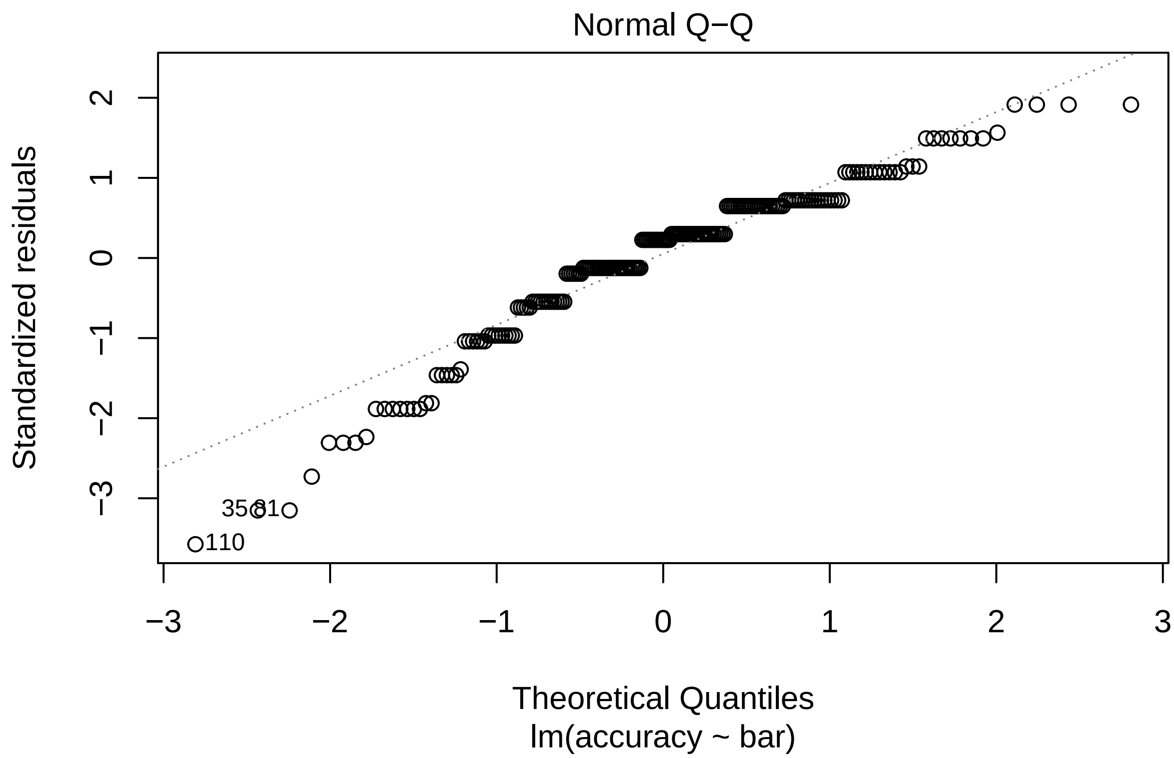
Results of the simple linear regression indicate a positive significant relationship between AI recommendations and accuracy as well as Unceraintry information and accuracy (F(1,199) = 29.67, p <0.001, R2 = 0.13).

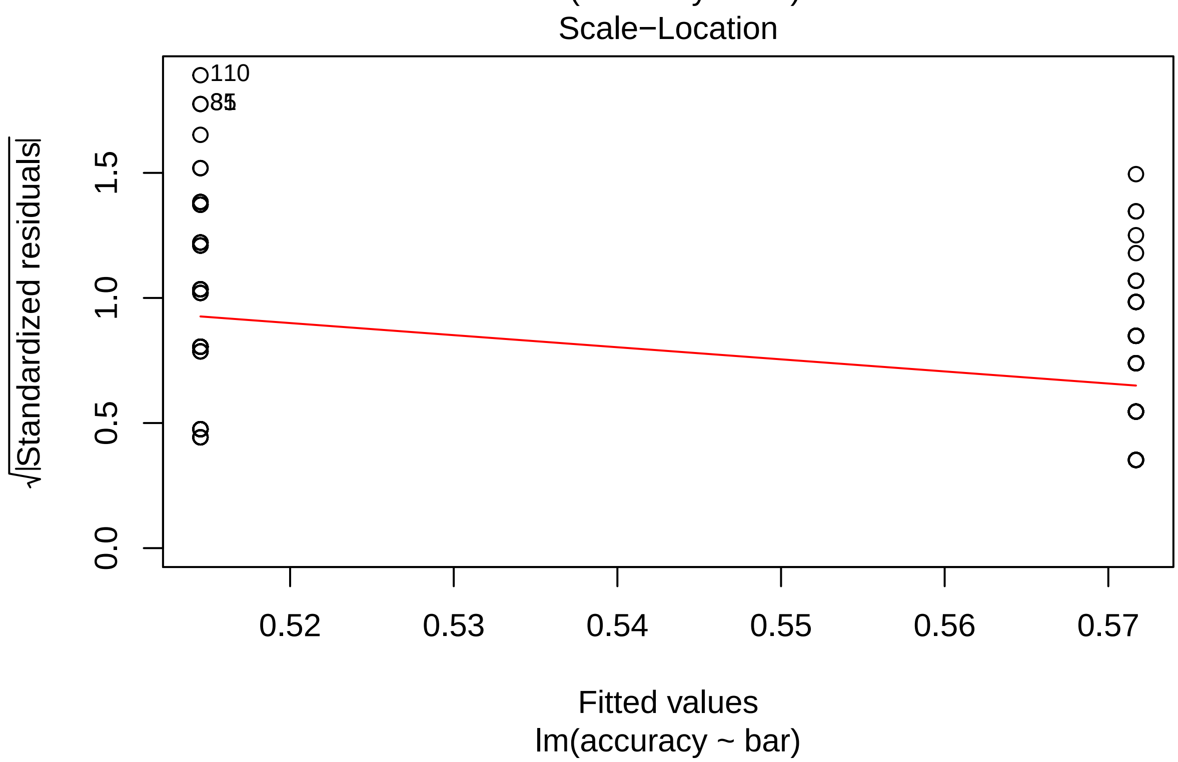
```
lm.2.acc <- lm(accuracy ~ bar, data = person_AI)

summary(lm.2.acc)
```

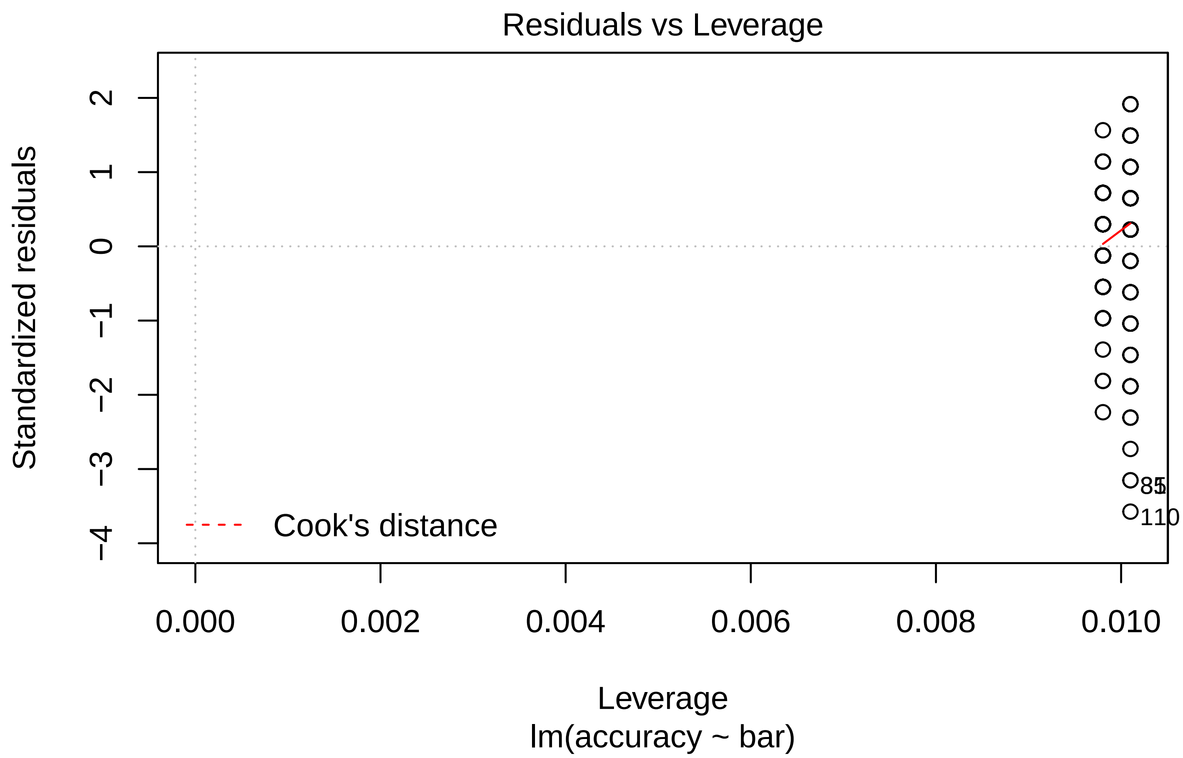
```
##
## Call:
## lm(formula = accuracy ~ bar, data = person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.26452 -0.04044  0.01673  0.04798  0.14173
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.514520   0.007477  68.815  < 2e-16 ***
## bar         0.057171   0.010496   5.447  1.5e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07439 on 199 degrees of freedom
## Multiple R-squared:  0.1297, Adjusted R-squared:  0.1254
## F-statistic: 29.67 on 1 and 199 DF,  p-value: 1.502e-07
```

```
plot(lm.2.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ bar)
110
35
81
−3
−2
−1
0
1
2
3


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ bar)
110
85
0.0
0.5
1.0
1.5
0.52
0.53
0.54
0.55
0.56
0.57



Effect of AI on accuracy with rest of the predictor variables

AI information is still significant. But none of the other variables are significant. Presence of AI information is significantly affecting the accuracy of the participants positively (F(8,389) = 69.13, p <0.001, R2 = 0.58).

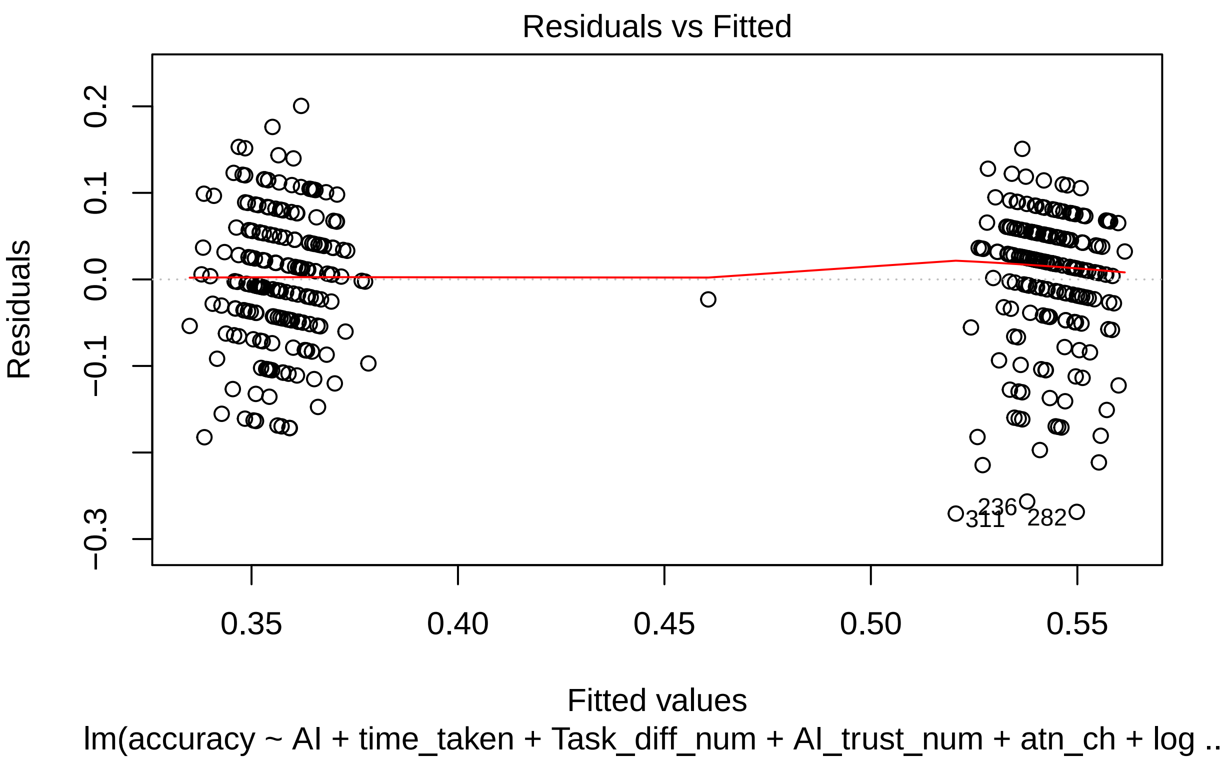
```
lm.3.acc <- lm(accuracy ~ AI + time_taken + Task_diff_num +
                 AI_trust_num + atn_ch + log(age) + male_num + college,
               data = person)

summary(lm.3.acc)
```

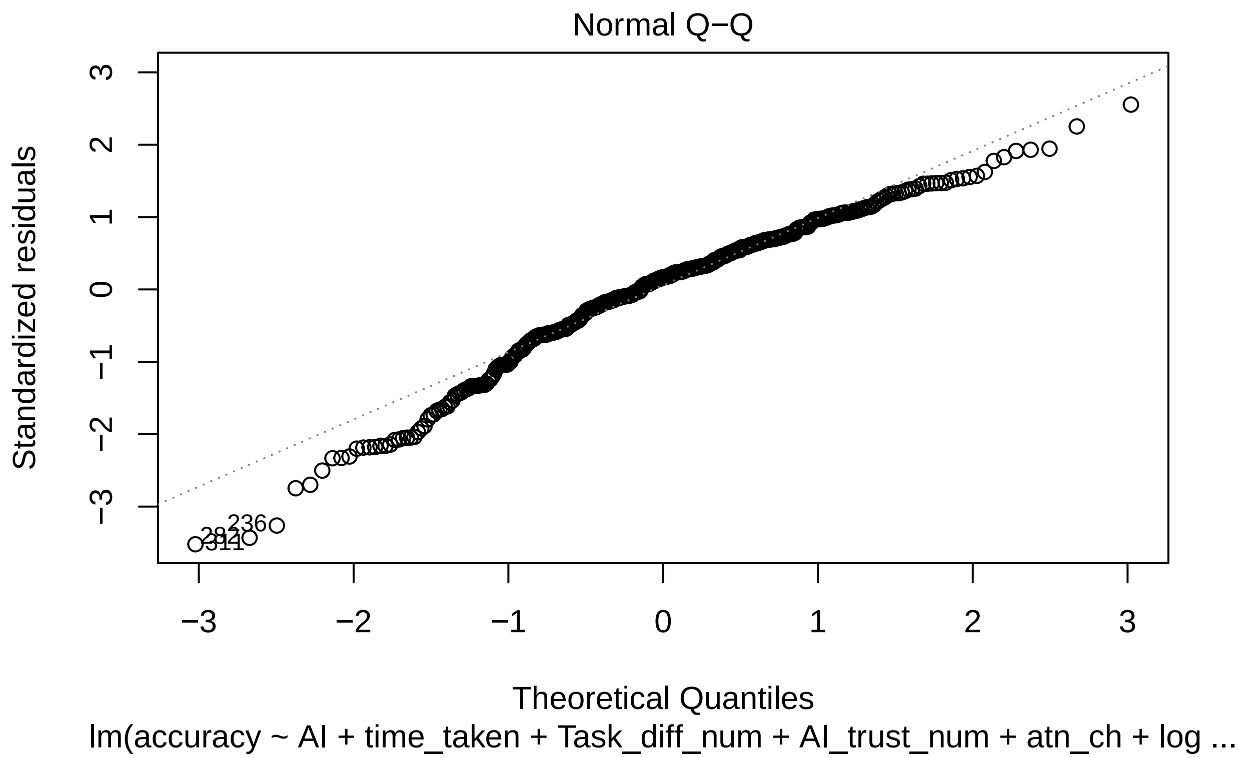
```
##
## Call:
## lm(formula = accuracy ~ AI + time_taken + Task_diff_num + AI_trust_num +
##     atn_ch + log(age) + male_num + college, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27056 -0.04375  0.01311  0.05387  0.20048
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4150302  0.0490526   8.461 5.45e-16 ***
## AI             0.1882646  0.0080981  23.248  < 2e-16 ***
## time_taken     0.0004357  0.0002879   1.514    0.131
## Task_diff_num  0.0005238  0.0039811   0.132    0.895
## AI_trust_num  -0.0005769  0.0039720  -0.145    0.885
## atn_ch        -0.0039077  0.0091791  -0.426    0.671
## log(age)      -0.0176953  0.0135581  -1.305    0.193
## male_num      -0.0109264  0.0080894  -1.351    0.178
```

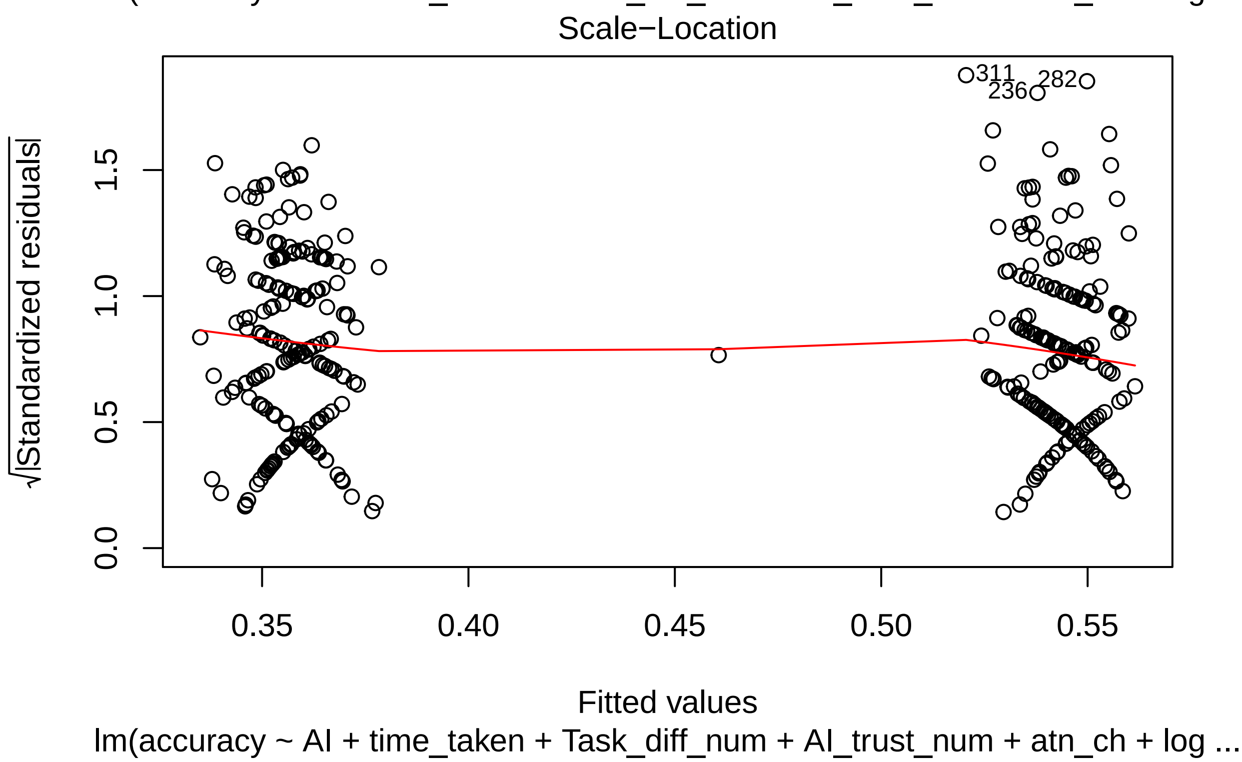
```
## college          0.0082197  0.0087534   0.939    0.348
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0793 on 389 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.5871, Adjusted R-squared:  0.5786
## F-statistic: 69.13 on 8 and 389 DF,  p-value: < 2.2e-16
```

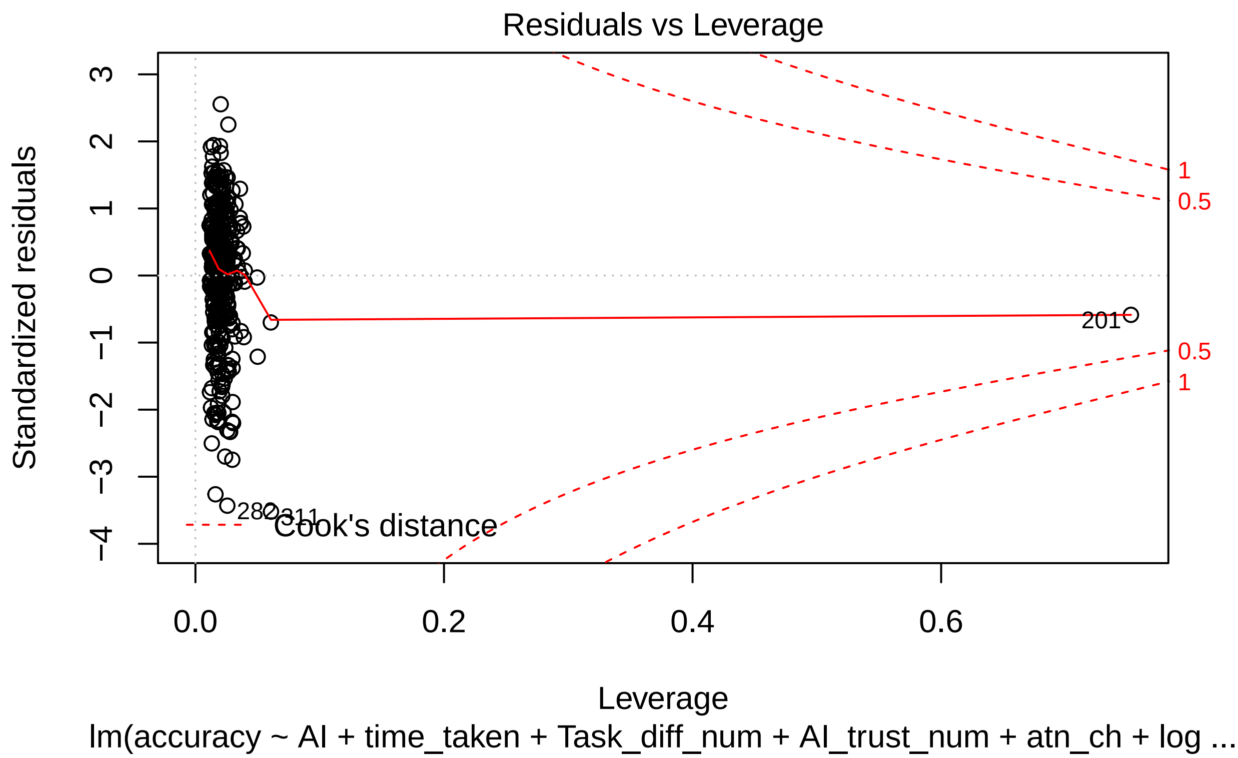
Jackknife plot and the Q-Q plot look acceptable. In the Q-Q plot, there is some deviation at tails but it is still a acceptable fit.

```
plot(lm.3.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + log ...
236
282
311


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + log ...
311
236
282



Effect of Uncertainty Information on accuracy with other predictor variables

Just like the simple linear regression (lm.2.acc), provision of uncertainty information is positively and significantly affecting the participants accuracy. Along with that, perceived AI usefulness rating and Task difficulty rating are also significant. $F(9, 189) = 6.39$, $p < 0.001$, $R\hat{}2 = 0.20$.

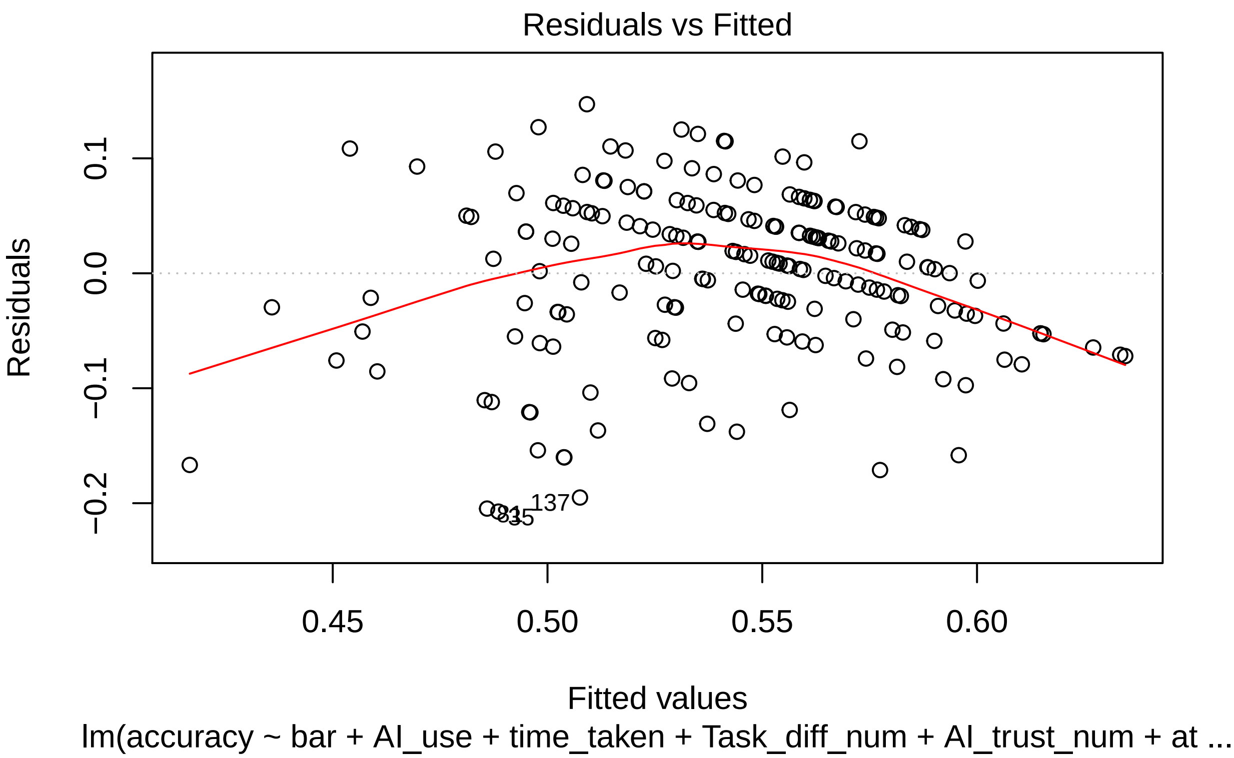
```
lm.4.acc <- lm(accuracy ~ bar + AI_use + time_taken + Task_diff_num +
                 AI_trust_num + atn_ch + log(age) + male_num + college,
               data = person_AI)

summary(lm.4.acc)
```

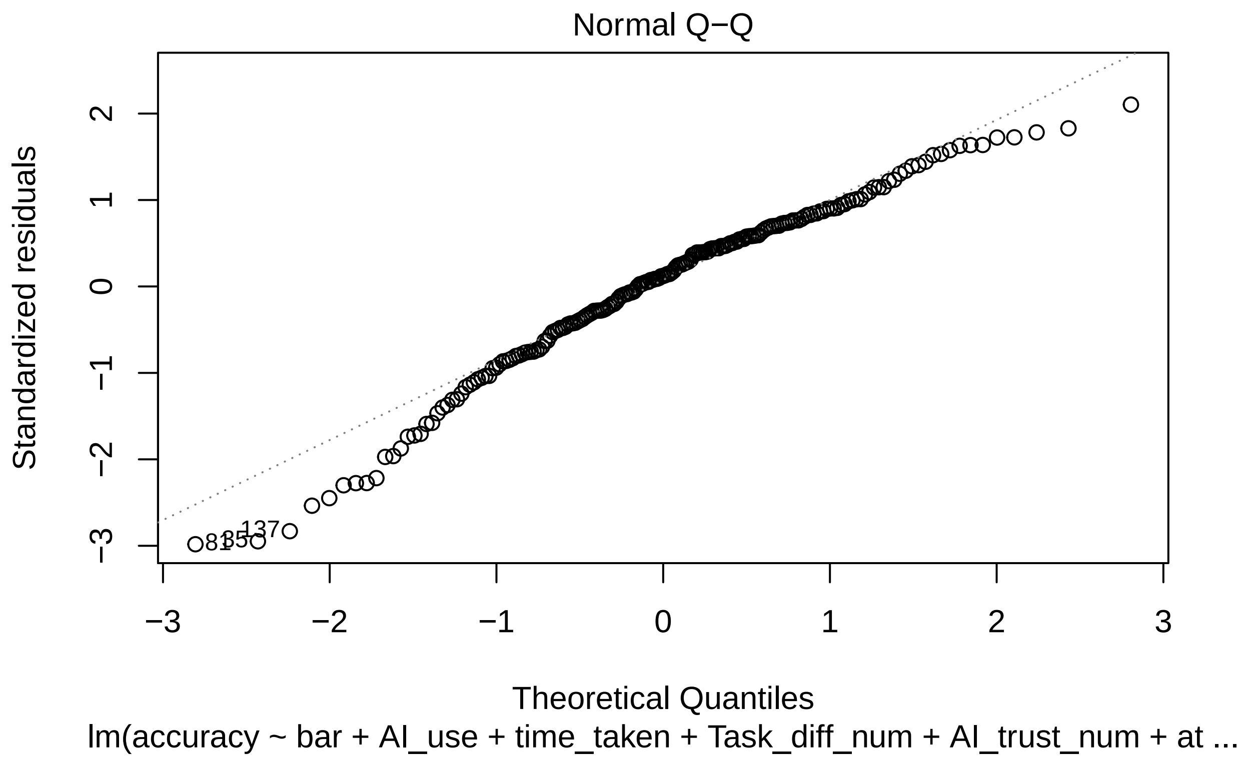
```
##
## Call:
## lm(formula = accuracy ~ bar + AI_use + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = person_AI)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.207322 -0.038506  0.008474  0.049019  0.147081
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.4255900  0.0654040   6.507 6.70e-10 ***
## bar             0.0607531  0.0106068   5.728 3.95e-08 ***
## AI_use          0.1342157  0.0340760   3.939 0.000115 ***
## time_taken     -0.0004790  0.0009214  -0.520 0.603774
## Task_diff_num   0.0122880  0.0050806   2.419 0.016528 *
## AI_trust_num   -0.0059819  0.0055307  -1.082 0.280815
## atn_ch          0.0108391  0.0121541   0.892 0.373631
```

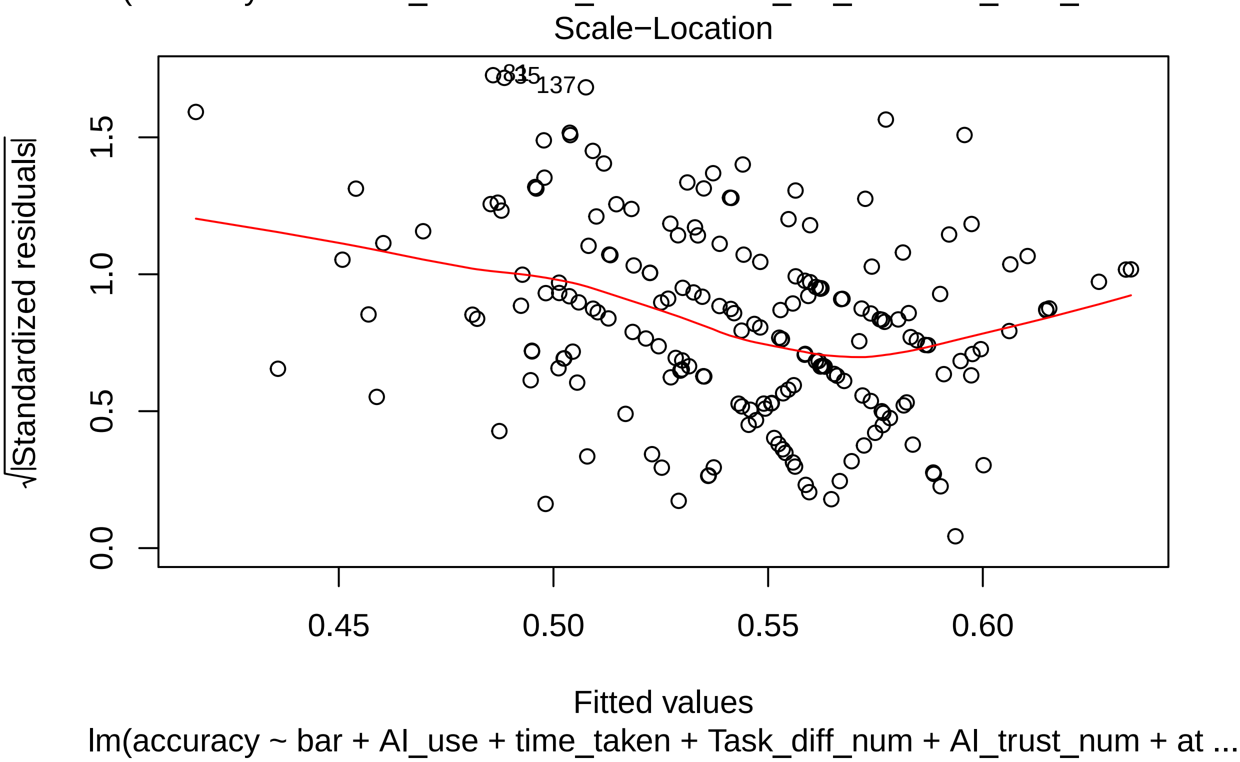
```
## log(age)       -0.0015355  0.0174070  -0.088 0.929804
## male_num       -0.0161267  0.0103727  -1.555 0.121684
## college         0.0076004  0.0112985   0.673 0.501963
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07157 on 189 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.2333, Adjusted R-squared:  0.1968
## F-statistic:  6.39 on 9 and 189 DF,  p-value: 6.639e-08
```

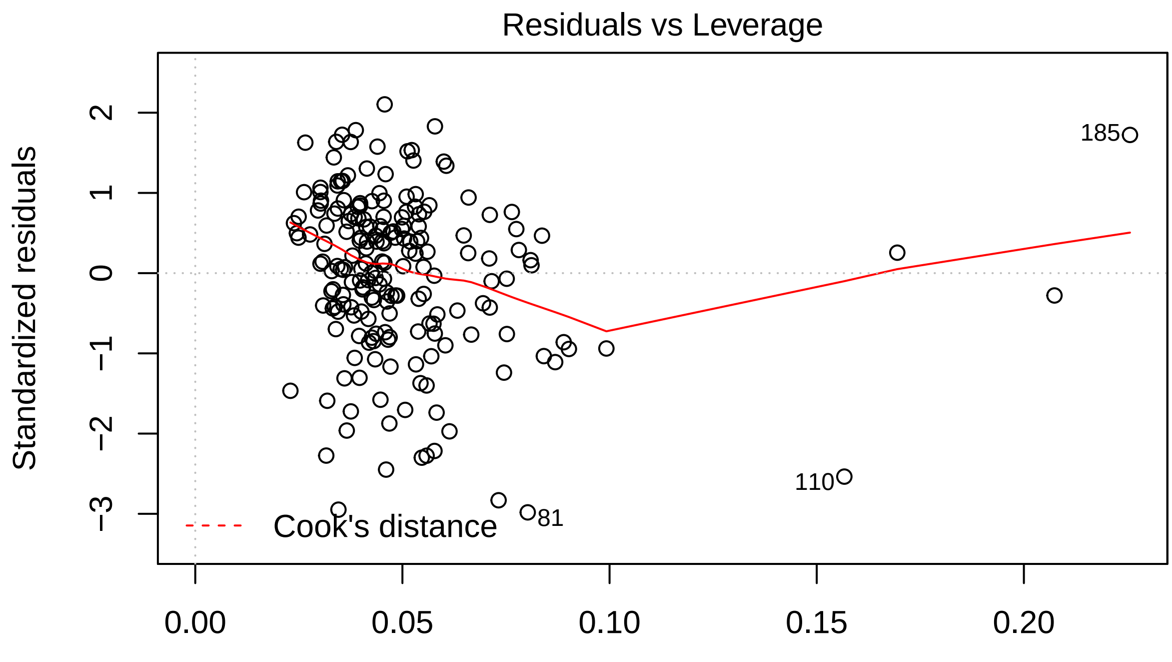
Jackknife and Q-Q plot - when you get to extremes, the model doesnt work well.

```
plot(lm.4.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + at ...
137
81
35


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + at ...
81
35
137



lm(accuracy ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + at ...

Effect of AI recommendations on accuracy with animal domain knowledge

In this model, to test the domain knowledge's effect on accuracy, the data is seperated to include animals stimuli only.

Like in the previous models, AI recommendations is significant. The perceived domain knowledge of animals of the participants is also significant. Domain knowledge affects the accuracy of the participants positively. The interaction between AI recommendations and domain knowledge however was not significant. F(10, 387) = 17.29, p <0.001, R^2 = .29

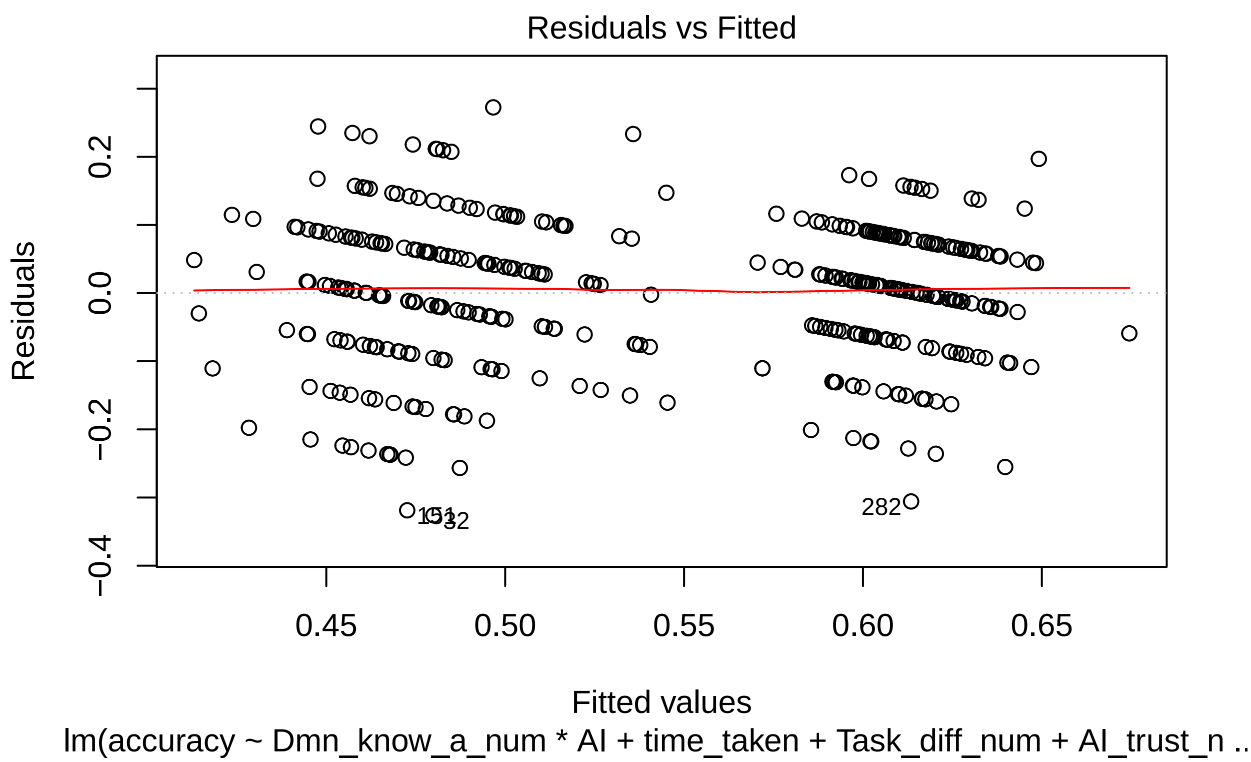
```
lm.5.a.acc <- lm(accuracy ~ Dmn_know_a_num*AI +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college , data = animals_person)

summary(lm.5.a.acc)
```

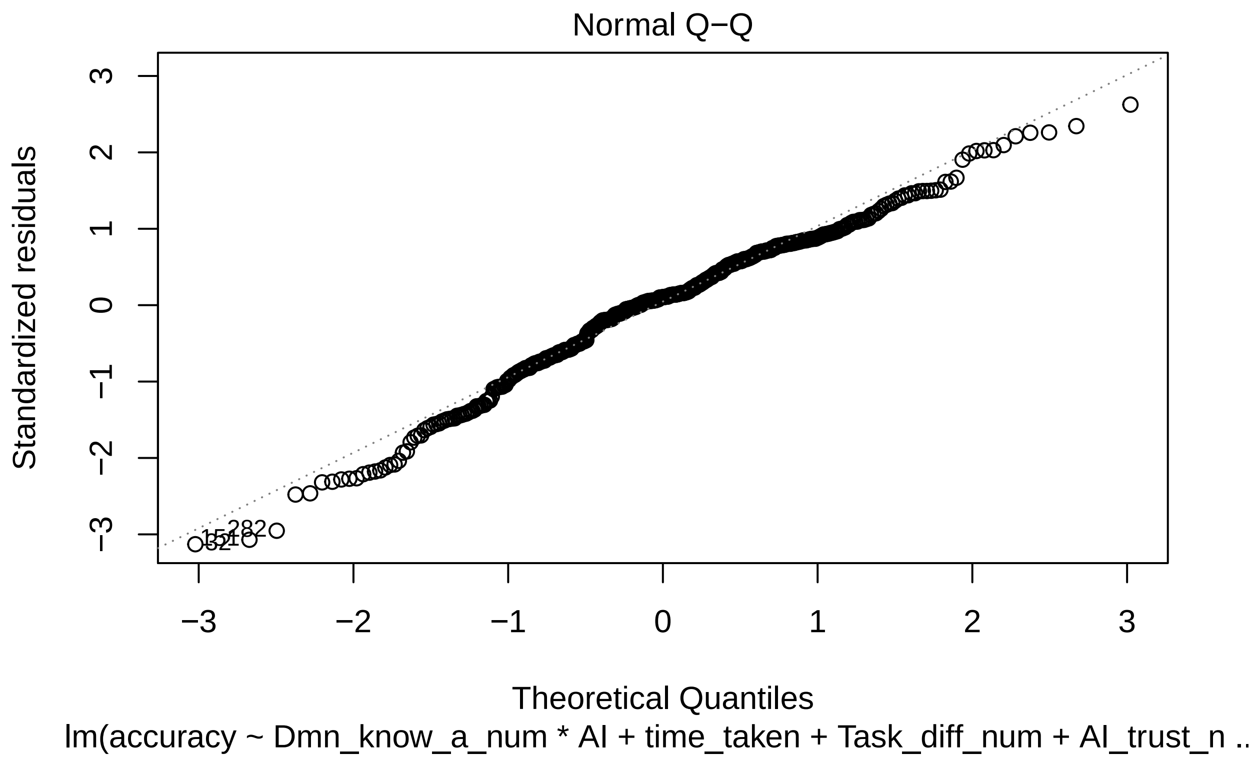
```
##
## Call:
## lm(formula = accuracy ~ Dmn_know_a_num * AI + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = animals_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.32610 -0.06463  0.01093  0.07446  0.27257
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.729e-01  6.818e-02   5.469 8.15e-08 ***
## Dmn_know_a_num     1.017e-01  3.219e-02   3.160   0.0017 **
## AI                 1.575e-01  2.396e-02   6.576 1.57e-10 ***
## time_taken         2.855e-05  1.733e-04   0.165   0.8692
```

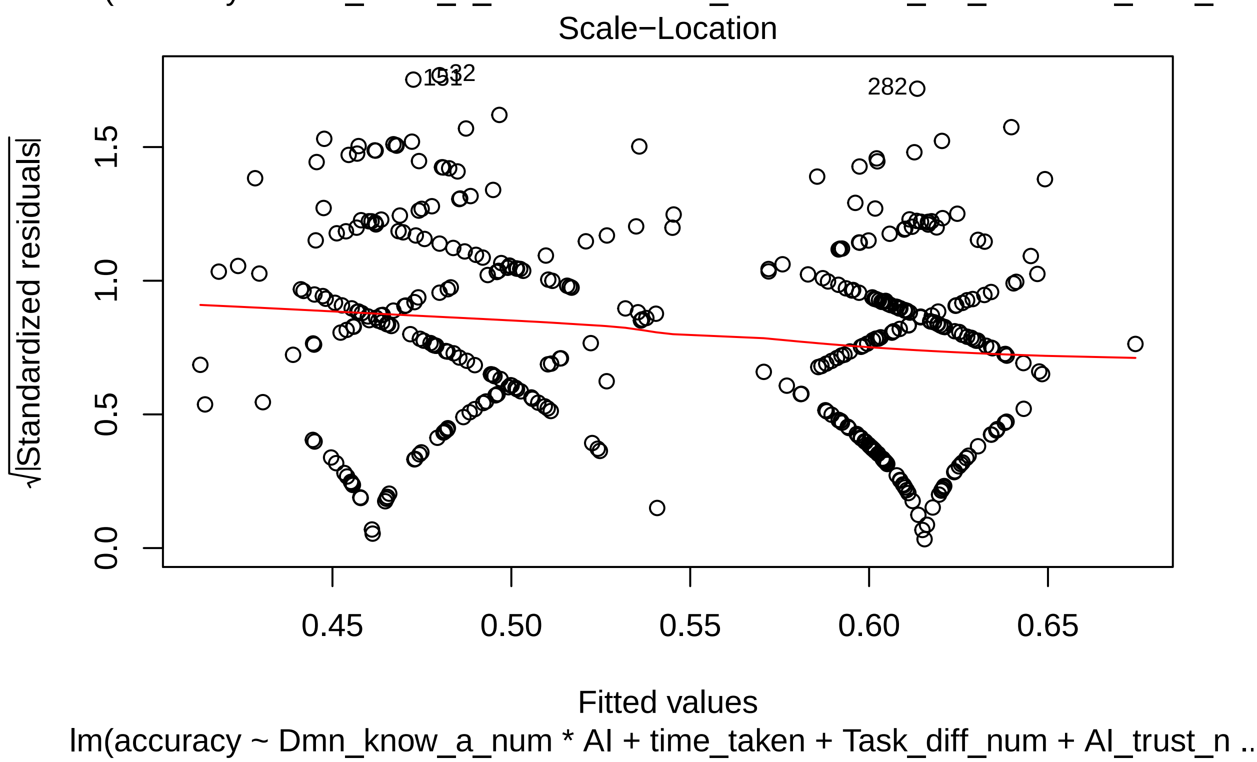
```
## Task_diff_num      -1.420e-03  5.321e-03  -0.267   0.7898
## AI_trust_num       -2.584e-03  5.263e-03  -0.491   0.6237
## atn_ch              2.143e-04  1.237e-02   0.017   0.9862
## log(age)            2.505e-02  1.791e-02   1.398   0.1629
## male_num           -1.668e-02  1.074e-02  -1.552   0.1214
## college            -1.156e-02  1.163e-02  -0.994   0.3209
## Dmn_know_a_num:AI -5.479e-02  4.493e-02  -1.219   0.2235
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1053 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.3088, Adjusted R-squared:  0.2909
## F-statistic: 17.29 on 10 and 387 DF,  p-value: < 2.2e-16
```

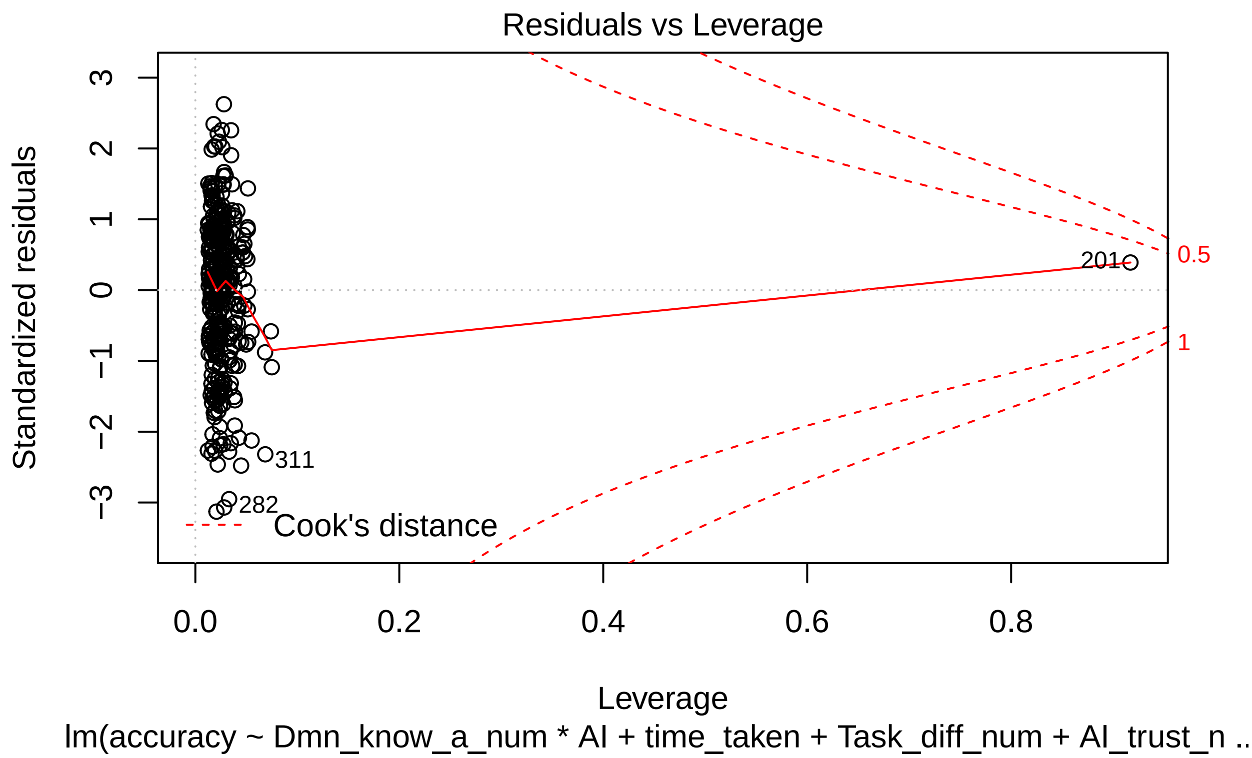
Jackknife plot and Q-Q plot are good fits.

```r
plot(lm.5.a.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust_n ..
151
32
282


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust_n ..
151
32
282



Effect of AI recommendations on accuracy with plant domain knowledge

Similar to previous models, AI recommendations is significant. On top of that, other predictor variables like Log(age) and education level (college) are also significant. Mainly, participants' perceived plant's domain knowledge is also significant; affecting their accuracy positively. The interaction between AI recommendations and domain knowledge however was not significant. F(10, 387) = 55.28, p < 0.001, R^2 = 0.58

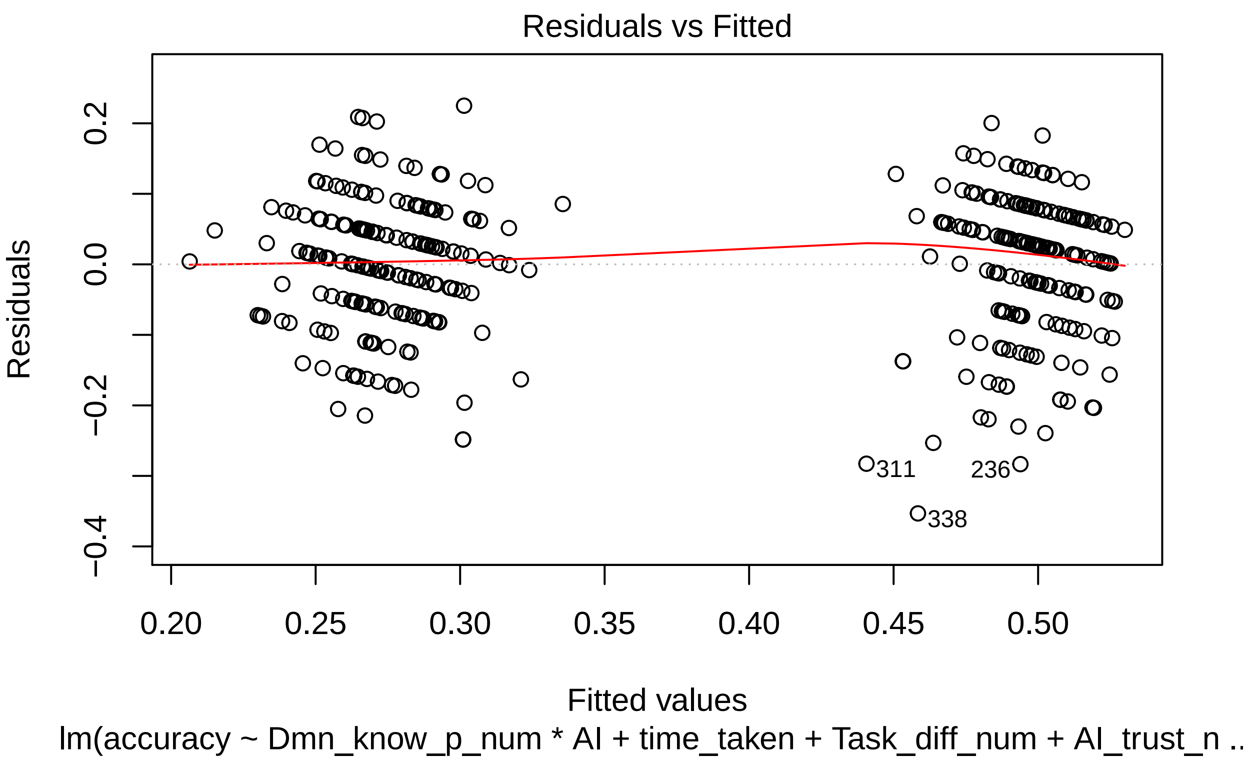
```
lm.5.p.acc <- lm(accuracy ~ Dmn_know_p_num*AI +
                 time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                 male_num + college, data = plants_person)

summary(lm.5.p.acc)
```

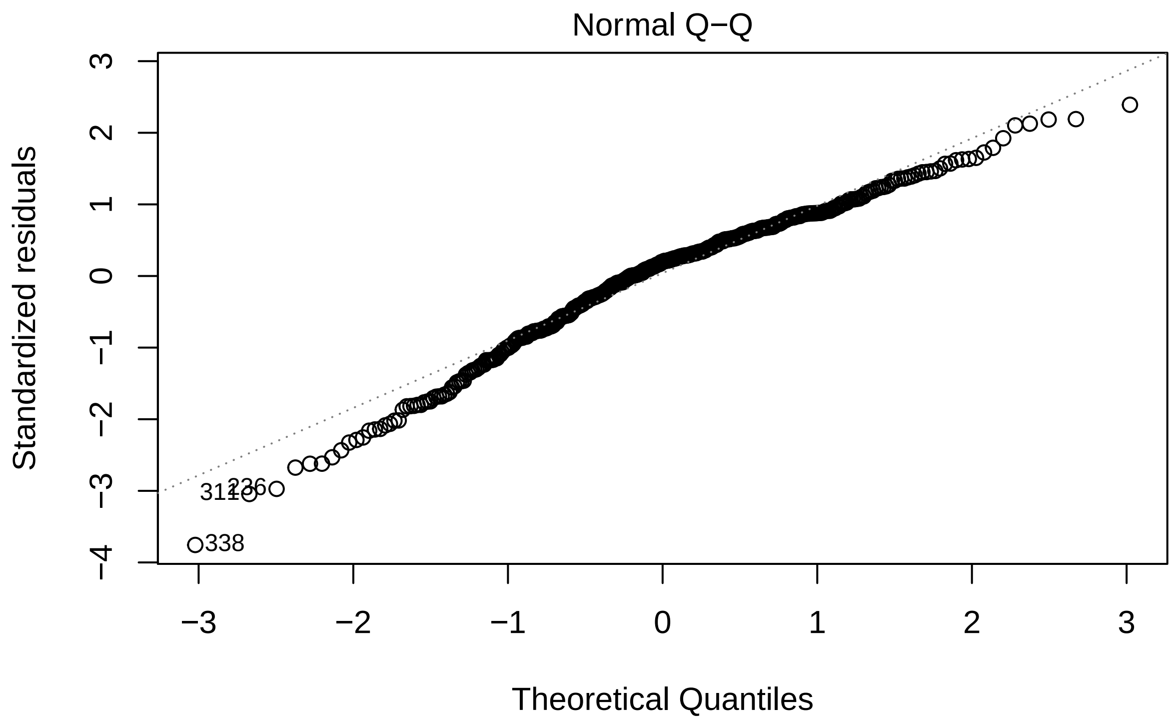
```
##
## Call:
## lm(formula = accuracy ~ Dmn_know_p_num * AI + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = plants_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35316 -0.05664  0.01847  0.06396  0.22495
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.3848568  0.0595793   6.460 3.16e-10 ***
## Dmn_know_p_num    0.0715952  0.0336734   2.126  0.03412 *
## AI                0.2365433  0.0131086  18.045  < 2e-16 ***
## time_taken        0.0004082  0.0006552   0.623  0.53364
## Task_diff_num     0.0038072  0.0048954   0.778  0.43722
## AI_trust_num      0.0003091  0.0048837   0.063  0.94957
```

```
## atn_ch             0.0013111  0.0112721   0.116  0.90746
## log(age)          -0.0452588  0.0165004  -2.743  0.00637 **
## male_num          -0.0068940  0.0097912  -0.704  0.48179
## college            0.0243808  0.0106988   2.279  0.02322 *
## Dmn_know_p_num:AI -0.0608204  0.0456199  -1.333  0.18325
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09611 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.5882, Adjusted R-squared:  0.5776
## F-statistic: 55.28 on 10 and 387 DF,  p-value: < 2.2e-16
```

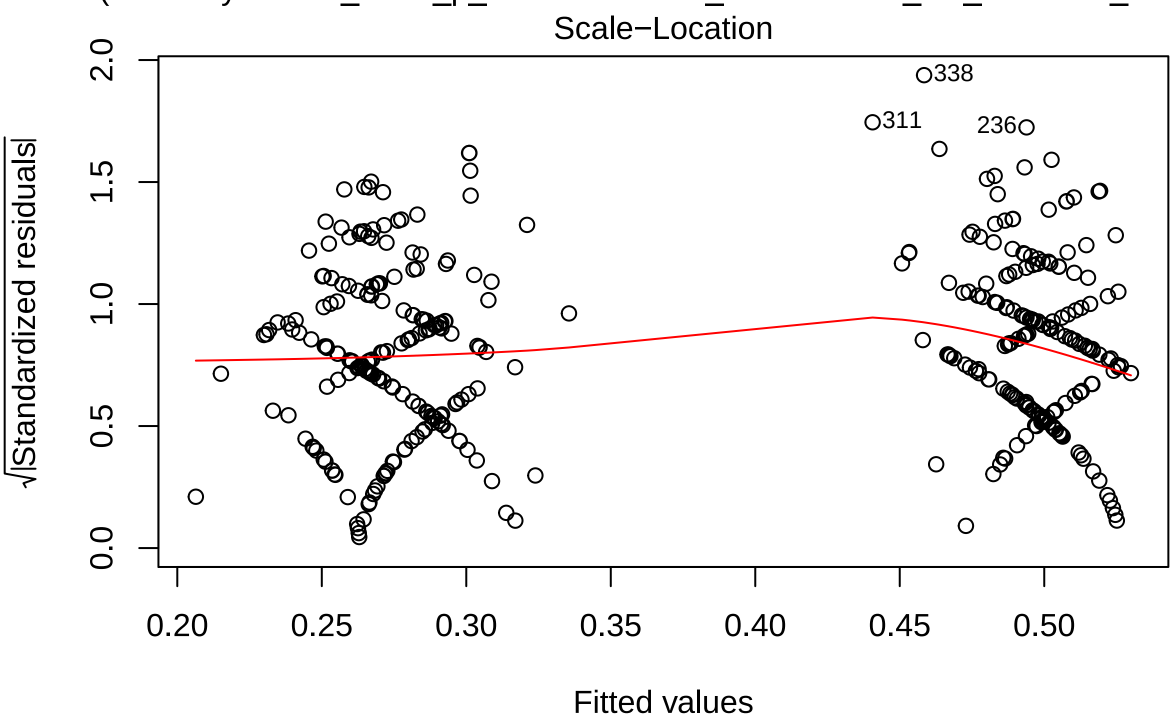
Jackknife plot and Q-Q plot are good fits.

```r
plot(lm.5.p.acc)
```

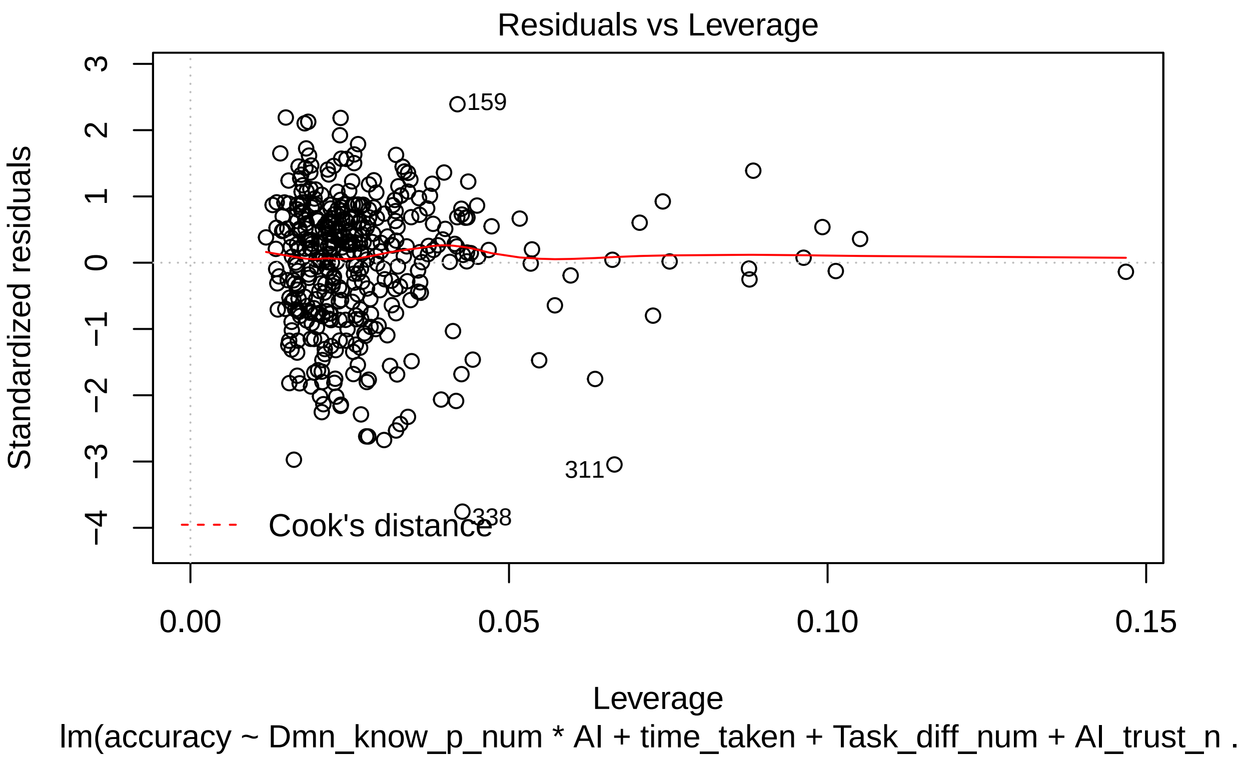




lm(accuracy ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust_n ..



lm(accuracy ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust_n ..



Effect of Uncertainty Information on accuracy with animal domain knowledge

Contrasting to previous models, provision of uncertainty information was not significantly affecting the accuracy of the participants. Domain knowledge was also not significant; neither was the interaction between domain knowlede and Uncertainty informtation. Only perceived AI usefulness rating was significant which positively affected the accuracy.

F(11, 187) = 2.89, p <0.01, R^2 = 0.10

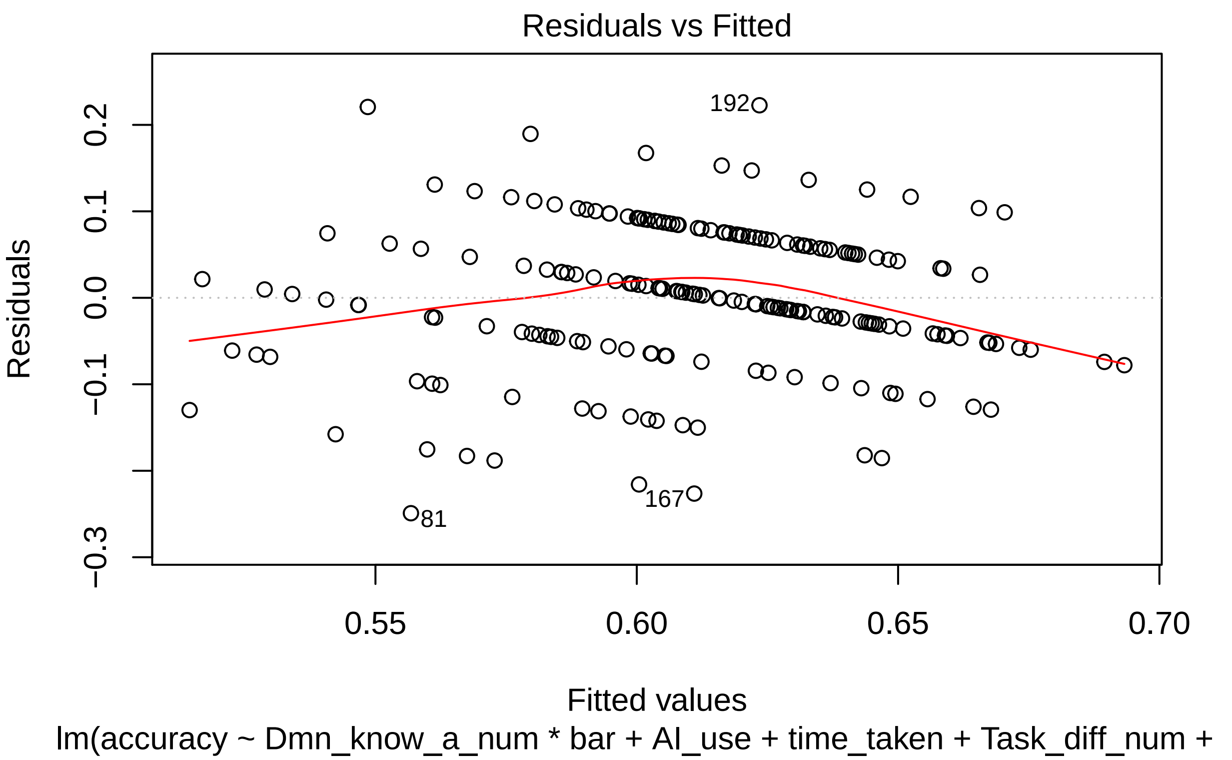
```
lm.6.a.acc <- lm(accuracy ~ Dmn_know_a_num*bar + AI_use +
                 time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                 male_num + college , data = animals_person_AI)

summary(lm.6.a.acc)
```

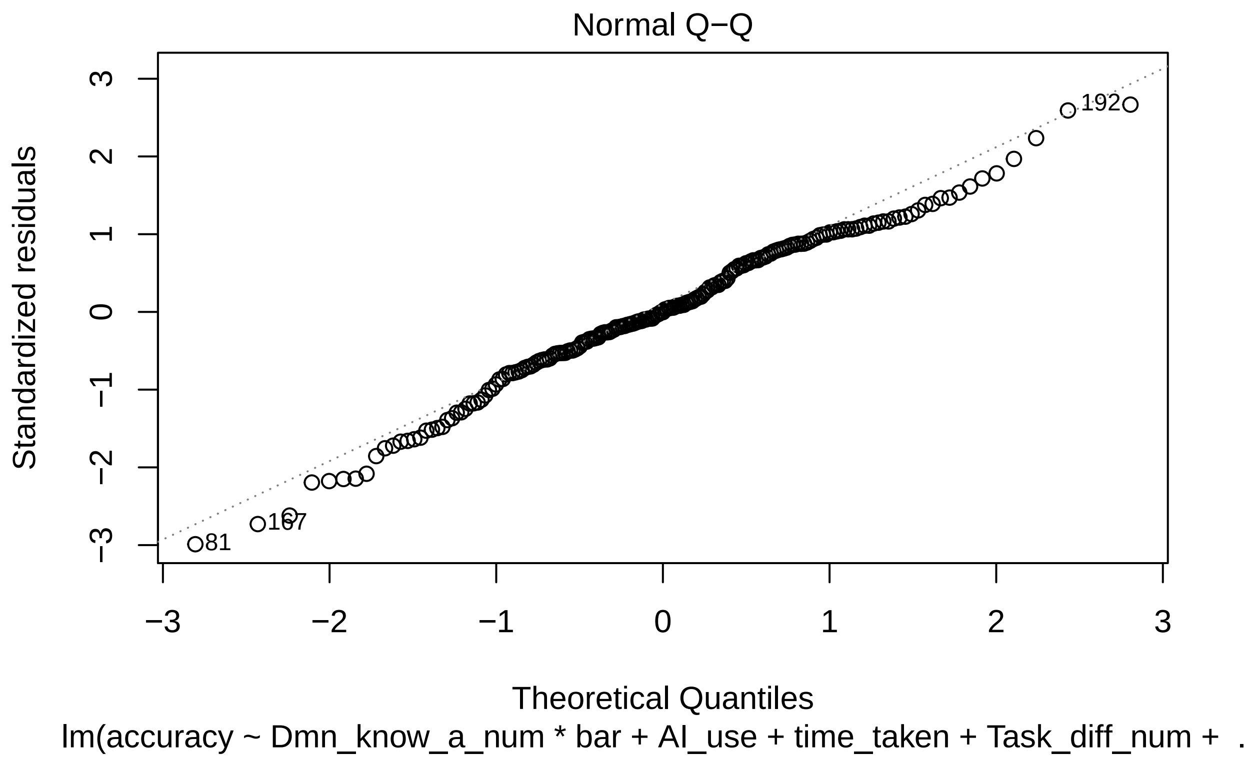
```
##
## Call:
## lm(formula = accuracy ~ Dmn_know_a_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = animals_person_AI)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.249096 -0.048340 -0.000305  0.065000  0.222674
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.417082   0.082363   5.064  9.8e-07 ***
## Dmn_know_a_num      0.054847   0.037563   1.460 0.145937
## bar                 0.036457   0.029213   1.248 0.213606
## AI_use              0.143926   0.039176   3.674 0.000312 ***
```

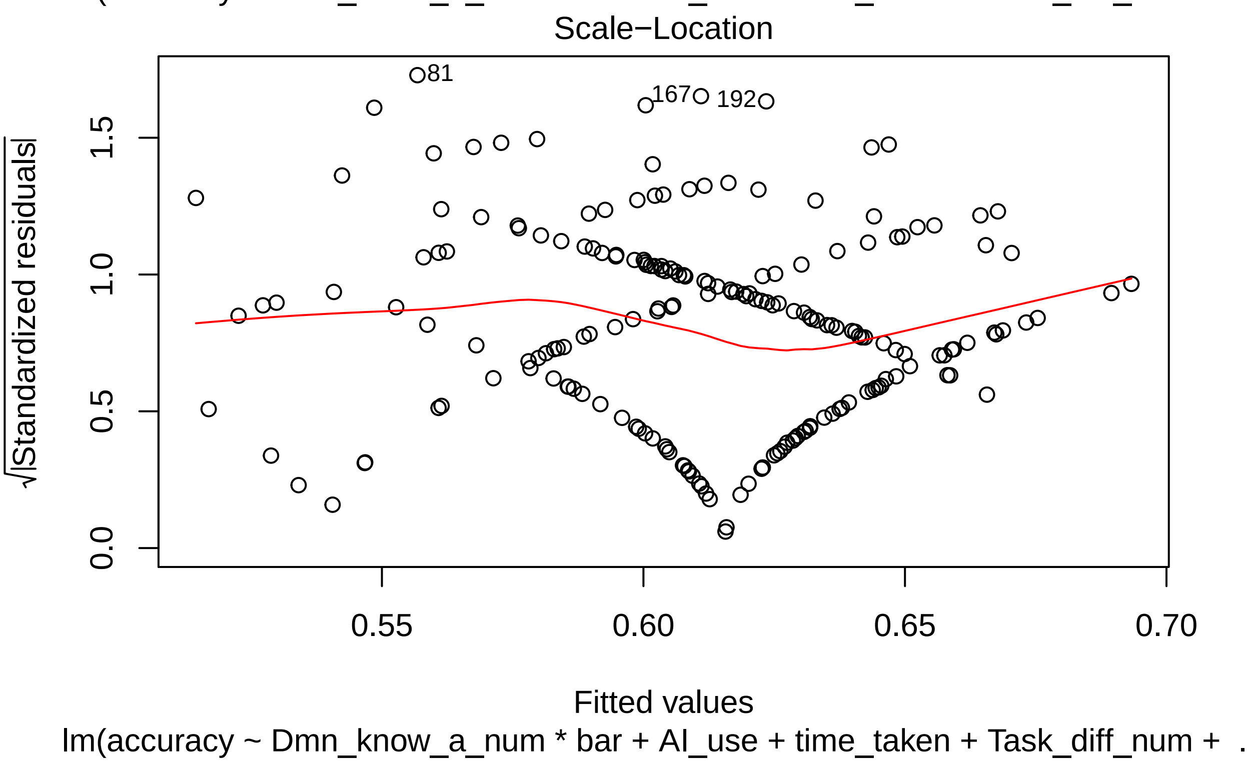
```
## time_taken          0.001090   0.001262   0.863 0.389057
## Task_diff_num       0.007095   0.006254   1.135 0.258019
## AI_trust_num       -0.008728   0.006605  -1.322 0.187947
## atn_ch              0.004336   0.015082   0.287 0.774067
## log(age)            0.023461   0.021362   1.098 0.273504
## male_num           -0.020493   0.012669  -1.618 0.107450
## college            -0.020892   0.013919  -1.501 0.135063
## Dmn_know_a_num:bar  0.003082   0.054409   0.057 0.954883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08739 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.1454, Adjusted R-squared:  0.09511
## F-statistic: 2.892 on 11 and 187 DF,  p-value: 0.001545
```

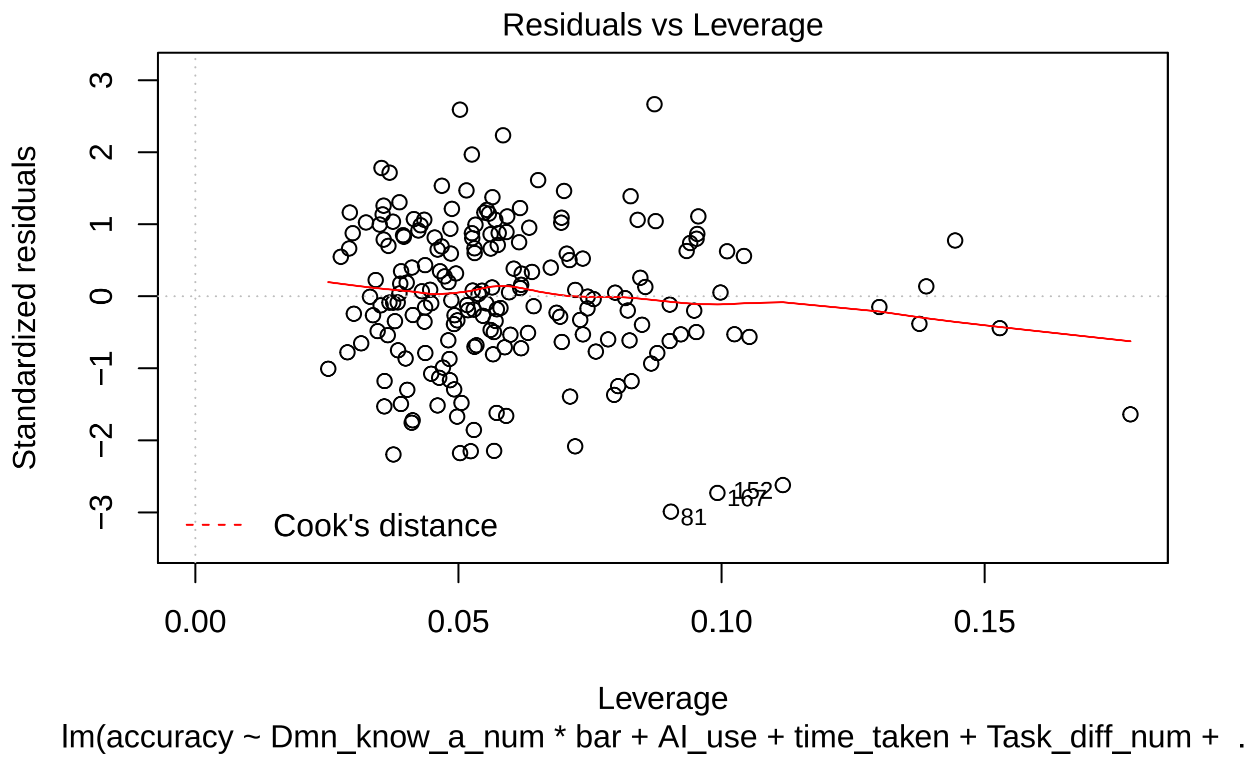
Jackknife and Q-Q plots look acceptable.

```
plot(lm.6.a.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num + .
192
167
81


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num + .
81
167
192



Effect of Uncertainty Information on accuracy with plants domain knowledge

Domain knowledge, the interaction between domain knowledge and uncertainty information were not significant. Uncertainty information, perceived AI usefulness rating, task difficulty rating, and college were all significant. F(11, 187) = 5.25, p < 0.001, R^2 = 0.19

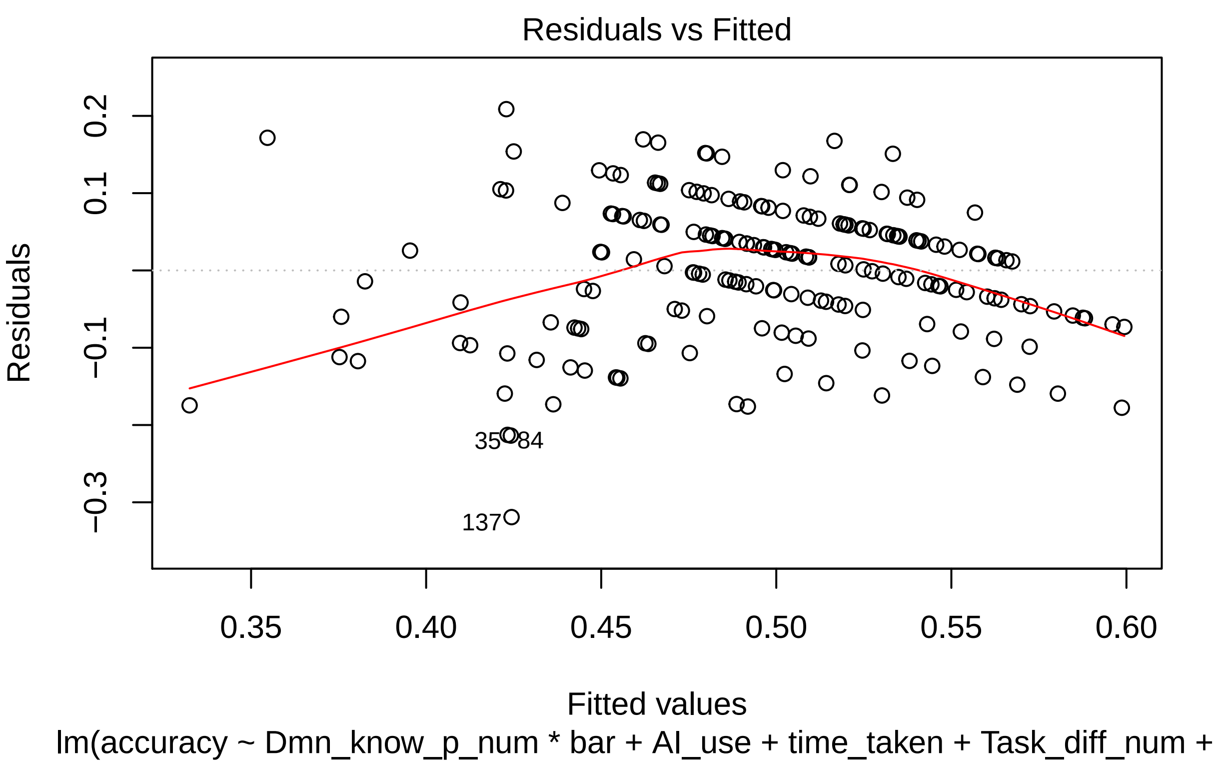
```
lm.6.p.acc <- lm(accuracy ~ Dmn_know_p_num*bar + AI_use +
                 time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                 male_num + college, data = plants_person_AI)

summary(lm.6.p.acc)
```

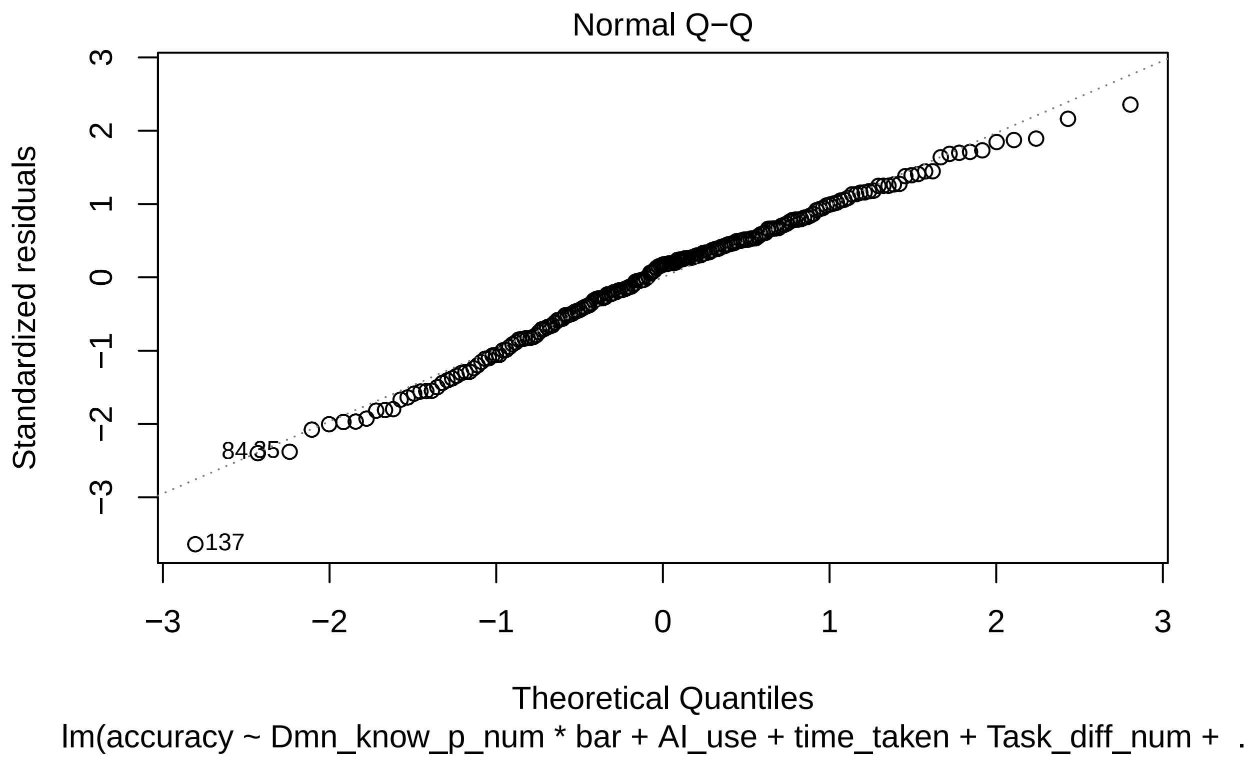
```
##
## Call:
## lm(formula = accuracy ~ Dmn_know_p_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = plants_person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.31914 -0.05877  0.01567  0.05931  0.20868
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.3983112  0.0839951   4.742 4.18e-06 ***
## Dmn_know_p_num      0.0365389  0.0415939   0.878  0.38082
## bar                 0.0810032  0.0186884   4.334 2.38e-05 ***
## AI_use              0.1097233  0.0433496   2.531  0.01219 *
## time_taken         -0.0010223  0.0009623  -1.062  0.28945
## Task_diff_num       0.0183507  0.0066458   2.761  0.00633 **
```

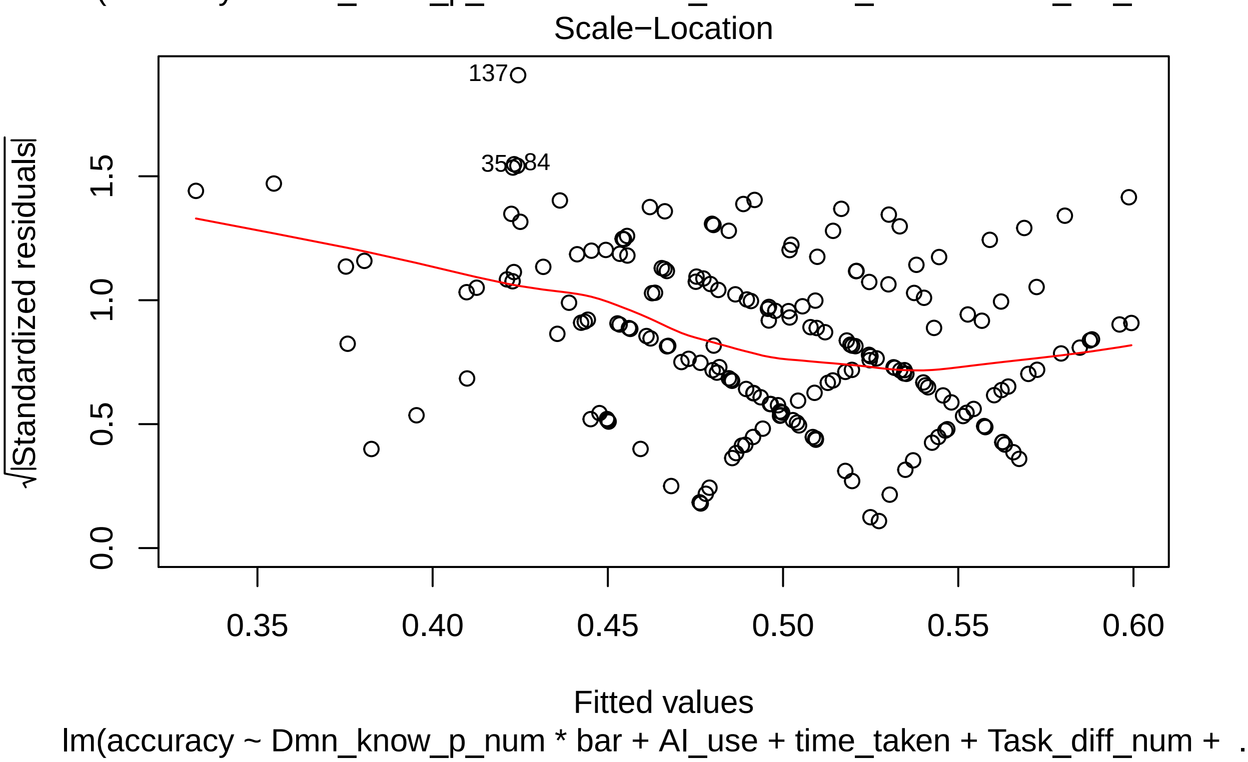
```
## AI_trust_num       -0.0034267  0.0071101  -0.482  0.63040
## atn_ch              0.0219377  0.0157692   1.391  0.16583
## log(age)           -0.0205709  0.0225588  -0.912  0.36301
## male_num           -0.0127921  0.0134050  -0.954  0.34117
## college             0.0293405  0.0145625   2.015  0.04536 *
## Dmn_know_p_num:bar  0.0060319  0.0684392   0.088  0.92986
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09144 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.2361, Adjusted R-squared:  0.1912
## F-statistic: 5.254 on 11 and 187 DF,  p-value: 3.299e-07
```

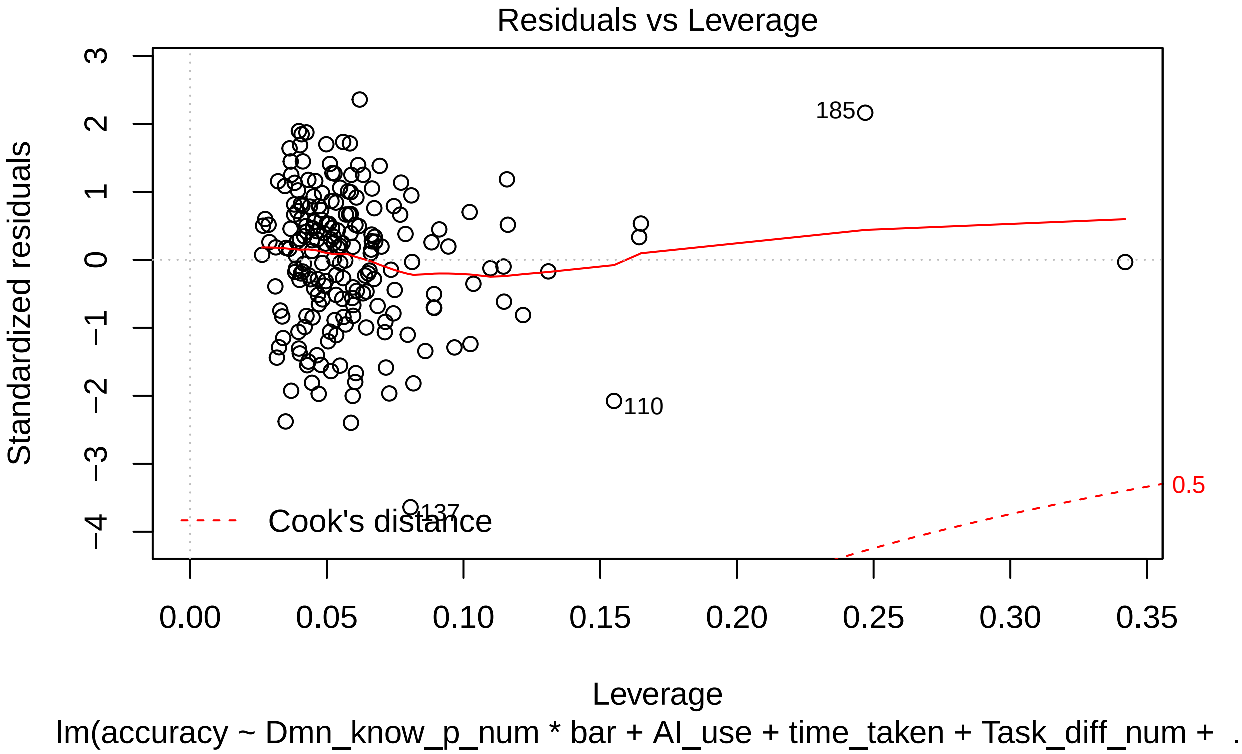
Q-Q plot is a great fit, jackknife plot looks acceptable.

```
plot(lm.6.p.acc)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(accuracy ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num + .
137
84
35


Scale−Location
√|Standardized residuals|
Fitted values
lm(accuracy ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num + .
137
35
84



END OF ANALYSIS WITH ACCURACY AS THE RESPONSE VARIABLE

CONFIDENCE ANALYSIS

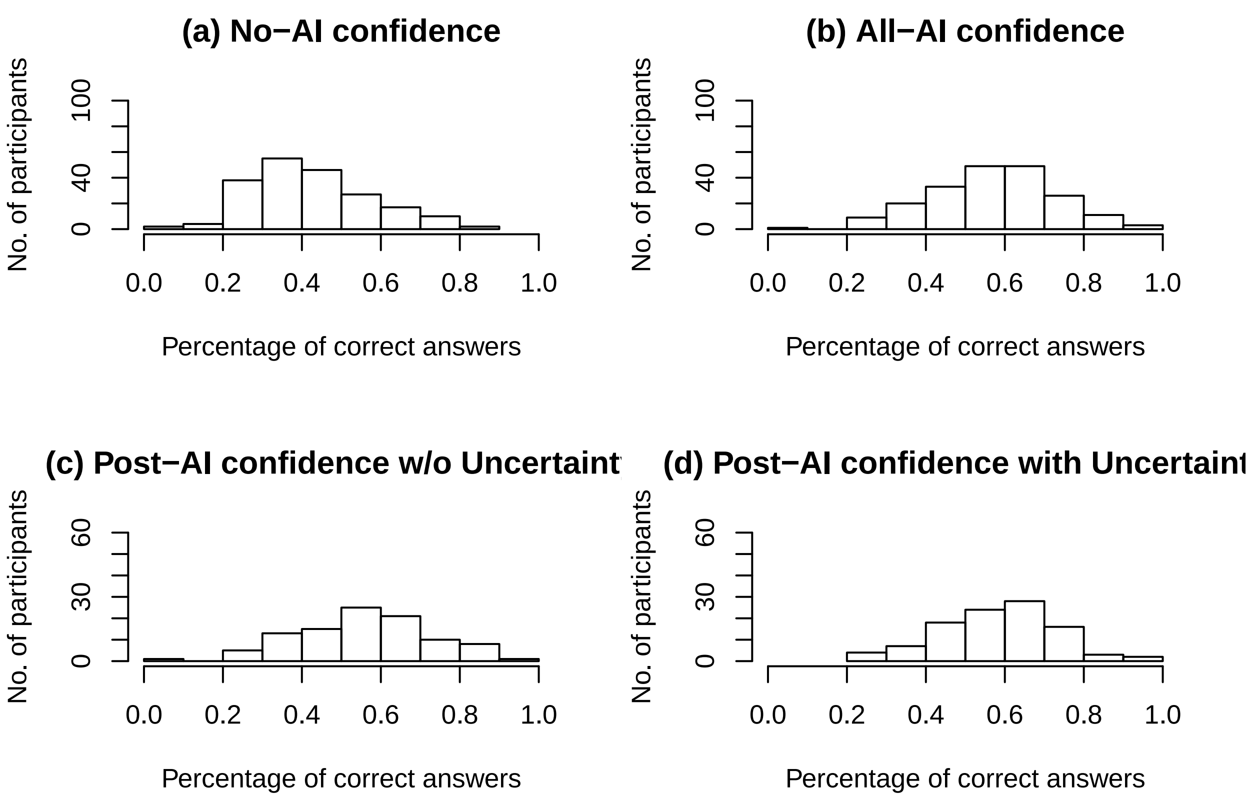
EXPLORATORY PLOTS

confidence Plots

The average confidence of the participants was 0.43 (SD = 0.15), when no AI recommendations were provided. In comparison, when AI recommendations were provided, the average confidence was 0.57(SD = 0.16). The means and the plots clearly indicate the positive relationship between AI recommendations and confidence of the participants.

The average confidence of the participants was 0.56 (SD = 0.16), when uncertainty information was not provided. In comparison, when the uncertainty information were provided, the average confidence was 0.58 (SD = 0.15). The means and the plots show that confidence of the participants increase slightly when uncertainty information is provided.

```
par(mfrow=c(2,2))

#confidence percentage for baseline - No AI condition
hist(person_noAI$confidence, #choosing column in a dataset
     main = "(a) No-AI confidence", #main plot label
     xlab = "Percentage of correct answers", #x-axis label
     ylab = "No. of participants", #y-axis label
     ylim = c(0,100), xlim = c(0,1)) #limits for x- & y-axis in the plot

#confidence percentage in all AI Condition
hist(person_AI$confidence, main = "(b) All-AI confidence",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,100), xlim = c(0,1))

#confidence percentage in AI Condition without bars
hist(person_nobar$confidence,
     main = "(c) Post-AI confidence w/o Uncertainty",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,60), xlim = c(0,1))

#confidence percentage in AI with Uncertianty Information (AI_bars)
hist(person_bar$confidence, main ="(d) Post-AI confidence with Uncertainty",
     xlab = "Percentage of correct answers", ylab = "No. of participants",
     ylim = c(0,60), xlim = c(0,1))
```




```
par(mfrow=c(1,1))
```

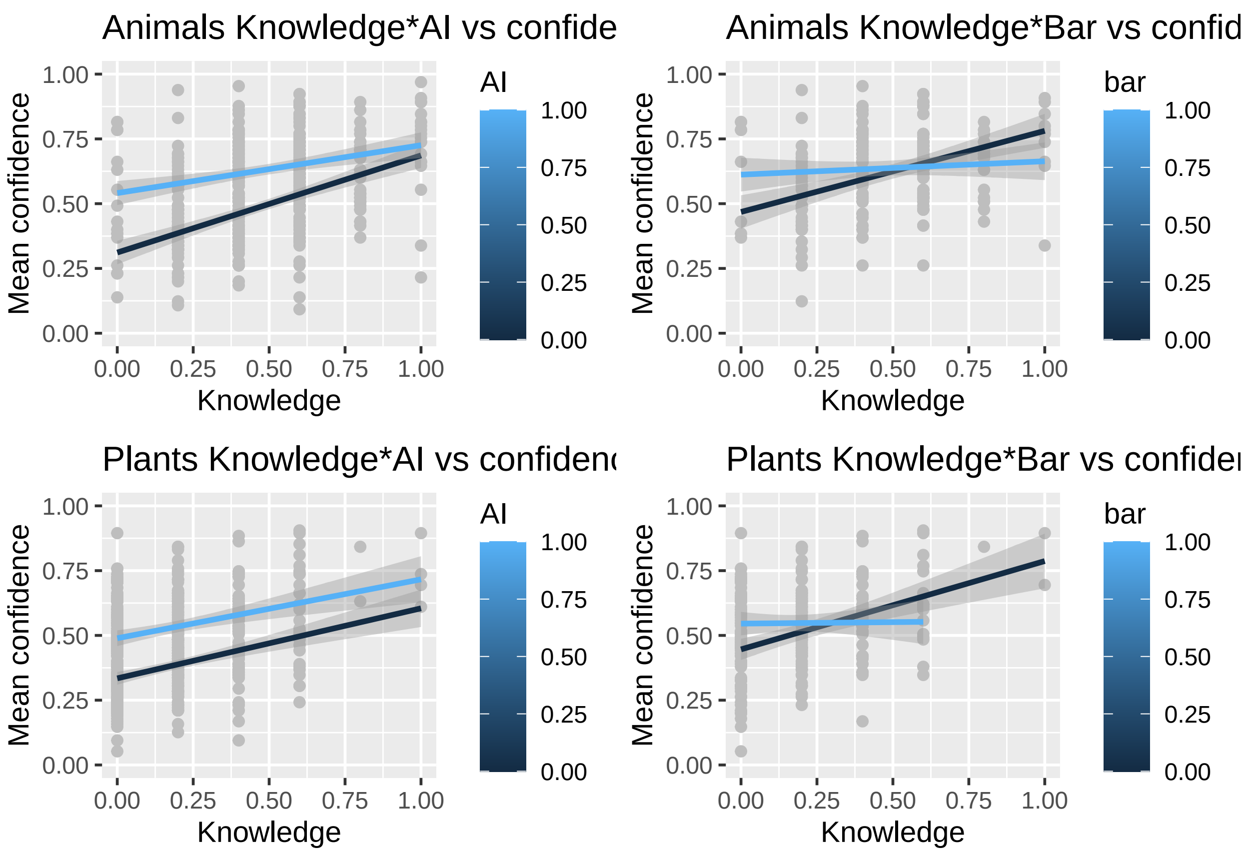
Domain Knowledge interaction with AI and Bar vs. confidence

Domain Knoelwedge interaction with AI and Bar for animals

```
#Filtering for AI vs No-AI
animals_person_noAI <- filter(animals_person, AI == 0)
animals_person_AI <- filter(animals_person, AI == 1)

#Filtering for Bar vs No-bar
animals_person_bar <- filter(animals_person_AI, bar == 1)
animals_person_nobar <- filter(animals_person_AI, bar == 0)

animal_dmn_AI_plot <- ggplot(animals_person) +
  aes(x = Dmn_know_a_num, y = confidence, color = AI) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = animals_person_noAI) +
  geom_smooth(method = "lm", data = animals_person_AI) +
  xlab("Knowledge") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Animals Knowledge*AI vs confidence") #main plot title

animal_dmn_bar_plot <- ggplot(animals_person_AI) +
  aes(x = Dmn_know_a_num, y = confidence, color = bar) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = animals_person_nobar) +
  geom_smooth(method = "lm", data = animals_person_bar) +
  xlab("Knowledge") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Animals Knowledge*Bar vs confidence") #main plot title

#Filtering for AI vs No-AI
plants_person_noAI <- filter(plants_person, AI == 0)
plants_person_AI <- filter(plants_person, AI == 1)

#Filtering for Bar vs No-bar
plants_person_bar <- filter(plants_person_AI, bar == 1)
plants_person_nobar <- filter(plants_person_AI, bar == 0)

plant_dmn_AI_plot <- ggplot(plants_person) +
  aes(x = Dmn_know_p_num, y = confidence, color = AI) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = plants_person_noAI) +
  geom_smooth(method = "lm", data = plants_person_AI) +
  xlab("Knowledge") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Plants Knowledge*AI vs confidence") #main plot title

plant_dmn_bar_plot <- ggplot(plants_person_AI) +
  aes(x = Dmn_know_p_num, y = confidence, color = bar) +
  geom_point(color = "grey") +
  geom_smooth(method = "lm", data = plants_person_nobar) +
  geom_smooth(method = "lm", data = plants_person_bar) +
```

```
  xlab("Knowledge") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Plants Knowledge*Bar vs confidence") #main plot title

cowplot::plot_grid(animal_dmn_AI_plot, animal_dmn_bar_plot, plant_dmn_AI_plot, plant_dmn_bar_plot)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```

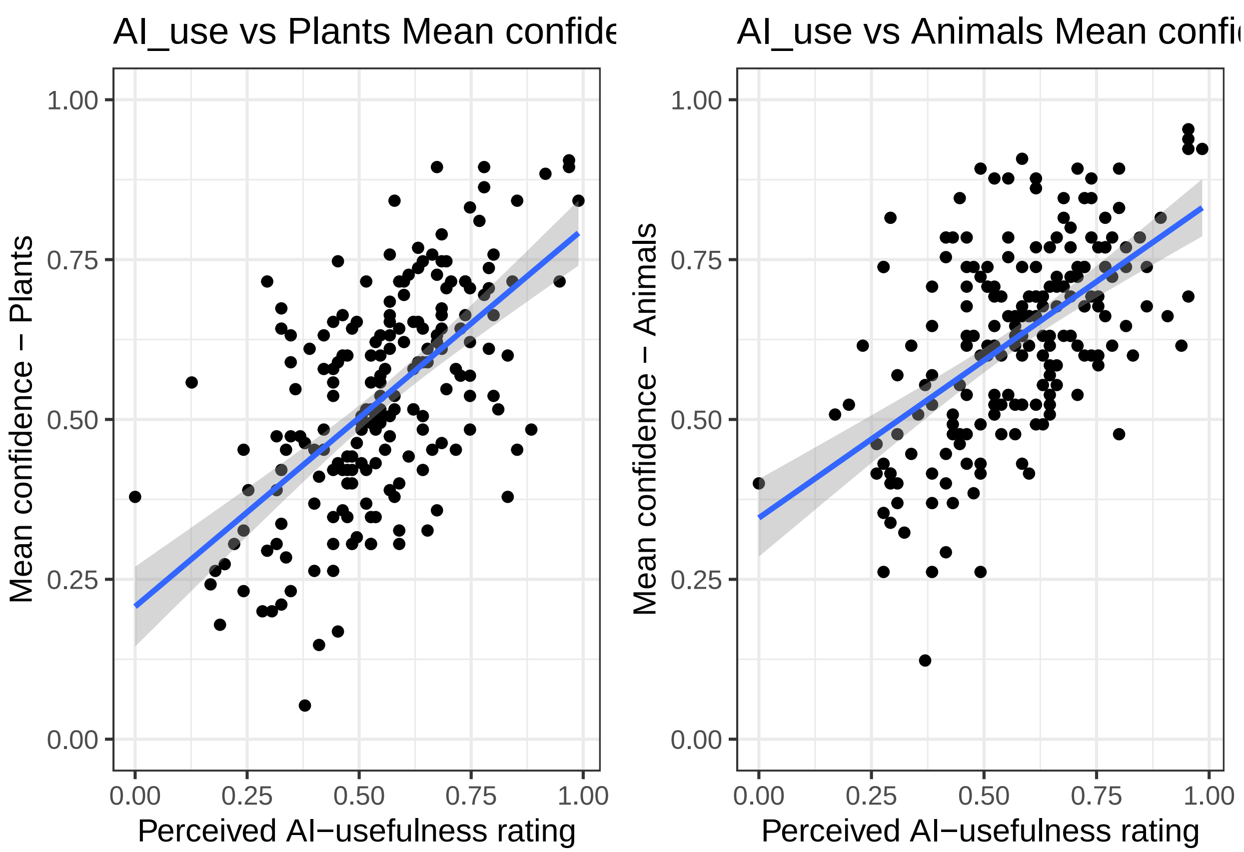


AI-usefulness rating vs confidence

The mean confidence of the partcipants vs perceived AI-usefulness rating does not show a linear relationship. The fitted line is almost horizontal. The data points also do not indicate any relationship. But the analysis is split between plants and animals images. When looking at separate datasets, it is clear that there exists a linear relationship between AI usefulness ratings and confidence.

```
# AI-use vs. confidence - all AI- conditions
AI_use_plot.1 <- filter(plants_person, AI ==1) %>%
  ggplot(aes(x=AI_use, y = confidence)) +
  geom_point() +
```

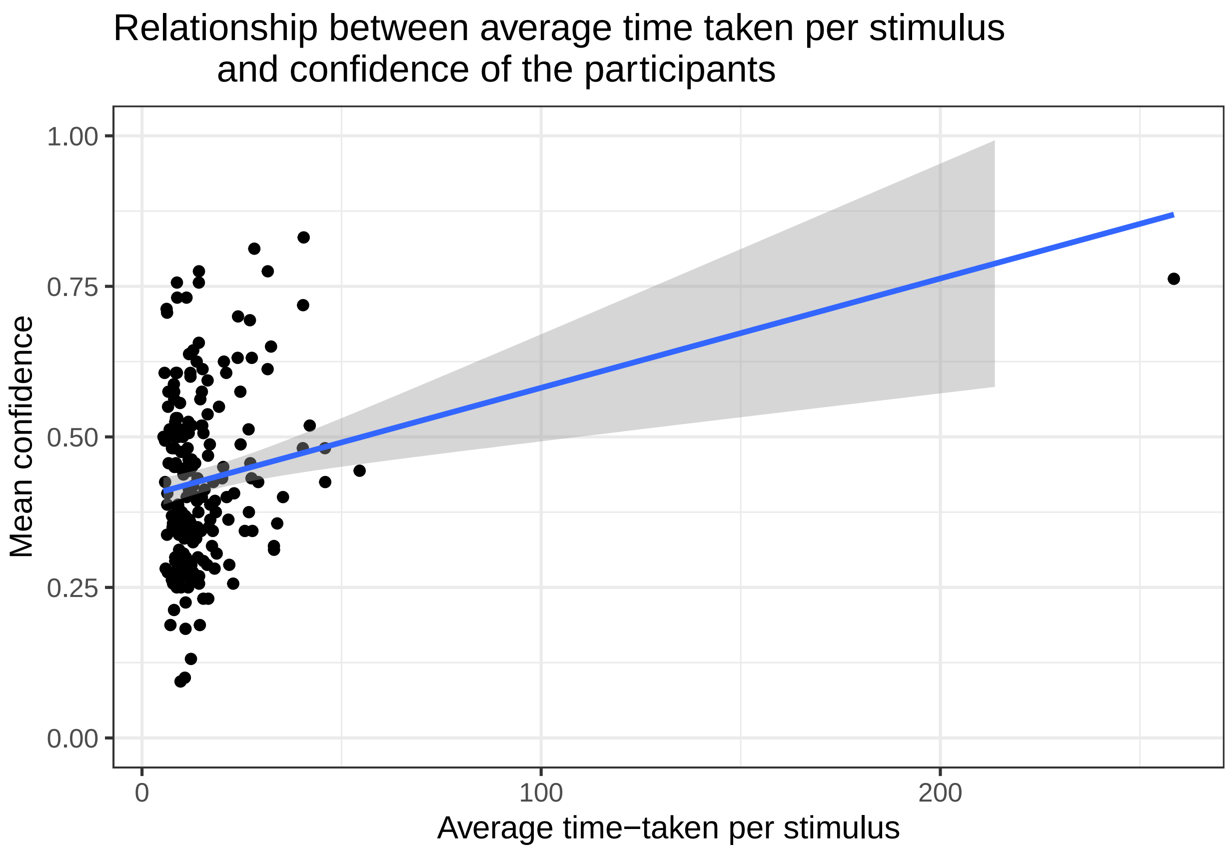
```
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI-usefulness rating") +
  ylab("Mean confidence - Plants") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI_use vs Plants Mean confidence") #main plot title

# AI-use vs. confidence - all AI- conditions
AI_use_plot.2 <- filter(animals_person, AI ==1) %>%
  ggplot(aes(x=AI_use, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI-usefulness rating") +
  ylab("Mean confidence - Animals") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI_use vs Animals Mean confidence") #main plot title


cowplot::plot_grid(AI_use_plot.1, AI_use_plot.2)
```



Time taken vs confidence

There is a clear outlier that is affecting the fit of the data. To find a relationship between these two variables, the outlier will be removed. During the analysis, linear models will be fit with and without the outlier to see its effect on the results.

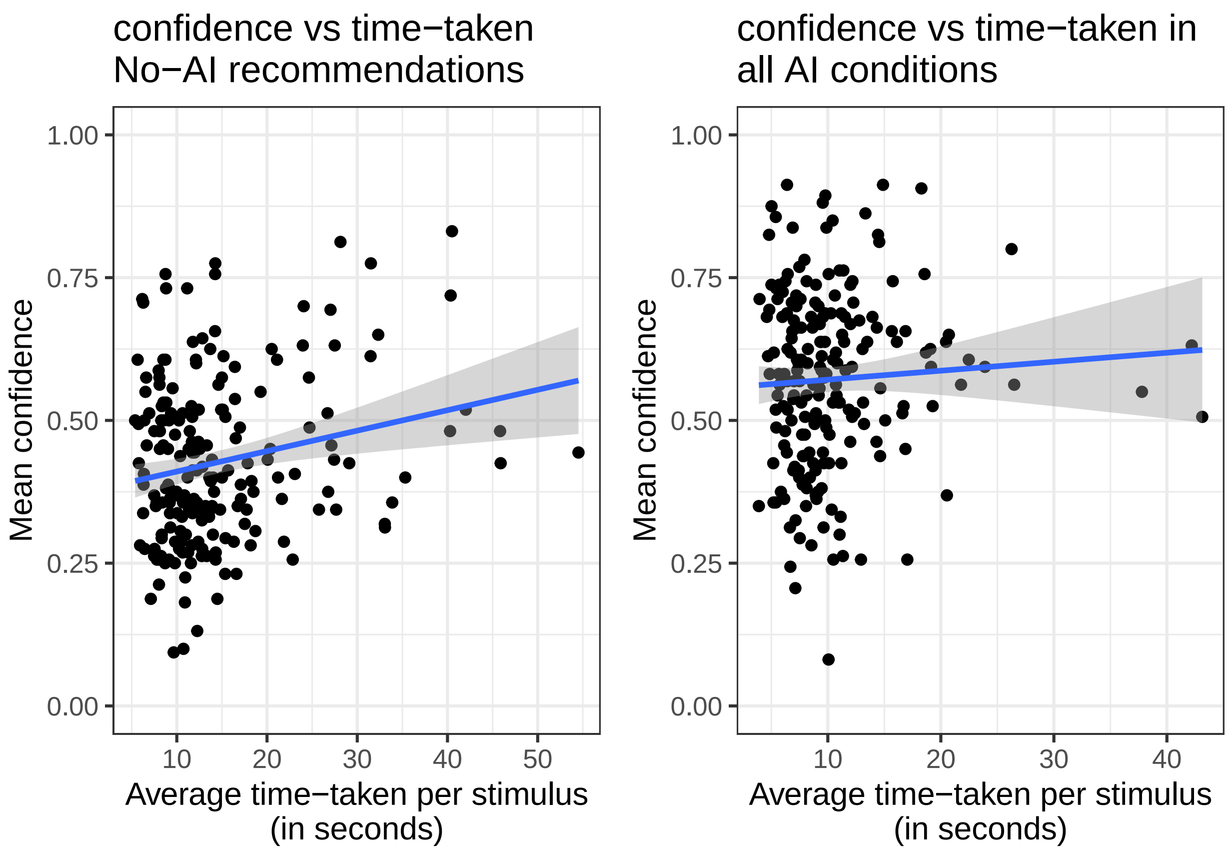
```r
#Time taken vs confidence
time_taken_plot1 <- ggplot(person_noAI, aes(x=time_taken, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Relationship between average time taken per stimulus
          and confidence of the participants") #main plot title

time_taken_plot1
```



There is no clear relationship between time taken and confidence in AI or No-AI conditions. Even though an extreme outlier was removed, new outliers will come up since the range in average time taken in wide. Graphs supports that as well. Based on the plots generated, time taken does not have any effect on the confidence of the participants.

```r
#replotting the same graph without the outlier
person_noAI2 <- person_noAI[-201,]

#Time taken in no-AI condition
time_taken_plot_1 <- ggplot(person_noAI2, aes(x=time_taken, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus\n(in seconds)") +
```

```
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("confidence vs time-taken\nNo-AI recommendations") #main plot title

#Time-taken in AI-condition
time_taken_plot_2 <- ggplot(person_AI, aes(x=time_taken, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #adjusting x-axis title place
  xlab("Average time-taken per stimulus\n(in seconds)") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("confidence vs time-taken in\nall AI conditions") #main plot title


cowplot::plot_grid(time_taken_plot_1, time_taken_plot_2)
```



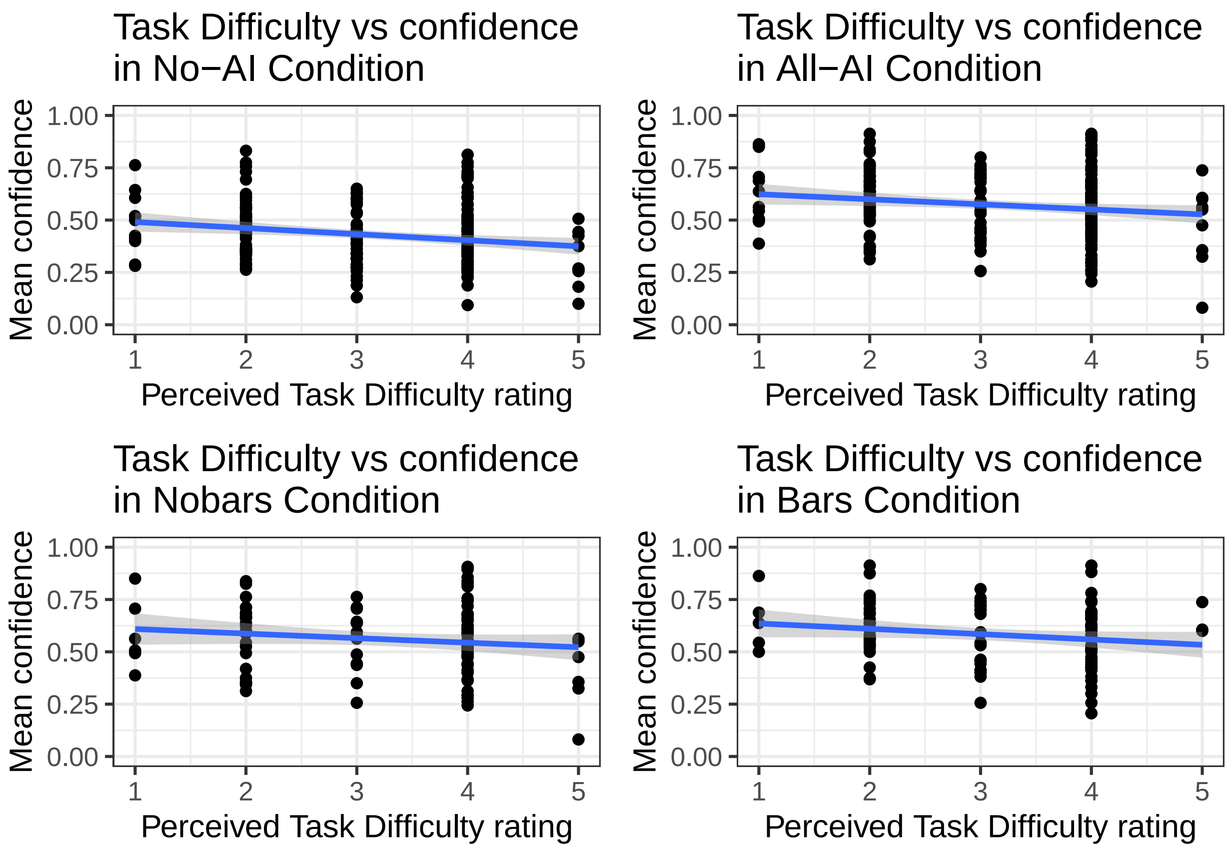
Task Difficulty vs confidence

Compared to No-AI condition, the confidence should improve in all AI conditions when the perceived task difficulty rating increases. confidence for participants placed in AI-nobar condition increases as participants task difficulty rating increases. However, there does not seem to be any relationship between their perceived ratings and confidence for participants placed in AI-bars condition.

```
Task_diff_plot_1 <- ggplot(person_noAI, aes(x=Task_diff_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
```

```r
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs confidence\nin No-AI Condition") #main plot title

Task_diff_plot_2 <- ggplot(person_AI, aes(x=Task_diff_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs confidence\nin All-AI Condition") #main plot title

Task_diff_plot_3 <- ggplot(person_nobar, aes(x=Task_diff_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs confidence\nin Nobars Condition") #main plot title

Task_diff_plot_4 <- ggplot(person_bar, aes(x=Task_diff_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived Task Difficulty rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Task Difficulty vs confidence\nin Bars Condition") #main plot title

cowplot::plot_grid(Task_diff_plot_1, Task_diff_plot_2, Task_diff_plot_3, Task_diff_plot_4)
```

```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```



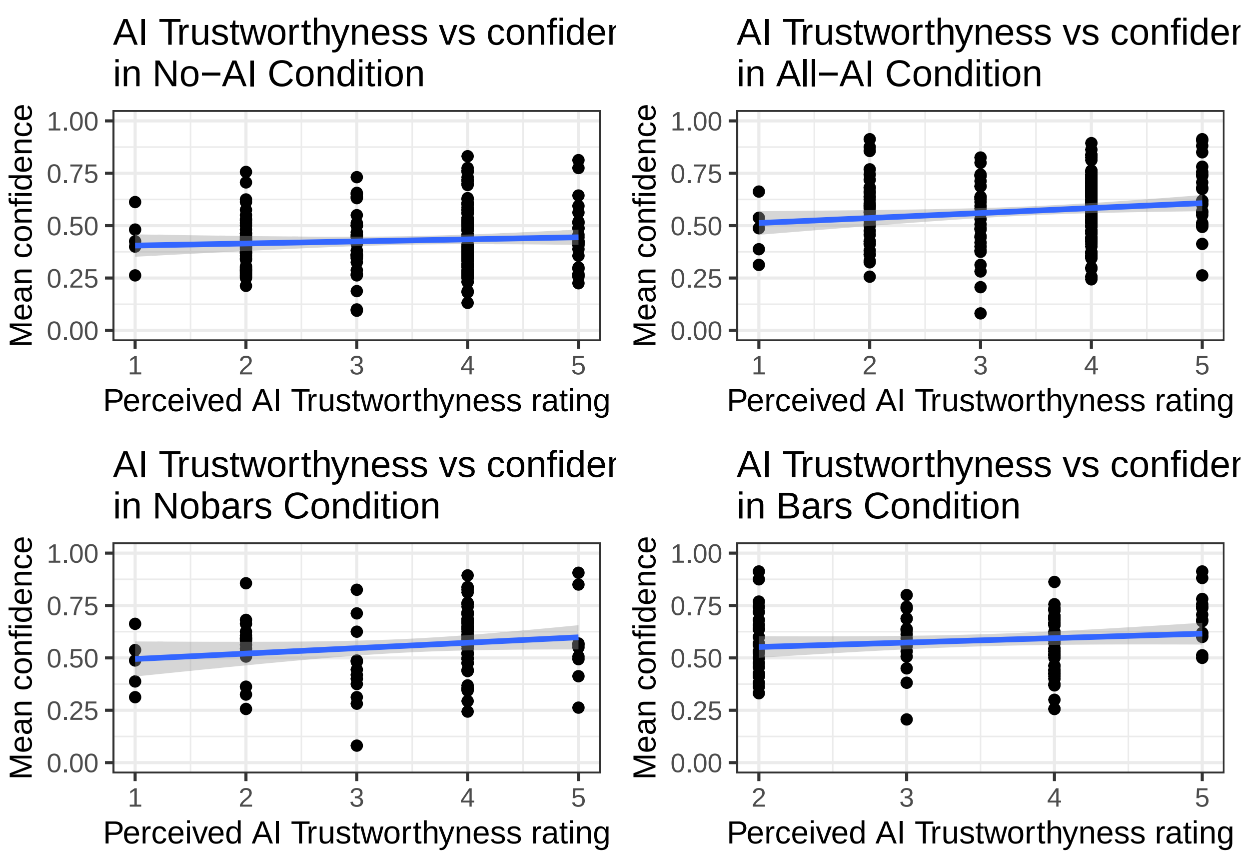
Perceived AI trustworthyness ratings vs confidence

There is no apparent relationship between AI trustworthyness and confidence in any and all conditions.

```
AI_trust_plot_1 <- ggplot(person_noAI, aes(x=AI_trust_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs confidence\nin No-AI Condition") #main plot title

AI_trust_plot_2 <- ggplot(person_AI, aes(x=AI_trust_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs confidence\nin All-AI Condition") #main plot title

AI_trust_plot_3 <- ggplot(person_nobar, aes(x=AI_trust_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
```

```
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs confidence\nin Nobars Condition") #main plot title

AI_trust_plot_4 <- ggplot(person_bar, aes(x=AI_trust_num, y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Perceived AI Trustworthyness rating") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("AI Trustworthyness vs confidence\nin Bars Condition") #main plot title

cowplot::plot_grid(AI_trust_plot_1, AI_trust_plot_2, AI_trust_plot_3, AI_trust_plot_4)
```

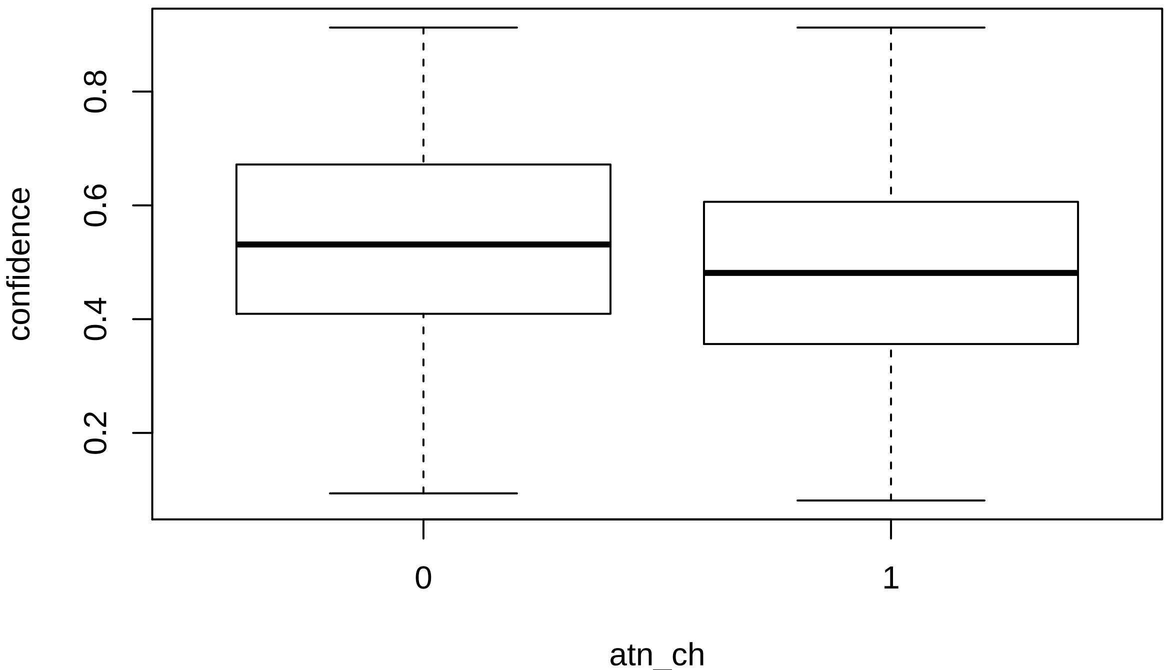
```
## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).

## Warning: Removed 1 rows containing non-finite values (stat_smooth).

## Warning: Removed 1 rows containing missing values (geom_point).
```



Attention Check vs confidence

There is no evidence suggesting confidence will be affected by participants performance on attention checks.

```r
boxplot(confidence ~ atn_ch, data = person)
```




```r
#Average confidence in no-bar vs bar
t.test(confidence ~ atn_ch, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  confidence by atn_ch
## t = 2.5908, df = 164.28, p-value = 0.01044
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.01225948 0.09082116
## sample estimates:
## mean in group 0 mean in group 1
##       0.5386418       0.4871015
```

```r
#Effect Size of the t.test
cohen.d(confidence ~ atn_ch, data = person)
```

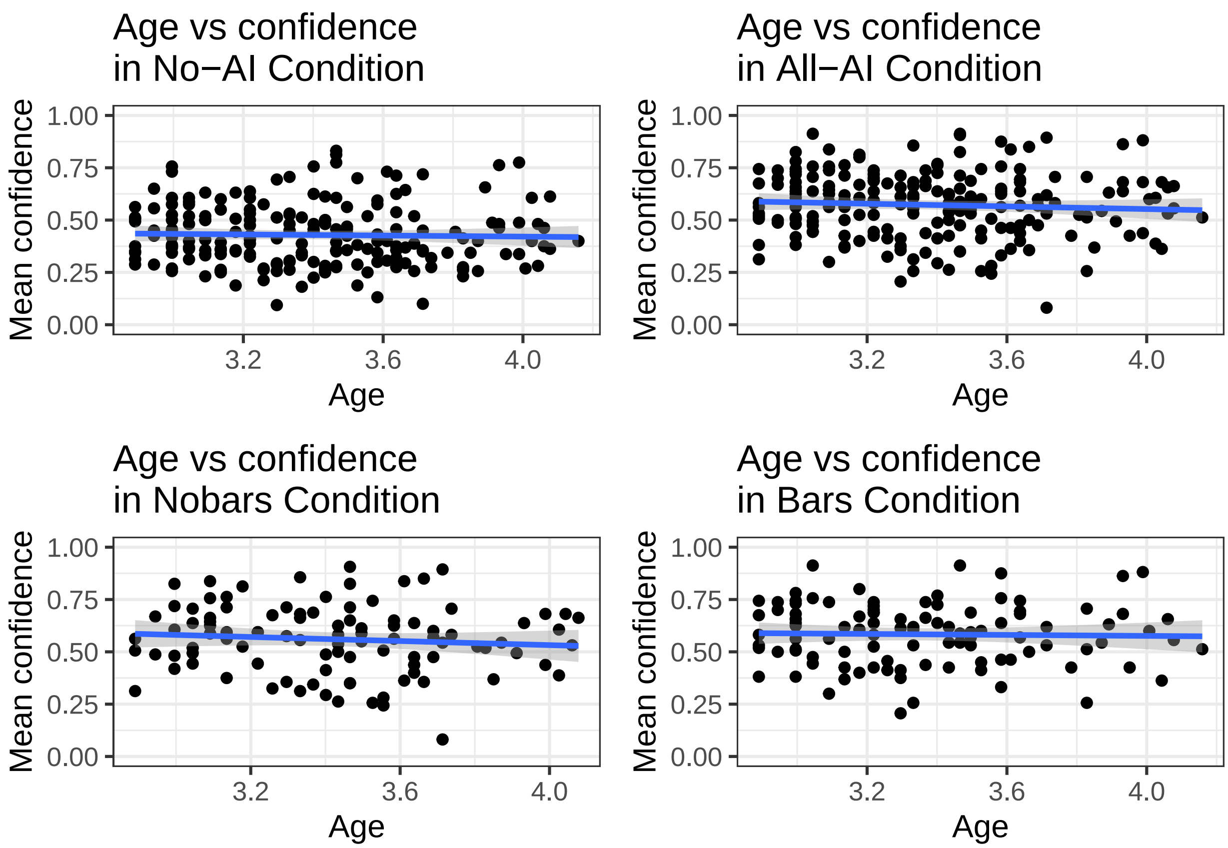
```
## Warning in cohen.d.formula(confidence ~ atn_ch, data = person): Cohercing rhs of
## formula to factor
```

```
##
## Cohen's d
##
## d estimate: 0.3111761 (small)
## 95 percent confidence interval:
##      lower      upper
## 0.08623996 0.53611228
```

Age vs confidence

There is no evidence suggesting confidence will be affected by participants age.

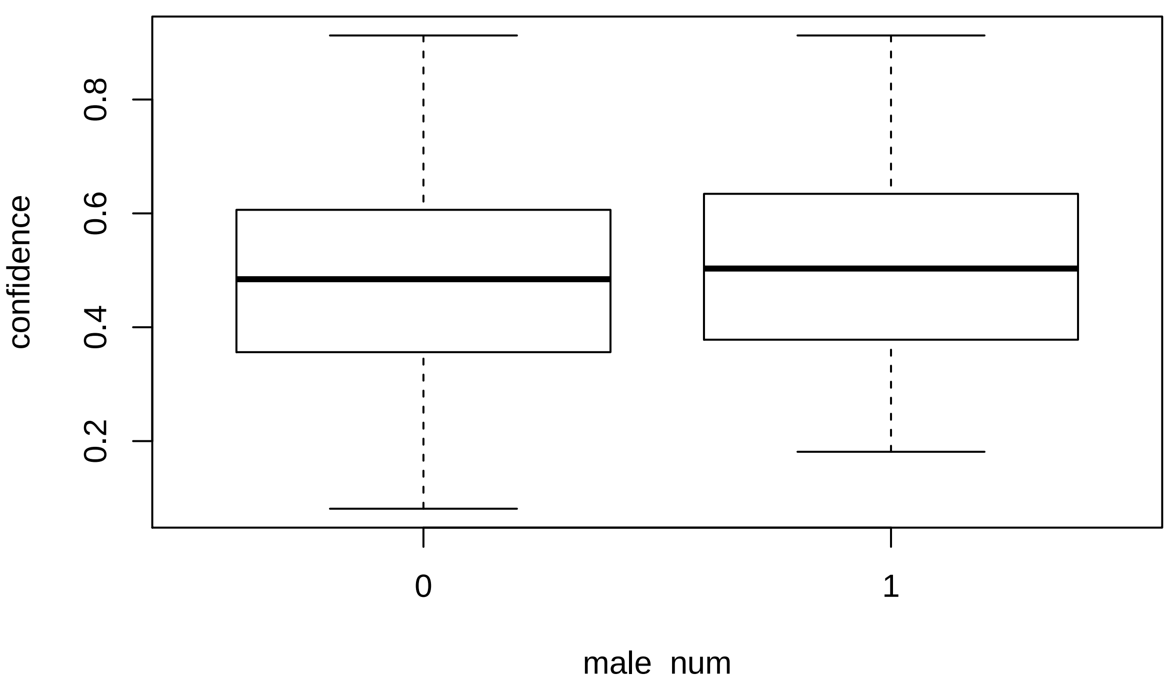
```
age_plot_1 <- ggplot(person_noAI, aes(x=log(age), y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs confidence\nin No-AI Condition") #main plot title

age_plot_2 <- ggplot(person_AI, aes(x=log(age), y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs confidence\nin All-AI Condition") #main plot title

age_plot_3 <- ggplot(person_nobar, aes(x=log(age), y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs confidence\nin Nobars Condition") #main plot title

age_plot_4 <- ggplot(person_bar, aes(x=log(age), y = confidence)) +
  geom_point() +
  geom_smooth(method = "lm", formula = y~x) +
  theme_bw(base_size = 12) + #styling the plot
  xlab("Age") +
  ylab("Mean confidence") + #axis labels
  ylim(0,1) + #providing the y-axis limits for the plot
  ggtitle("Age vs confidence\nin Bars Condition") #main plot title

cowplot::plot_grid(age_plot_1, age_plot_2, age_plot_3, age_plot_4)
```



Gender vs confidence

Although there is a slight difference, it is not significant and the effect size is negligible. Gender should not affect confidence.

```
boxplot(confidence ~ male_num, data = person)
```



```r
t.test(confidence ~ male_num, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  confidence by male_num
## t = -1.2695, df = 399.86, p-value = 0.205
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.05381796  0.01158433
## sample estimates:
## mean in group 0 mean in group 1
##       0.4901396       0.5112564
```

```r
cohen.d(confidence ~ male_num, data = person)
```

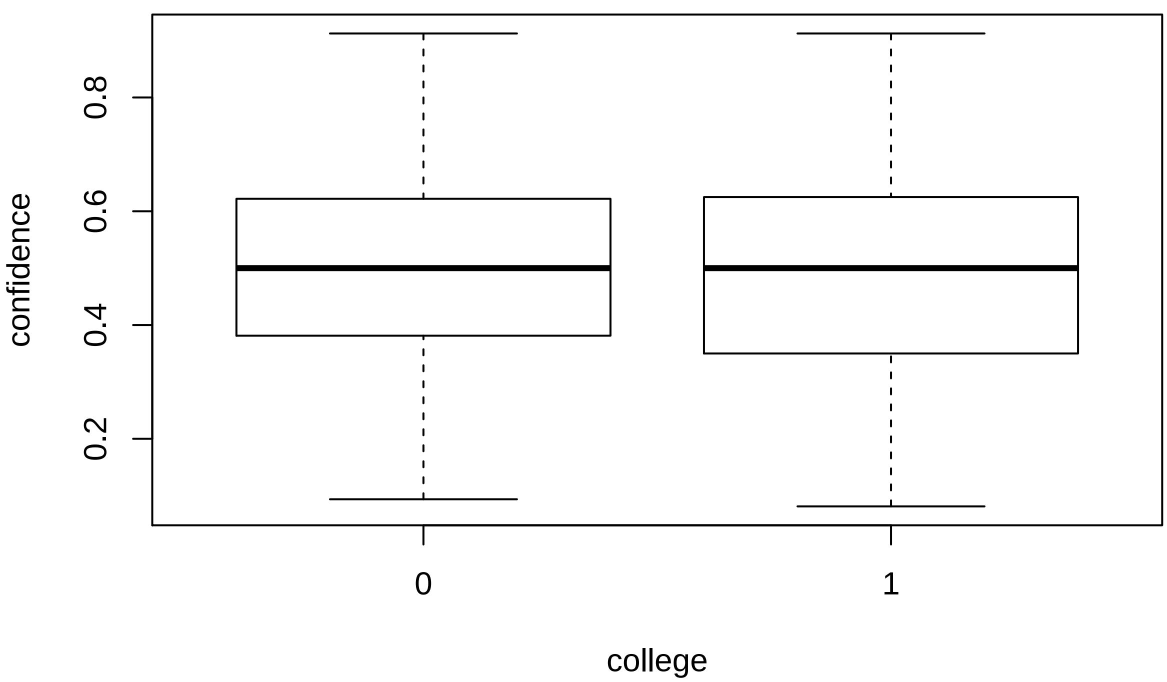
```
## Warning in cohen.d.formula(confidence ~ male_num, data = person): Cohercing rhs
## of formula to factor
```

```
##
## Cohen's d
##
## d estimate: -0.1265739 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.3229324  0.0697846
```

College vs confidence

Participants education level should not affect confidence.

```r
boxplot(confidence ~ college, data = person)
```



```r
t.test(confidence ~ college, data = person)
```

```
##
##  Welch Two Sample t-test
##
## data:  confidence by college
## t = 0.34001, df = 384.91, p-value = 0.734
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02725100  0.03864696
## sample estimates:
## mean in group 0 mean in group 1
##       0.5031851       0.4974871
```

```r
cohen.d(confidence ~ college, data = person)
```

```
## Warning in cohen.d.formula(confidence ~ college, data = person): Cohercing rhs
## of formula to factor
```

```
##
## Cohen's d
##
## d estimate: 0.03409011 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.1621447  0.2303250
```

END OF EXPLORATORY PLOTS FOR DATASET WHERE EACH OBSERVATION WAS A PERSON

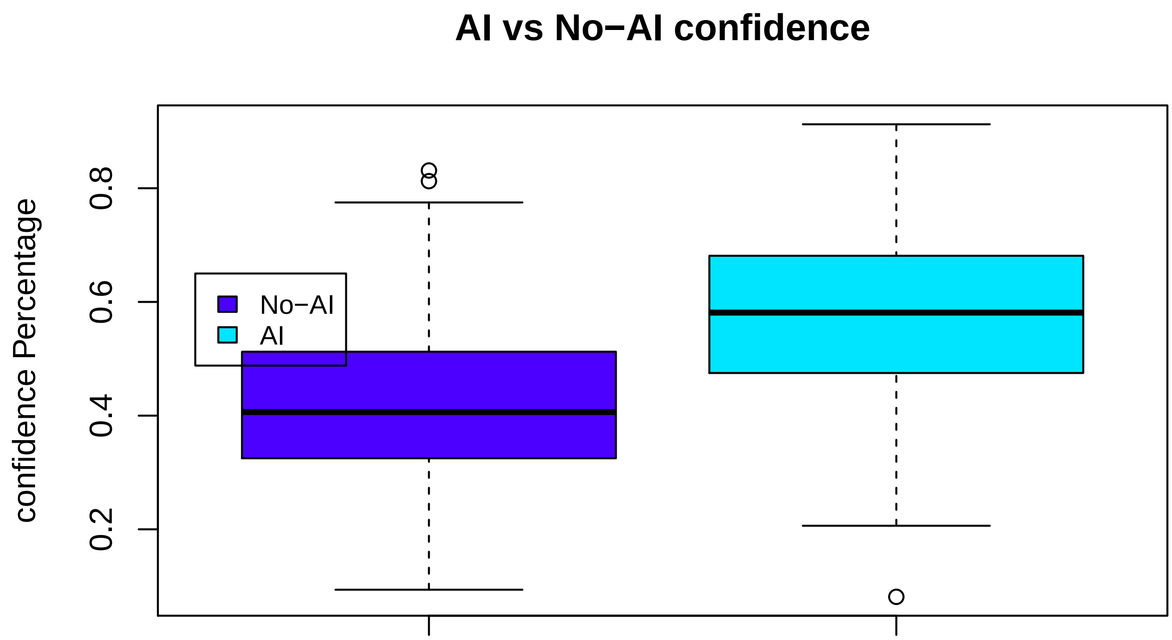
Box-Plots and T-test

Comparing confidence Pre-AI and Post-AI

The box plot clearly indicates a positive change in confidence when participants were provided AI recommendations. The t-test results along with the box plots shows that AI recommendations will improve confidence. The repeated measures paired t-test indicates that confidence was significantly higher when AI information was provided (M = -0.14). The significant t-test with $t(200) = -18.07$, $p < 0.001$. An effect size $d = -0.95$ proves the change in confidence is most likely not due to chance.

```r
#Boxplot
boxplot(person_noAI$confidence,#box plot. Comparing two sets of data
```

```
        person_AI$confidence, #comparing no-AI vs AI confidence
        col= topo.colors(2), #mentioning colors of the boxes
        main = "AI vs No-AI confidence", #main title
        ylab = "confidence Percentage") #y-axis title
legend(.5,.65, #legend, mentioning where in the graph to place
       inset = 0.2, c("No-AI","AI"), #mentioning text for the legend
       fill = topo.colors(2), cex=0.8) #providing color info for the boxes.
```




```
#Average confidence in AI vs no_AI
t.test(person_noAI$confidence, person_AI$confidence, #t-test to compare.
       paired = TRUE) #paired is true as this is a within-subjects comparison
```

```
##
##  Paired t-test
##
## data:  person_noAI$confidence and person_AI$confidence
## t = -18.067, df = 200, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1587841 -0.1275343
## sample estimates:
## mean of the differences
##              -0.1431592
```

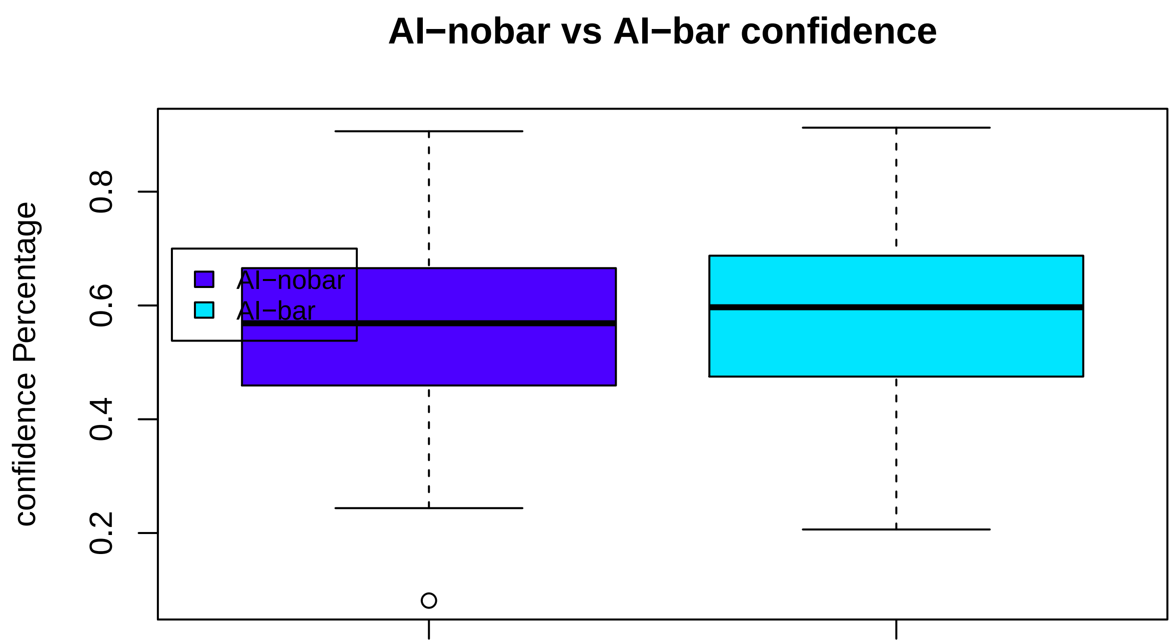
```
#Effect Size of the t.test
cohen.d(person_noAI$confidence, person_AI$confidence, #effect size of the comparison
        paired = TRUE)
```

```
##
## Cohen's d
##
## d estimate: -0.9455242 (large)
## 95 percent confidence interval:
##      lower      upper
## -1.0692863 -0.8217622
```

Comparing confidence in AI-nobar and AI-bar conditions

By looking at the box plot, it is clear confidence of the participants who received uncertainty information performed better compared to participants who did not. The average confidence changed significantly when comparing AI vs No-AI performance but the change, although significant, is less pronounced, when comparing between participants receiving uncertainty information vs participants not receiving uncertainty information. The Two sample t-test indicates that confidence was not significantly higher when uncertainty information was provided (M = 0.58) compared to when uncertinaty information was not provided (M = 0.56).

```
#Boxplot
boxplot(person_nobar$confidence, person_bar$confidence,
        col= topo.colors(2),
        main = "AI-nobar vs AI-bar confidence",
        ylab = "confidence Percentage")
legend(.45,.70, inset = 0.2, c("AI-nobar","AI-bar"), fill = topo.colors(2), cex=0.8)
```




```
#Average confidence in no-bar vs bar
t.test(person_nobar$confidence, person_bar$confidence)
```

```
##
##  Welch Two Sample t-test
##
## data:  person_nobar$confidence and person_bar$confidence
## t = -1.0666, df = 195.66, p-value = 0.2875
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06688671  0.01993187
## sample estimates:
## mean of x mean of y
## 0.5601010 0.5835784
```

```
#Effect Size of the t.test
cohen.d(person_nobar$confidence, person_bar$confidence)
```

```
##
## Cohen's d
```

```
##
## d estimate: -0.1507116 (negligible)
## 95 percent confidence interval:
##      lower      upper
## -0.4293194  0.1278961
```

LINEAR MODELS ON confidence

Effect of AI on confidence

Results of the simple linear regression indicate a positive significant relationship between AI recommendations and confidence (F(1,400) = 90.35, p <0.001, R2 = 0.18).

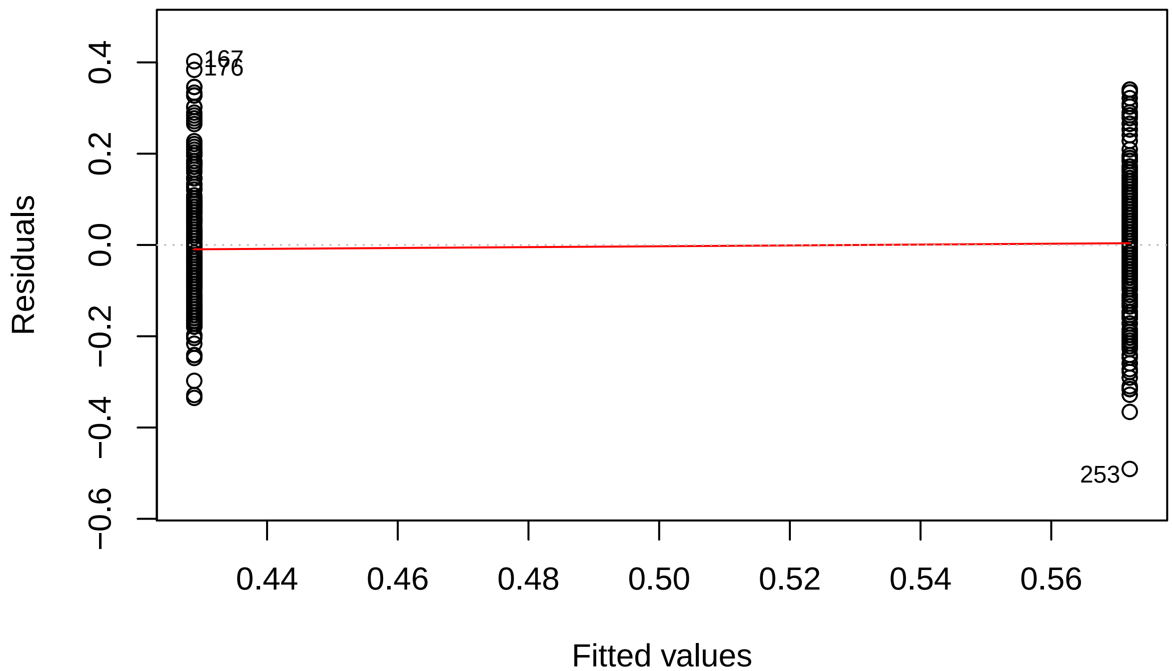
```r
lm.1.conf <- lm(confidence ~ AI, data = person) #linear model
summary(lm.1.conf)
```

```
##
## Call:
## lm(formula = confidence ~ AI, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.49076 -0.10229 -0.00386  0.09659  0.40239
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.42886    0.01065  40.270   <2e-16 ***
## AI           0.14316    0.01506   9.505   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.151 on 400 degrees of freedom
## Multiple R-squared:  0.1843, Adjusted R-squared:  0.1822
## F-statistic: 90.35 on 1 and 400 DF,  p-value: < 2.2e-16
```

Given the predictor variable is binary, we see a pattern in the residuals vs fitted plot. Otherwise, the model is confeptable. The patter in the Q-Q plot is confeptable given the binary predictor variable, but it does deviate from the line at the edges.

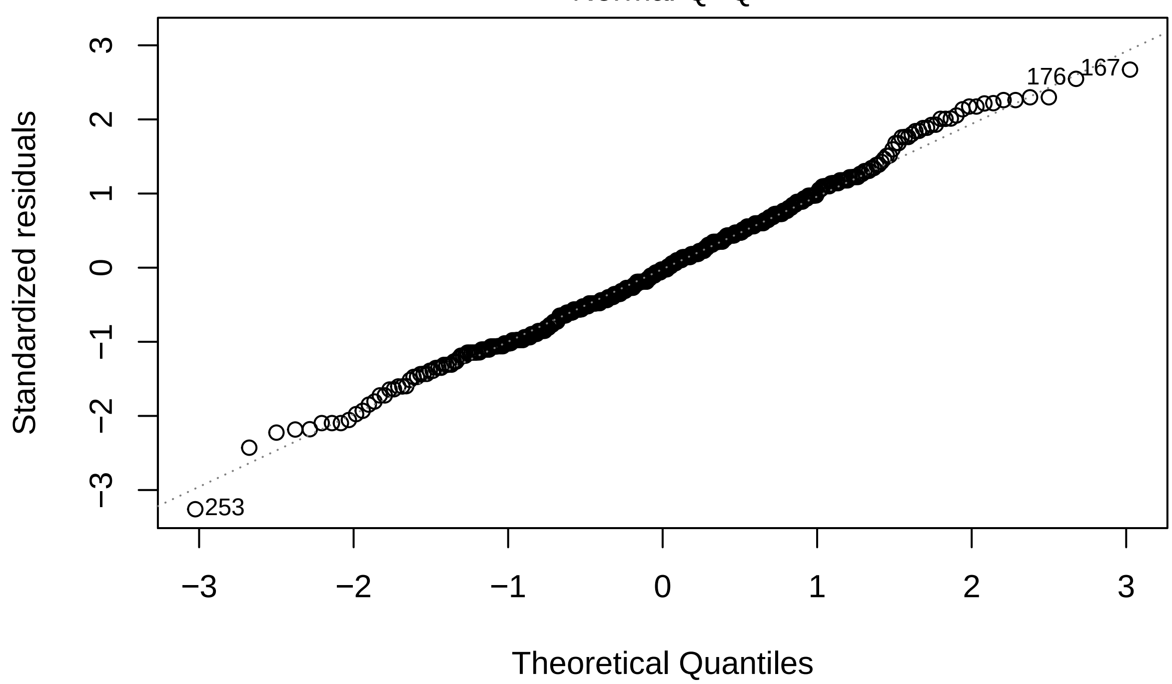
```r
plot(lm.1.conf)
```

Residuals vs Fitted



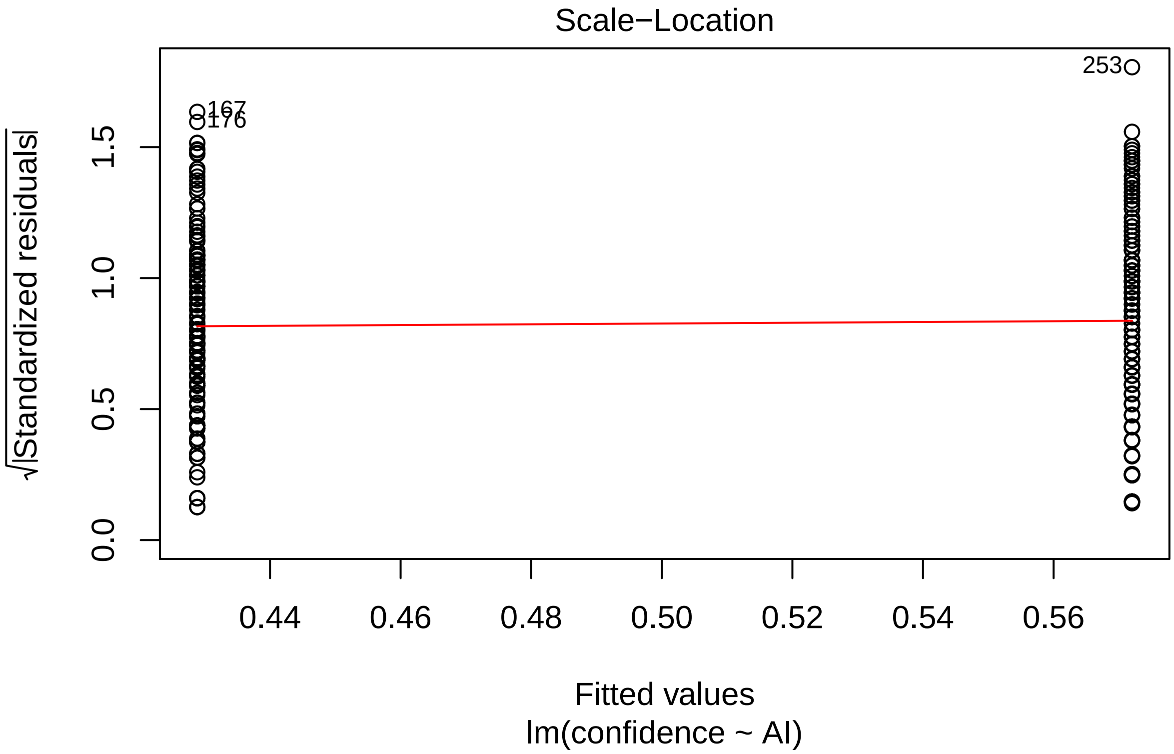
lm(confidence ~ AI)

Normal Q−Q



lm(confidence ~ AI)

```
## hat values (leverages) are all = 0.004975124
##  and there are no factor predictors; no plot no. 5
```



Effect of Uncertainty Information on confidence

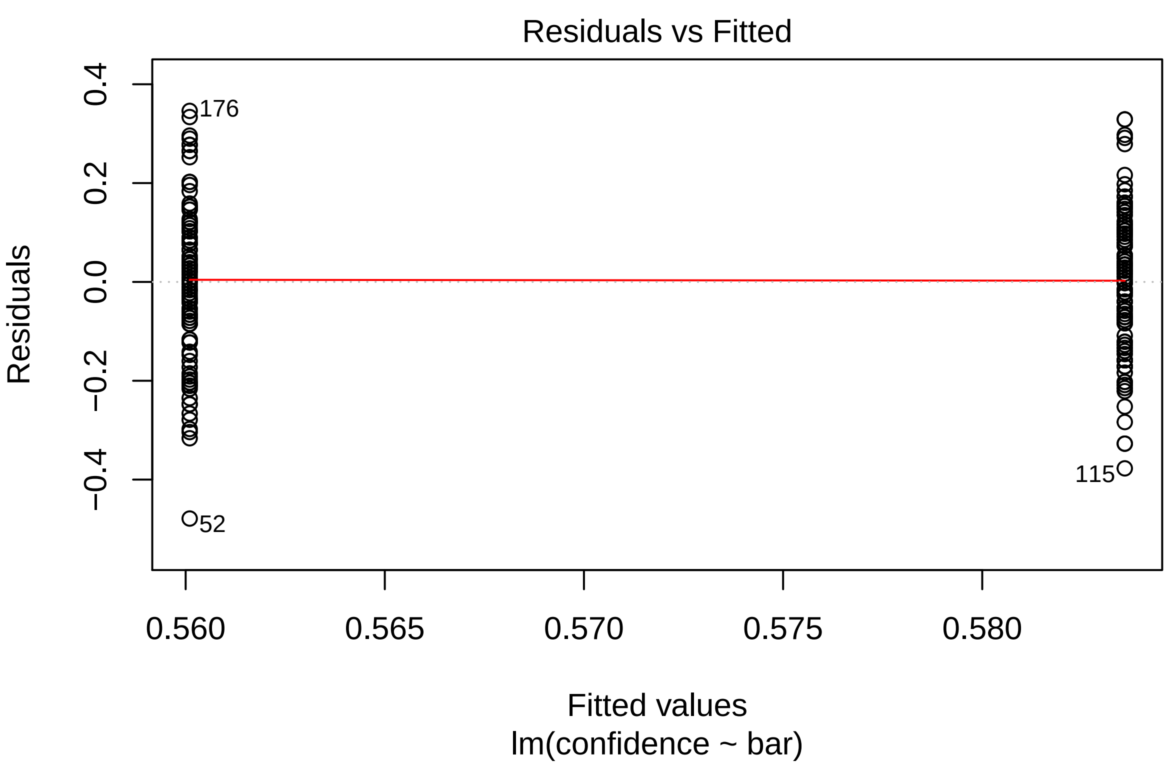
Results of the simple linear regression indicate an insignificant relationship between Unceraintny information and confidence.

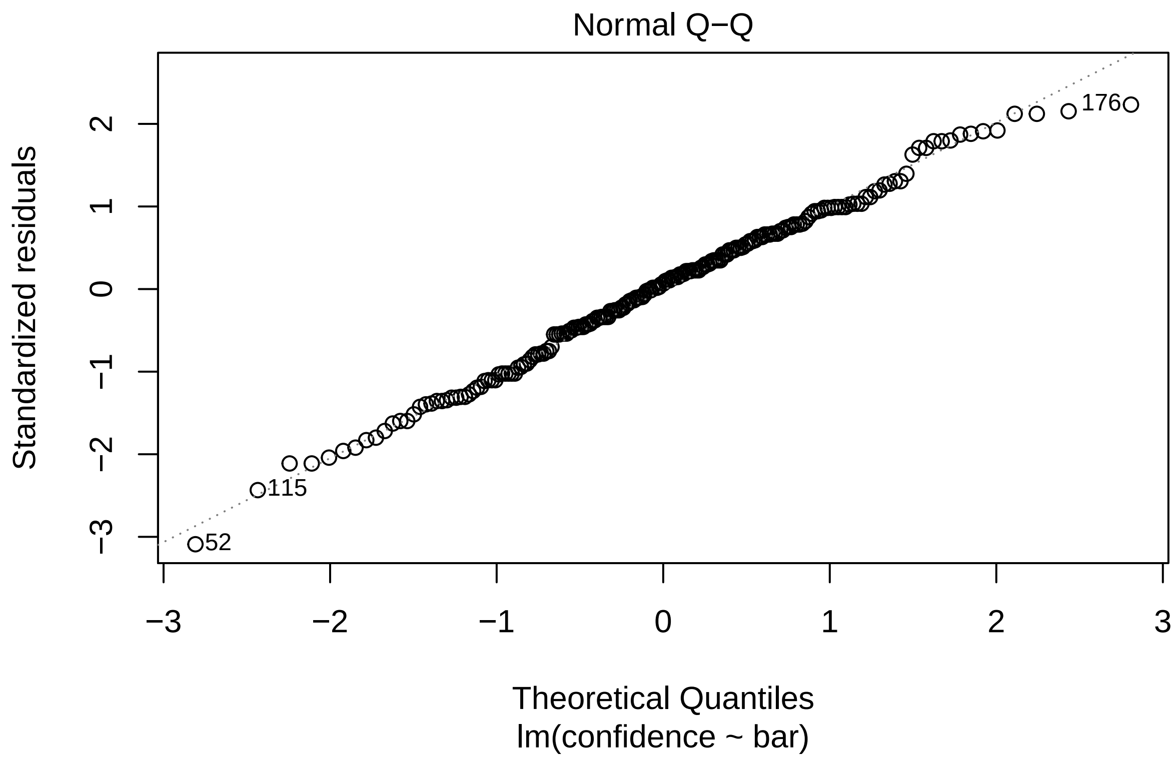
```r
lm.2.conf <- lm(confidence ~ bar, data = person_AI)

summary(lm.2.conf)
```

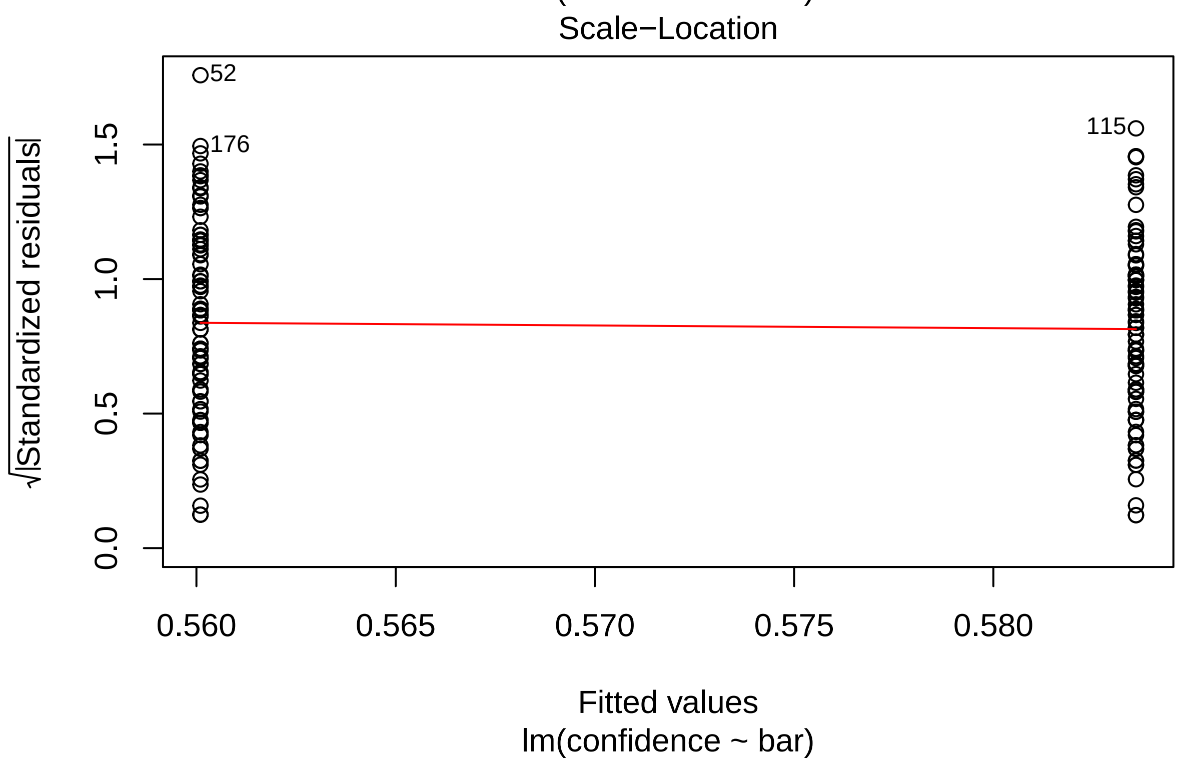
```
##
## Call:
## lm(formula = confidence ~ bar, data = person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.47885 -0.10858  0.01017  0.10392  0.34615
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.56010    0.01566  35.775   <2e-16 ***
## bar          0.02348    0.02198   1.068    0.287
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1558 on 199 degrees of freedom
## Multiple R-squared:  0.005702,   Adjusted R-squared:  0.0007051
## F-statistic: 1.141 on 1 and 199 DF,  p-value: 0.2867
```

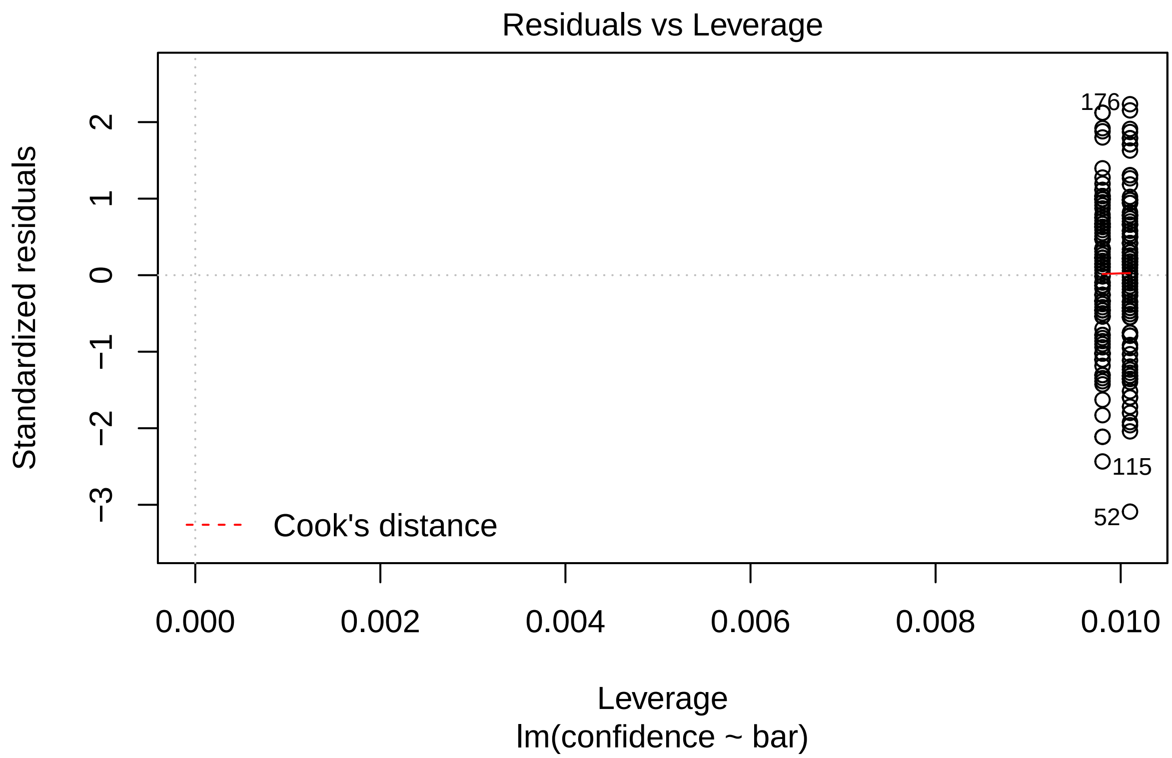
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.2.conf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(confidence ~ bar)
176
115
52


Scale−Location
√|Standardized residuals|
Fitted values
lm(confidence ~ bar)
52
176
115



Effect of AI on confidence with other predictor variables

AI recommendations, time taken, task difficulty, attention check, and log(age) are all significant. More time taken by the participants increases their confidence so does the provision of AI recommendations. Rest of the significant predictor variables all negatively affect confidence. In other words, they reduce the participants' confidence. F(8, 389) = 17.34, p < 0.001, R^2 = 0.25

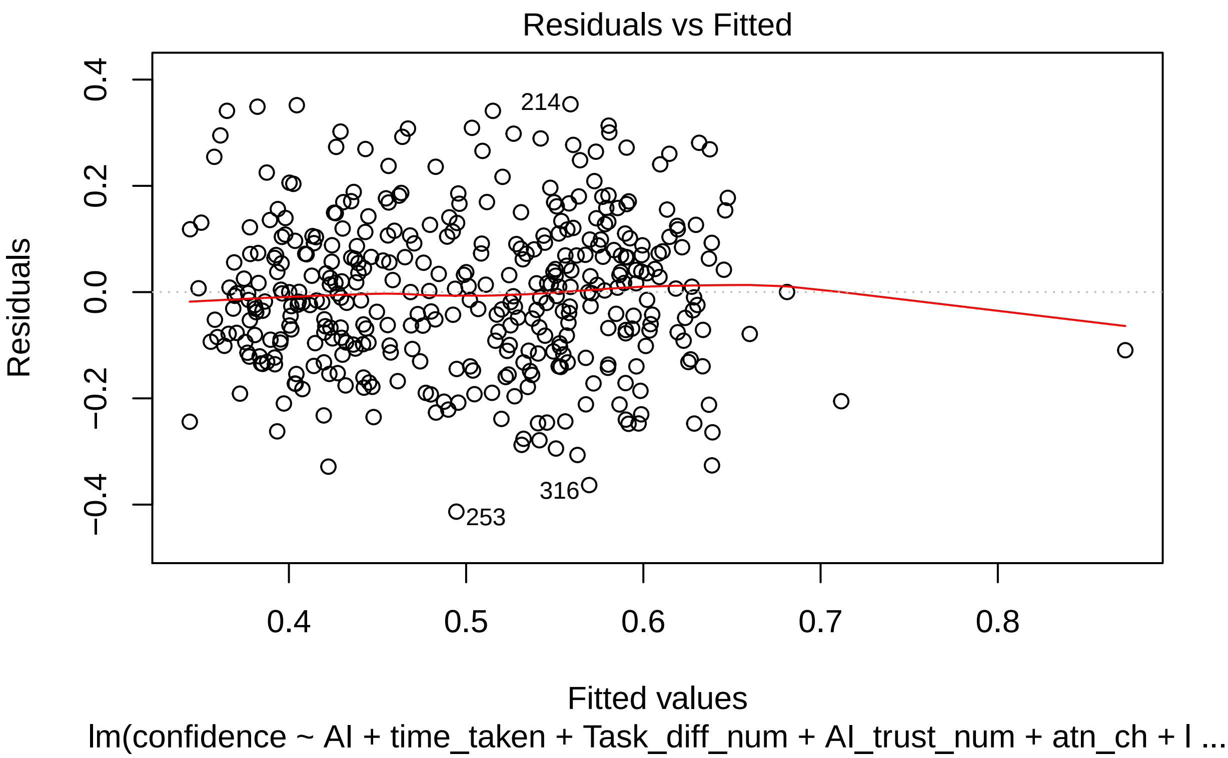
```r
lm.3.conf <- lm(confidence ~ AI + time_taken + Task_diff_num +
                 AI_trust_num + atn_ch + log(age) + male_num + college,
               data = person)

summary(lm.3.conf)
```

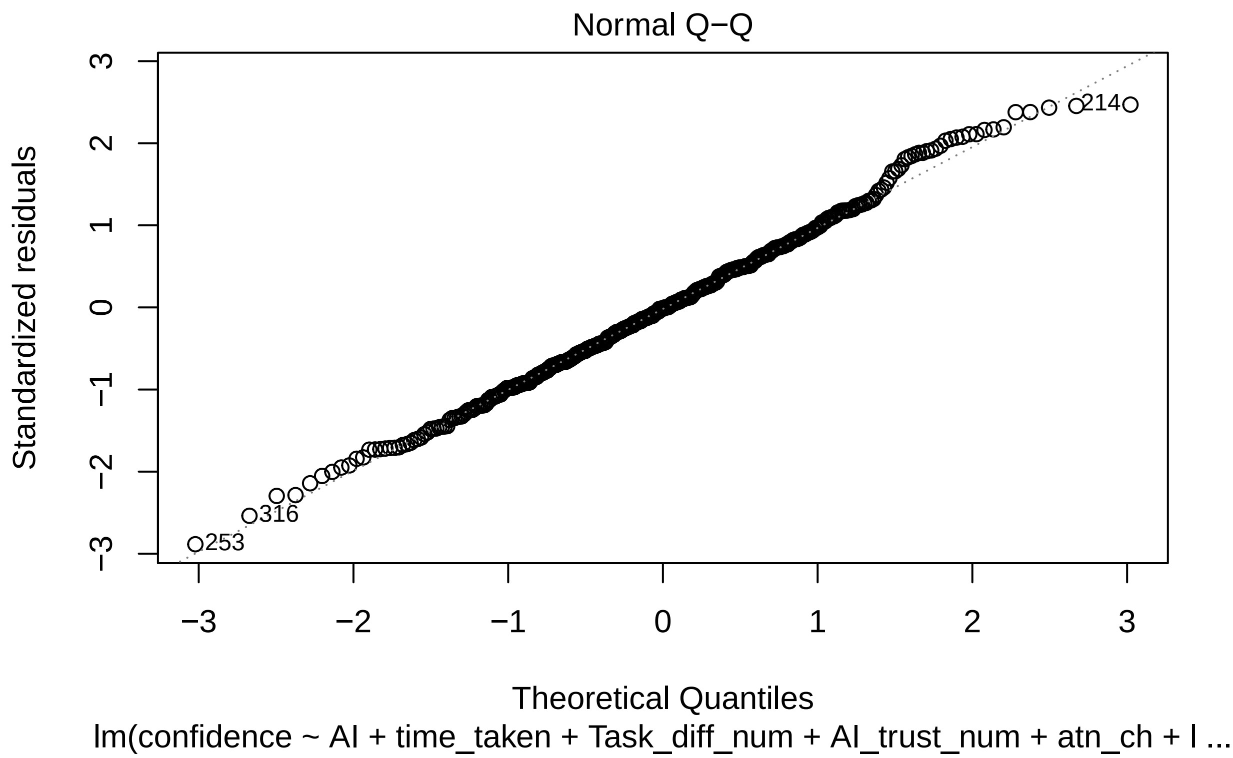
```
##
## Call:
## lm(formula = confidence ~ AI + time_taken + Task_diff_num + AI_trust_num +
##     atn_ch + log(age) + male_num + college, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41324 -0.09769 -0.00200  0.09251  0.35363
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6228537  0.0894197   6.966  1.4e-11 ***
## AI             0.1516125  0.0147623  10.270  < 2e-16 ***
## time_taken     0.0017660  0.0005247   3.366  0.00084 ***
## Task_diff_num -0.0221802  0.0072572  -3.056  0.00240 **
## AI_trust_num   0.0126099  0.0072407   1.742  0.08238 .
## atn_ch        -0.0537348  0.0167330  -3.211  0.00143 **
```

```
## log(age)      -0.0505868  0.0247155  -2.047  0.04135 *
## male_num       0.0148273  0.0147465   1.005  0.31529
## college        0.0169048  0.0159569   1.059  0.29007
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1446 on 389 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.2629, Adjusted R-squared:  0.2477
## F-statistic: 17.34 on 8 and 389 DF,  p-value: < 2.2e-16
```

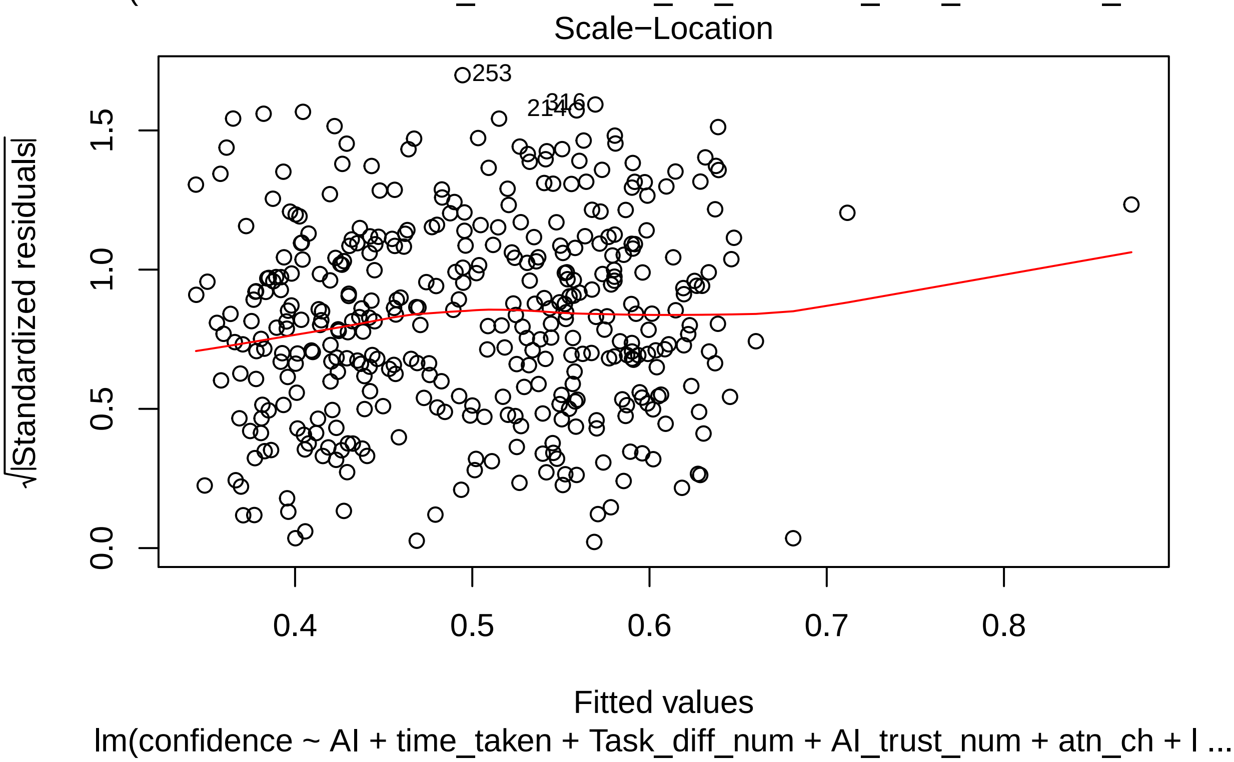
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.3.conf)
```

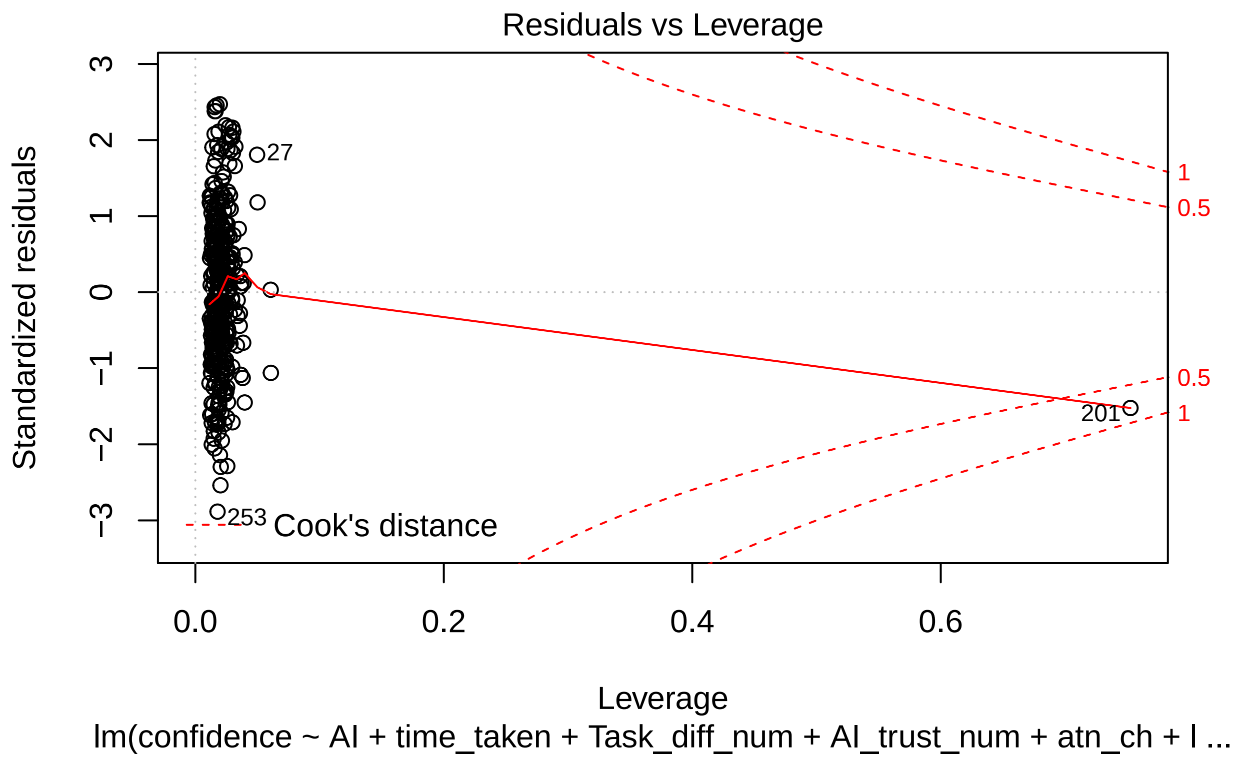




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(confidence ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + l ...
214
316
253



Scale−Location
√|Standardized residuals|
Fitted values
lm(confidence ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + l ...
253
214
316



Effect of Uncertainty Information on confidence with other predictor variables

Only perceived AI usefulness rating is significantly affecting the participants confidence. $F(9, 189) = 15.32$, $p < 0.001$, R^2 = 0.39.

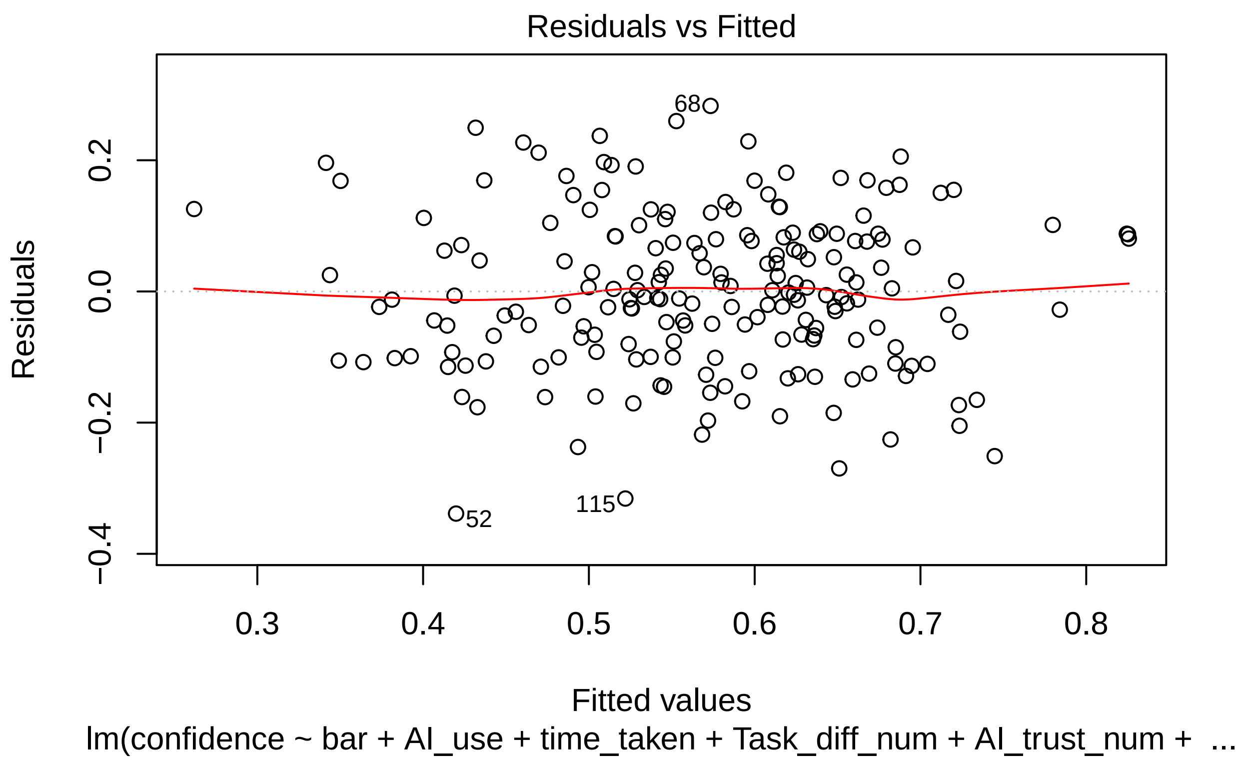
```
lm.4.conf <- lm(confidence ~ bar + AI_use + time_taken + Task_diff_num +
                 AI_trust_num + atn_ch + log(age) + male_num + college,
               data = person_AI)

summary(lm.4.conf)
```

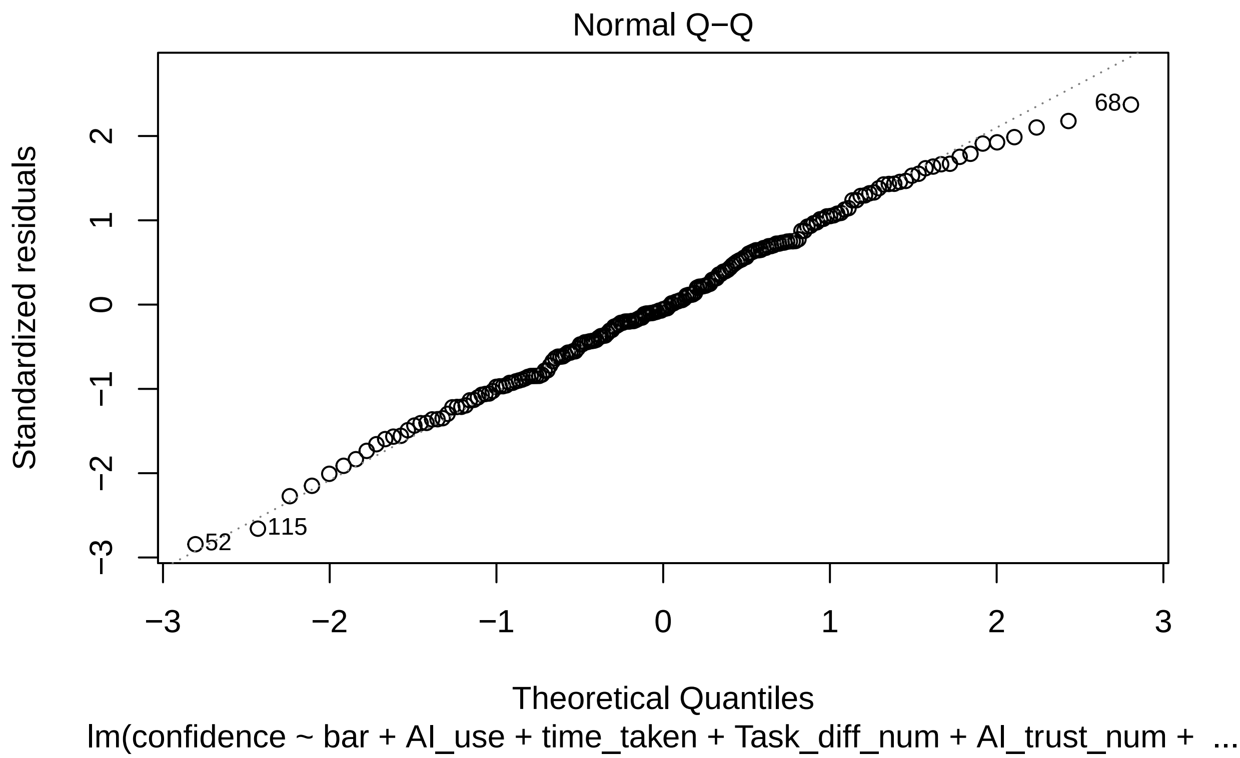
```
##
## Call:
## lm(formula = confidence ~ bar + AI_use + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33864 -0.08265 -0.00644  0.08410  0.28285
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     0.519206   0.110867   4.683 5.38e-06 ***
## bar            -0.002082   0.017980  -0.116   0.9079
## AI_use          0.599625   0.057763  10.381  < 2e-16 ***
## time_taken      0.001954   0.001562   1.251   0.2124
## Task_diff_num -0.016492   0.008612  -1.915   0.0570 .
## AI_trust_num  -0.017654   0.009375  -1.883   0.0612 .
## atn_ch         -0.036359   0.020603  -1.765   0.0792 .
## log(age)       -0.051946   0.029507  -1.760   0.0799 .
```

```
## male_num        0.007064   0.017583   0.402   0.6883
## college         0.020837   0.019152   1.088   0.2780
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1213 on 189 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.4218, Adjusted R-squared:  0.3943
## F-statistic: 15.32 on 9 and 189 DF,  p-value: < 2.2e-16
```

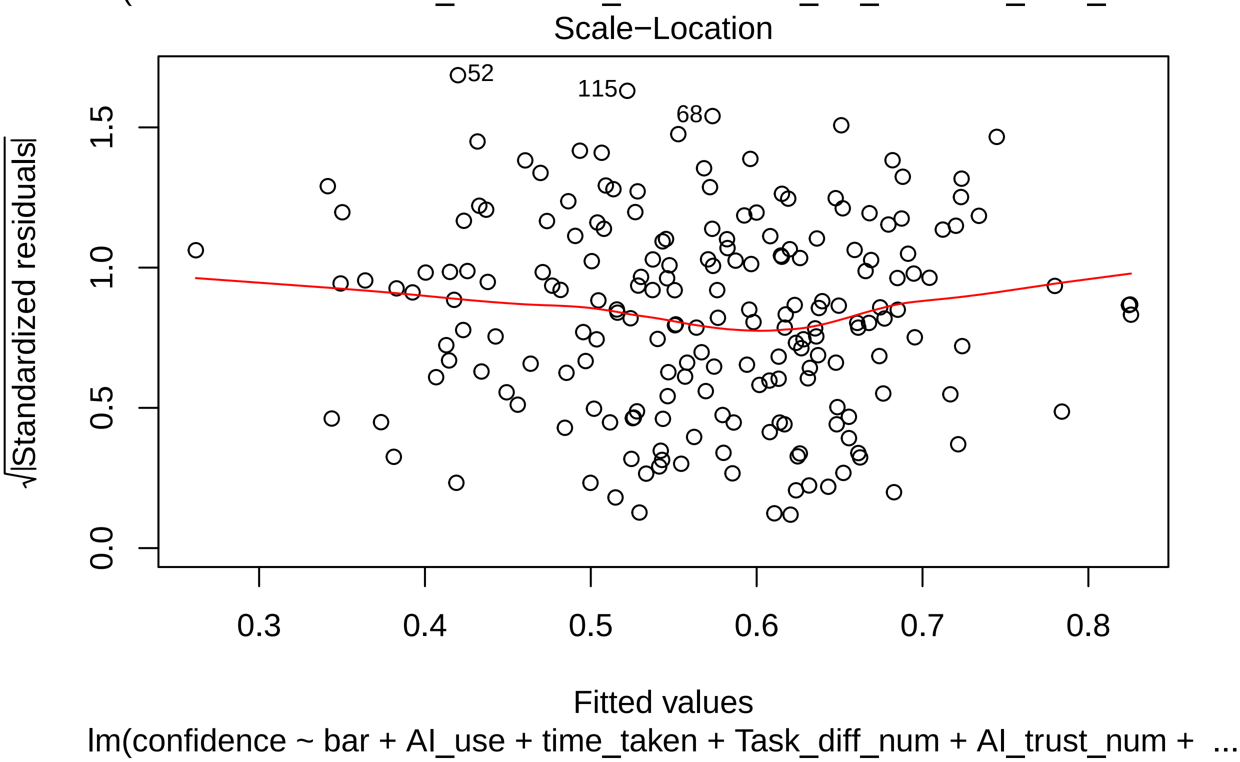
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.4.conf)
```

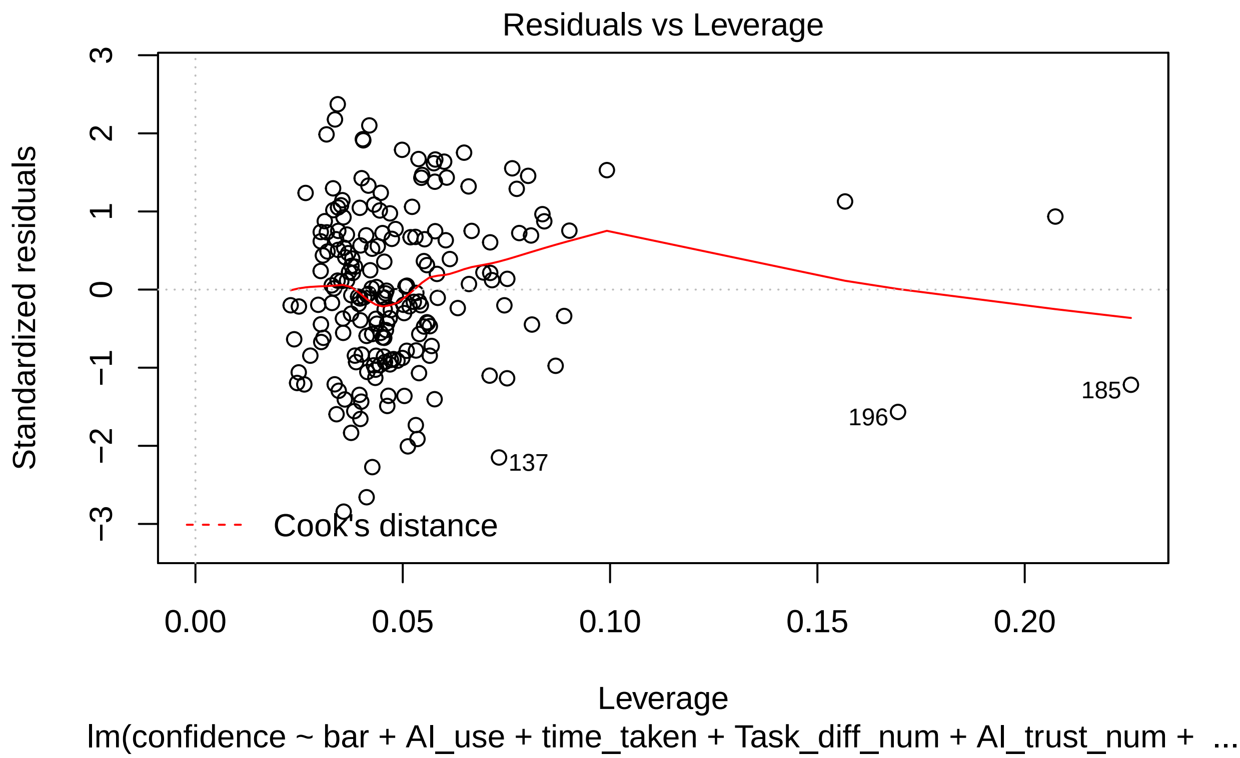




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(confidence ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + ...
52
115
68



Scale−Location
√|Standardized residuals|
Fitted values
lm(confidence ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + ...
52
115
68



Effect of AI recommendations on confidence with animal domain knowledge

Domain knowledge, AI recommendations, Time taken, Task difficulty rating, gender, and the interaction of domain knowledge and AI recommendations are all significantly affecting the confidence of the participants. The interaction however is negatively affecting confidence. In the presence of AI, as the domain knowledge rating increases, their confidence decreases.

F(10, 387) = 22, p < 0.001, R^2 0.35.

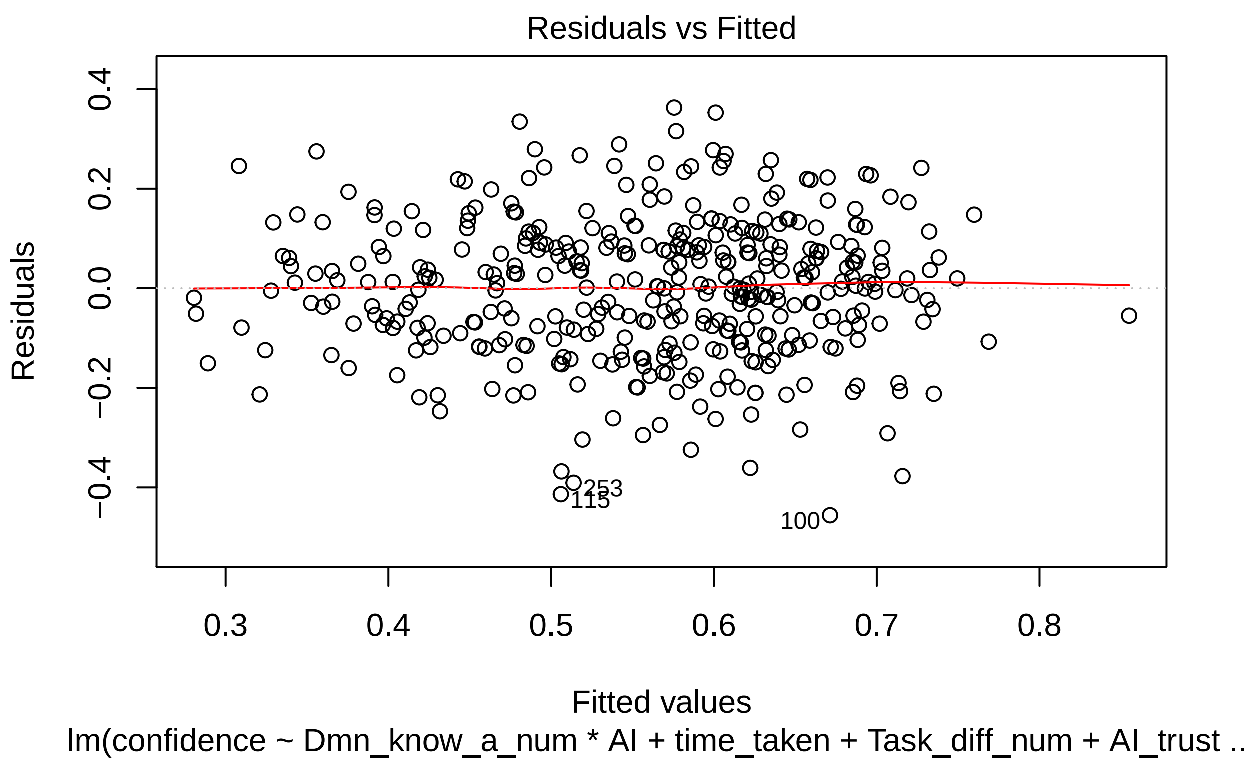
```
lm.5.a.conf <- lm(confidence ~ Dmn_know_a_num*AI +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = animals_person)

summary(lm.5.a.conf)
```

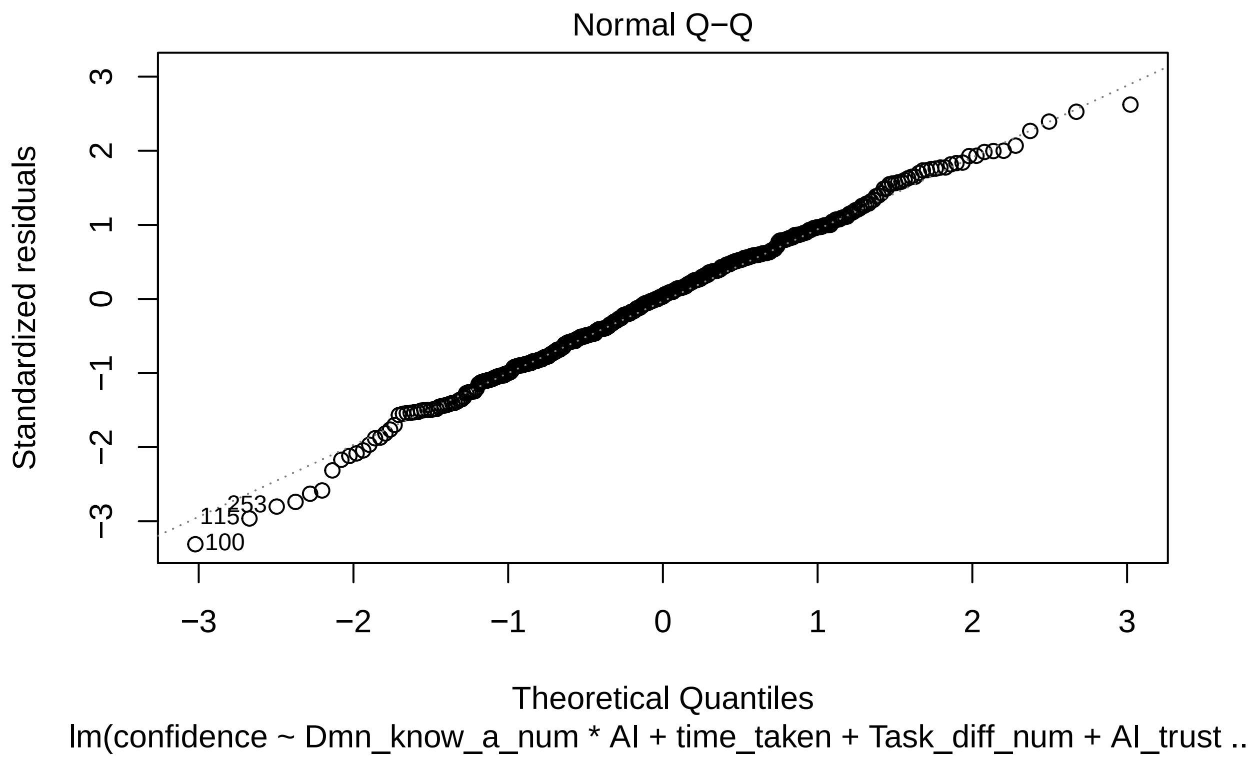
```
##
## Call:
## lm(formula = confidence ~ Dmn_know_a_num * AI + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = animals_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45591 -0.09395  0.00464  0.08682  0.36295
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.4498115  0.0913450   4.924 1.26e-06 ***
## Dmn_know_a_num      0.3515999  0.0431215   8.154 4.97e-15 ***
## AI                  0.2299658  0.0320966   7.165 3.96e-12 ***
## time_taken          0.0005070  0.0002322   2.183  0.02961 *
```

```
## Task_diff_num      -0.0209119  0.0071290  -2.933  0.00355 **
## AI_trust_num        0.0134094  0.0070519   1.902  0.05798 .
## atn_ch             -0.0252330  0.0165731  -1.523  0.12869
## log(age)           -0.0356204  0.0239989  -1.484  0.13856
## male_num            0.0294161  0.0143921   2.044  0.04164 *
## college             0.0176635  0.0155873   1.133  0.25783
## Dmn_know_a_num:AI -0.1846461  0.0602028  -3.067  0.00231 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1411 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.3624, Adjusted R-squared:  0.3459
## F-statistic: 21.99 on 10 and 387 DF,  p-value: < 2.2e-16
```

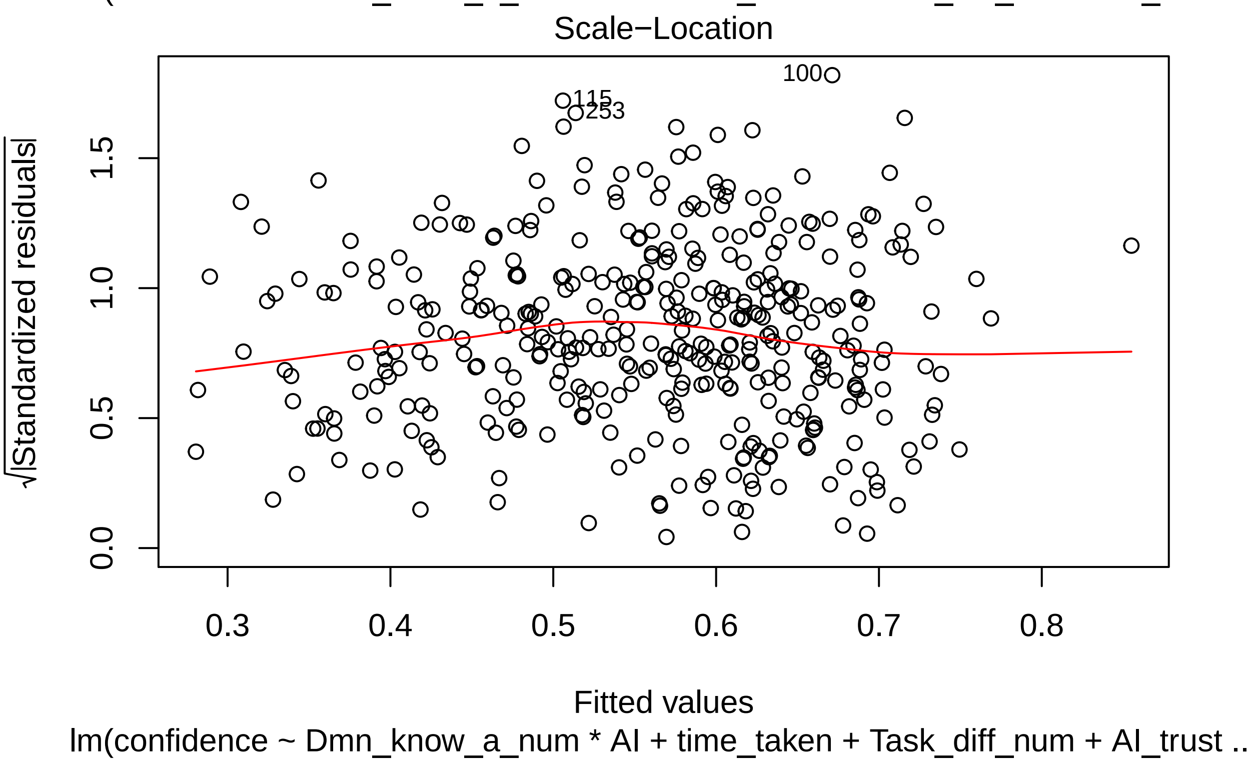
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.5.a.conf)
```

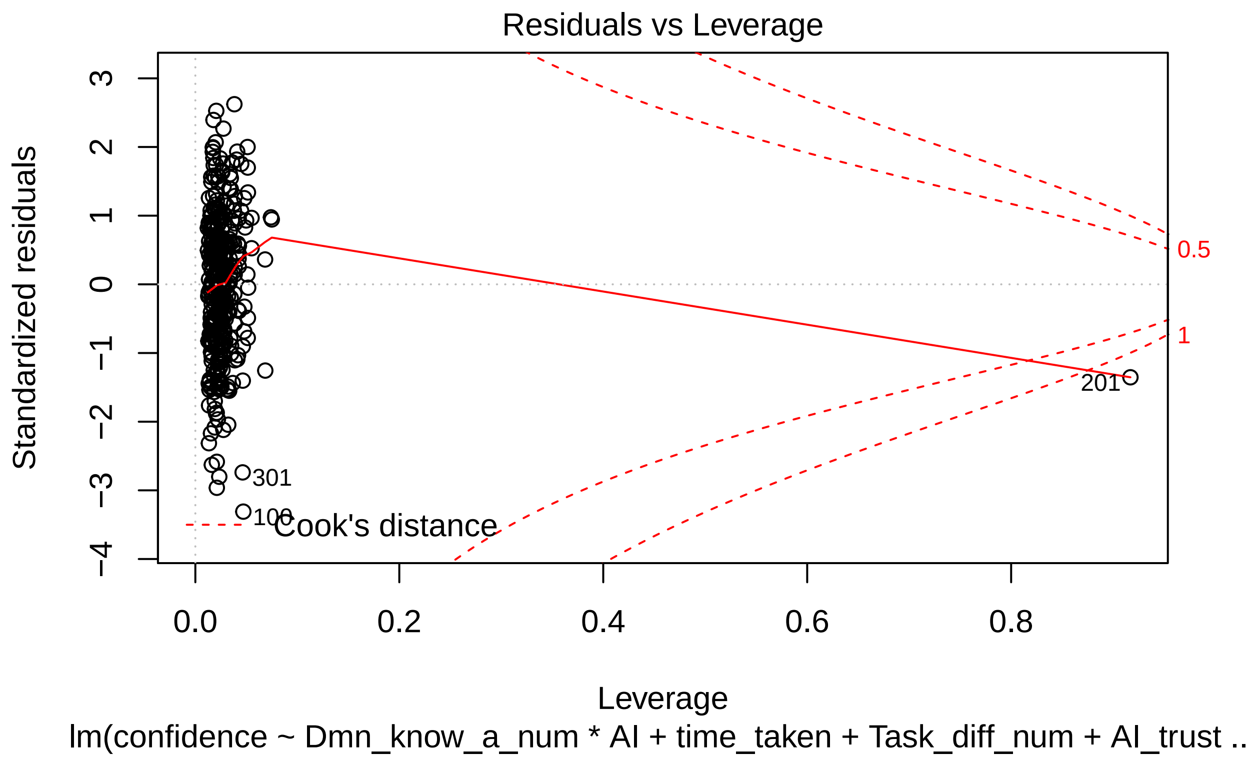




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(confidence ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust ..
253
115
100



Scale−Location
√|Standardized residuals|
Fitted values
lm(confidence ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust ..
100
115
253



Effect of AI recommendations on confidence with plants domain knowledge

Domain knowledge, AI recommendations, Attention checks, and log(age) were alll significantly affecting the confidence of the participants. The interaction term is not significant. Domain knowledge and AI recommendations affect the participants' confidence positively. $F(10, 387) = 17.07$, $p < 0.001$, $R\hat{}2 = 0.29$.

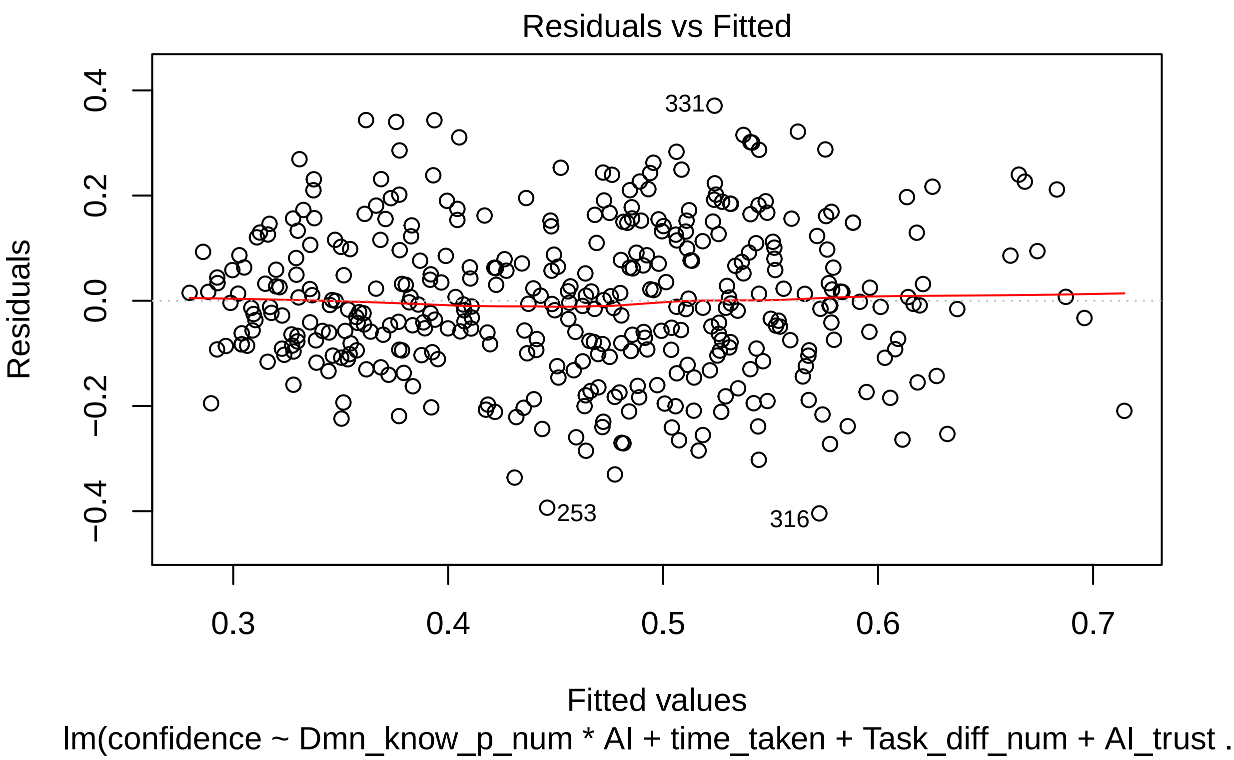
```r
lm.5.p.conf <- lm(confidence ~ Dmn_know_p_num*AI +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = plants_person)

summary(lm.5.p.conf)
```

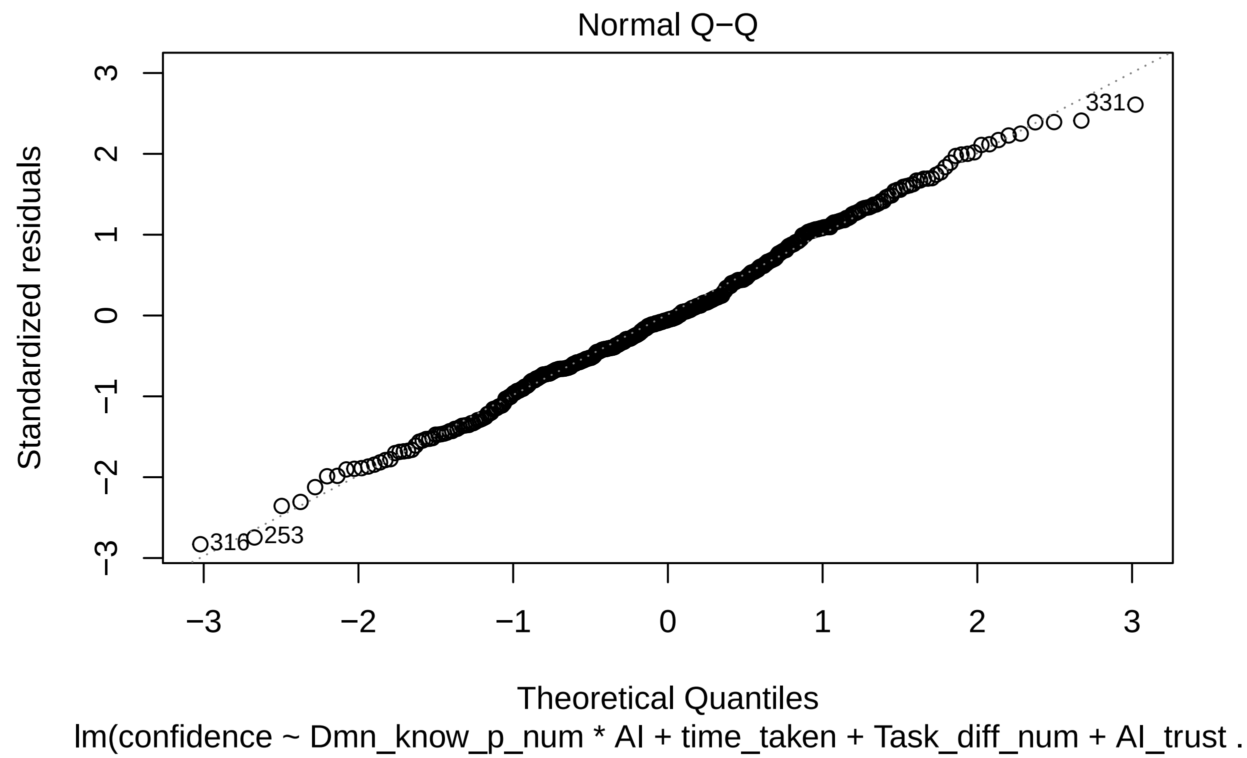
```
##
## Call:
## lm(formula = confidence ~ Dmn_know_p_num * AI + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = plants_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.40421 -0.09315 -0.00763  0.09802  0.37086
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.5501043  0.0897440   6.130 2.17e-09 ***
## Dmn_know_p_num     0.2362653  0.0507220   4.658 4.40e-06 ***
## AI                 0.1575158  0.0197454   7.977 1.71e-14 ***
## time_taken         0.0013794  0.0009869   1.398   0.1630
## Task_diff_num     -0.0129370  0.0073739  -1.754   0.0801 .
## AI_trust_num       0.0086578  0.0073564   1.177   0.2400
```

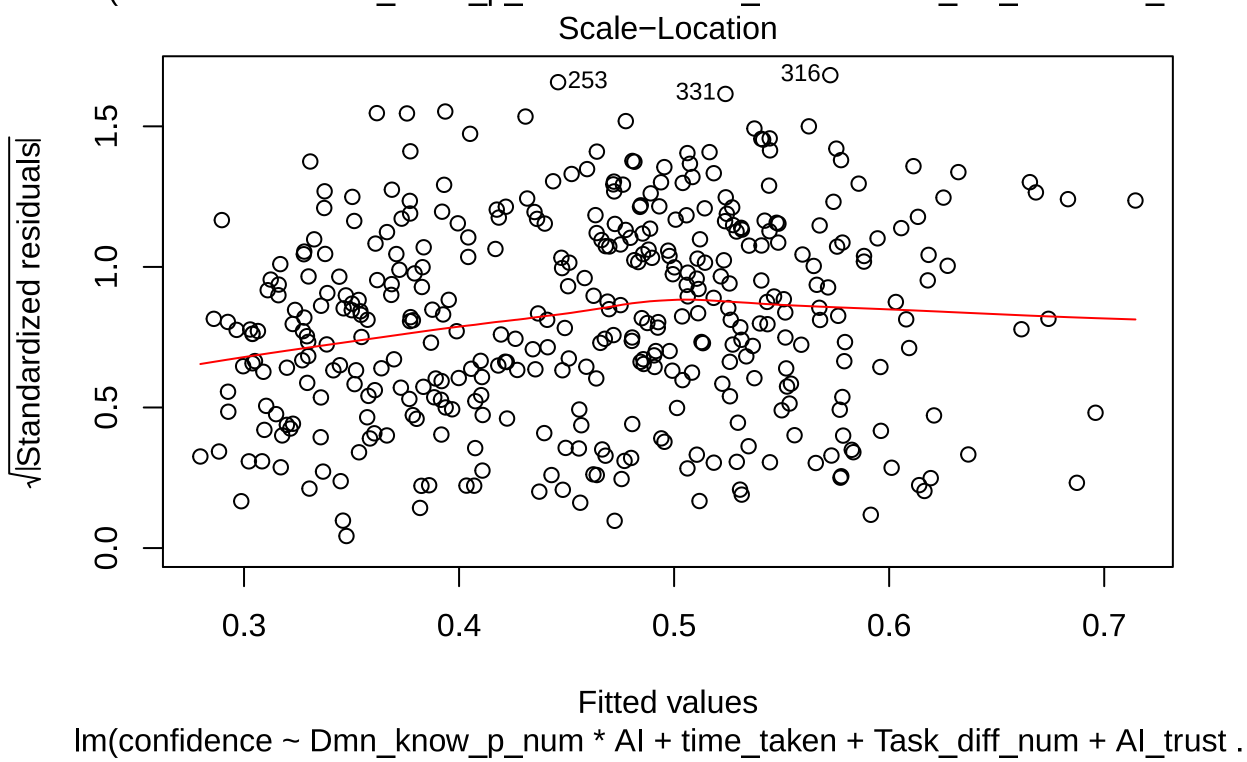
```
## atn_ch             -0.0406700  0.0169791  -2.395   0.0171 *
## log(age)           -0.0604544  0.0248544  -2.432   0.0155 *
## male_num            0.0065493  0.0147485   0.444   0.6572
## college             0.0278005  0.0161155   1.725   0.0853 .
## Dmn_know_p_num:AI -0.0378444  0.0687169  -0.551   0.5821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1448 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.3061, Adjusted R-squared:  0.2882
## F-statistic: 17.07 on 10 and 387 DF,  p-value: < 2.2e-16
```

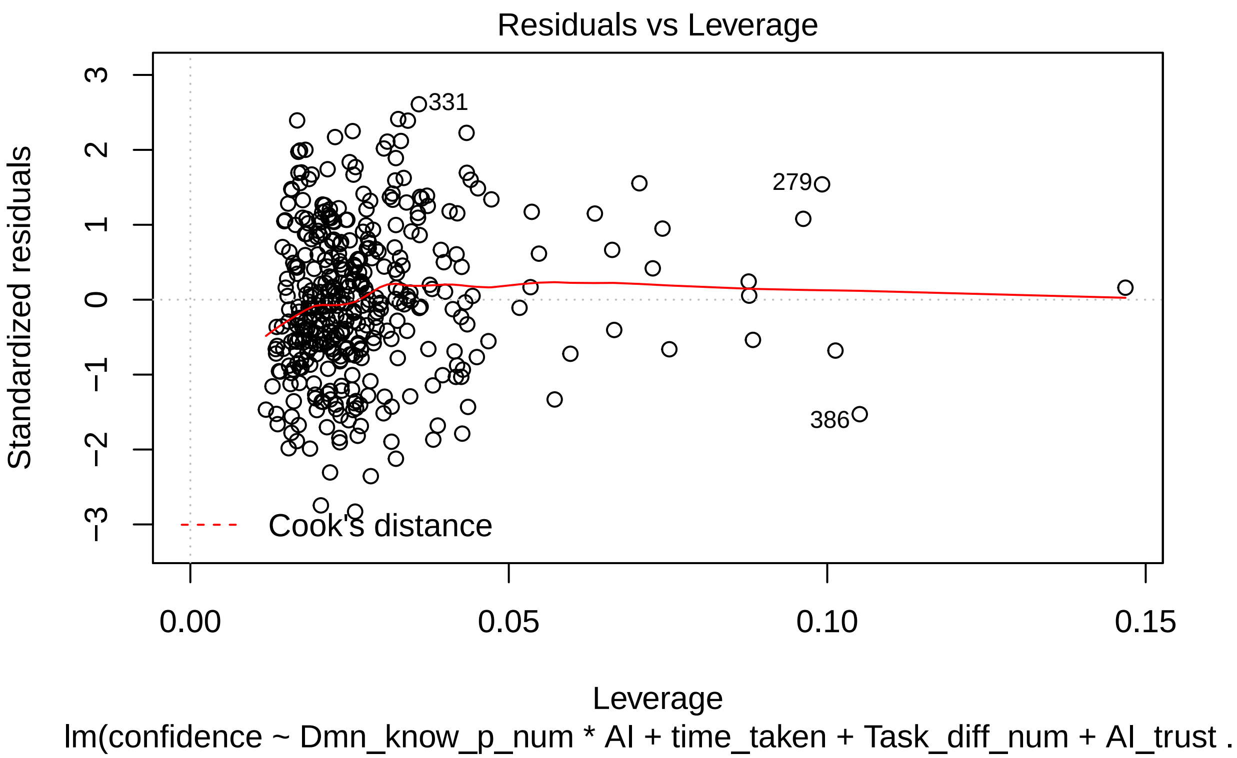
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.5.p.conf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
331
316
253
lm(confidence ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust ..


Scale−Location
√|Standardized residuals|
Fitted values
253
331
316
lm(confidence ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust ..



Effect of Uncertainty Information on confidence with animal domain knowledge

Domain knowledge, Uncertainty information, perceived AI usefulness rating, task difficulty rating, education level, and the interaction are all significantly affecting the participants' confidence. However, task difficulty rating and the interaction term is negatively affecting their confidence.

F(11, 187) = 14.07, p < 0.001, R^2 = 0.42

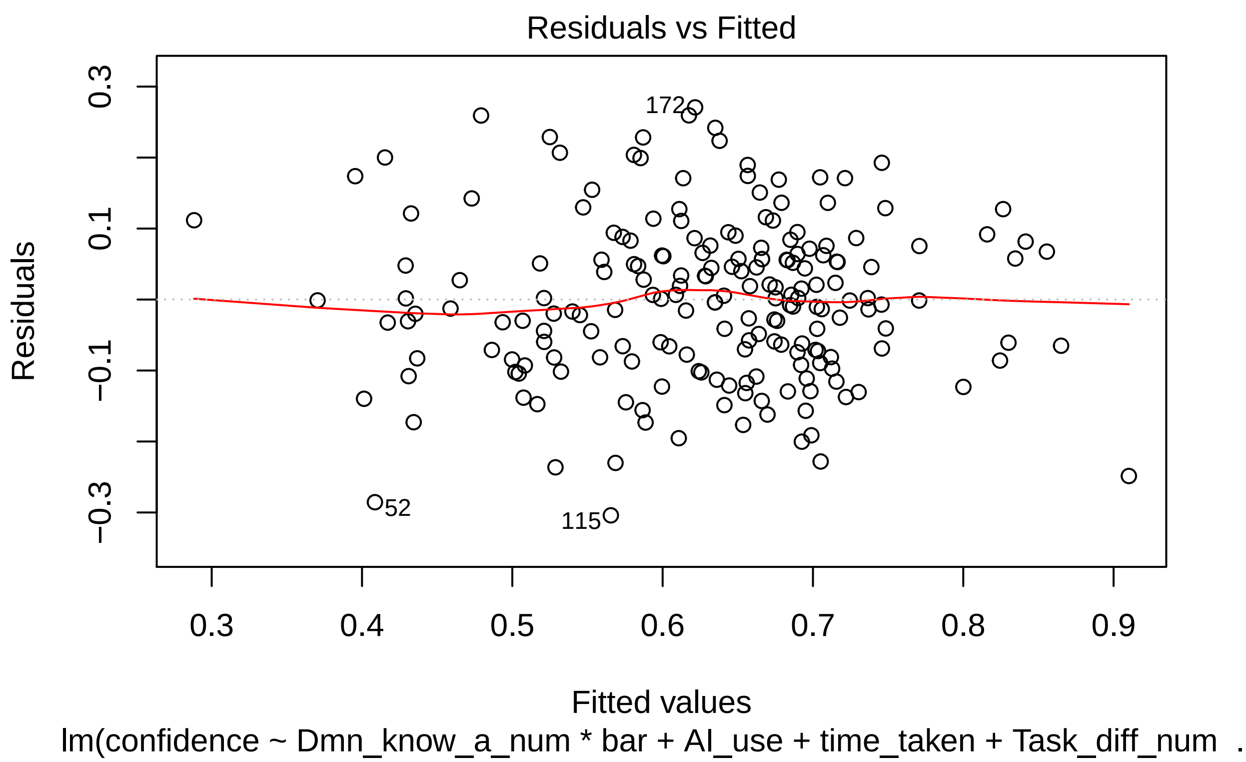
```
lm.6.a.conf <- lm(confidence ~ Dmn_know_a_num*bar + AI_use +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = animals_person_AI)

summary(lm.6.a.conf)
```

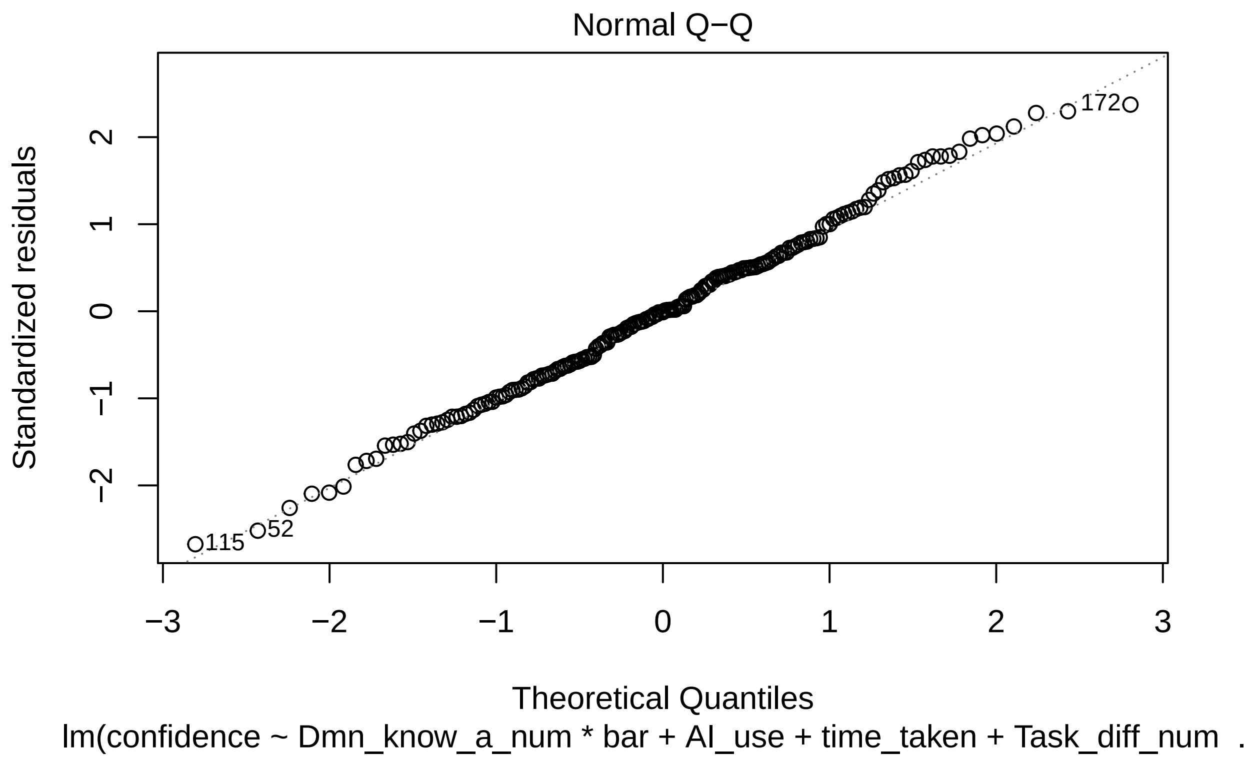
```
##
## Call:
## lm(formula = confidence ~ Dmn_know_a_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = animals_person_AI)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.304023 -0.081288 -0.001216  0.069500  0.270699
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.458973   0.109507   4.191 4.27e-05 ***
## Dmn_know_a_num      0.276650   0.049943   5.539 1.02e-07 ***
## bar                 0.109259   0.038841   2.813  0.00543 **
## AI_use              0.477781   0.052088   9.173  < 2e-16 ***
## time_taken          0.001875   0.001678   1.117  0.26536
```

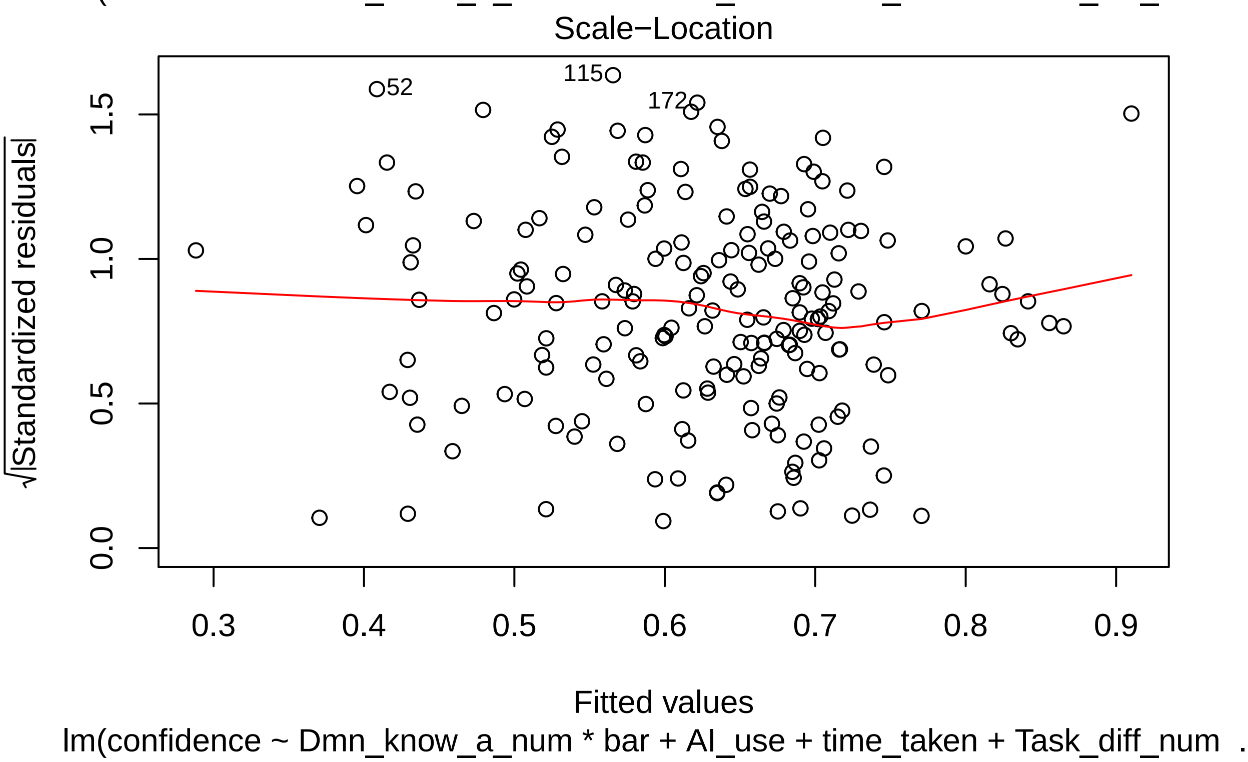
```
## Task_diff_num      -0.020096   0.008315  -2.417  0.01662 *
## AI_trust_num       -0.009420   0.008781  -1.073  0.28477
## atn_ch             -0.010168   0.020053  -0.507  0.61271
## log(age)           -0.054297   0.028403  -1.912  0.05744 .
## male_num            0.015489   0.016845   0.920  0.35901
## college             0.040786   0.018507   2.204  0.02876 *
## Dmn_know_a_num:bar -0.213292   0.072340  -2.948  0.00360 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1162 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.4529, Adjusted R-squared:  0.4207
## F-statistic: 14.07 on 11 and 187 DF,  p-value: < 2.2e-16
```

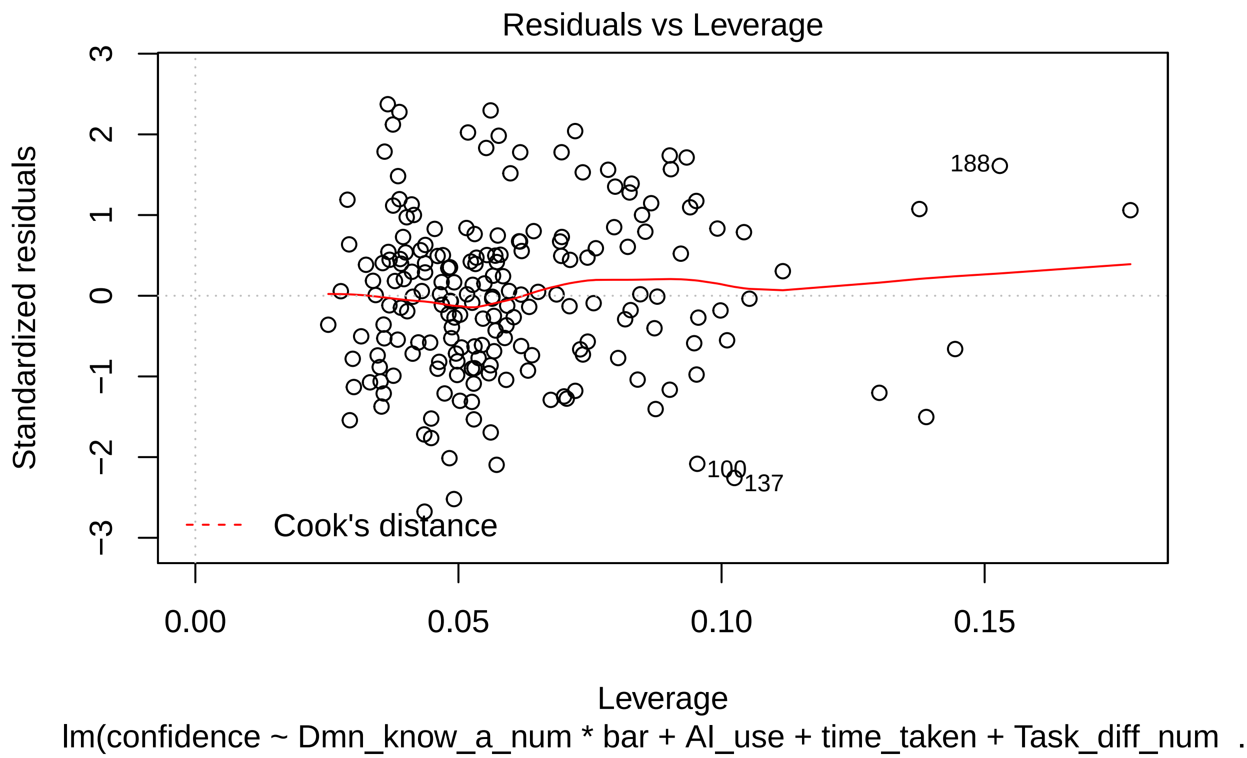
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.6.a.conf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(confidence ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num .
115
52
172


Scale−Location
√|Standardized residuals|
Fitted values
lm(confidence ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num .
52
115
172



Effect of Uncertainty Information on confidence with plant domain knowledge

Domain knowledge, perceived AI usefulness rating, log(age), and the interaction between plant domain knowledge and uncertainty information are all significantly affecting the participants' confidence. Similar to previous models, the interaction term decreases the confidence of the participants as the domain knowledge rating goes up when uncertainty information is provided.

$F(11, 187) = 13.13$, $p < 0.001$, $R^2 = 0.40$

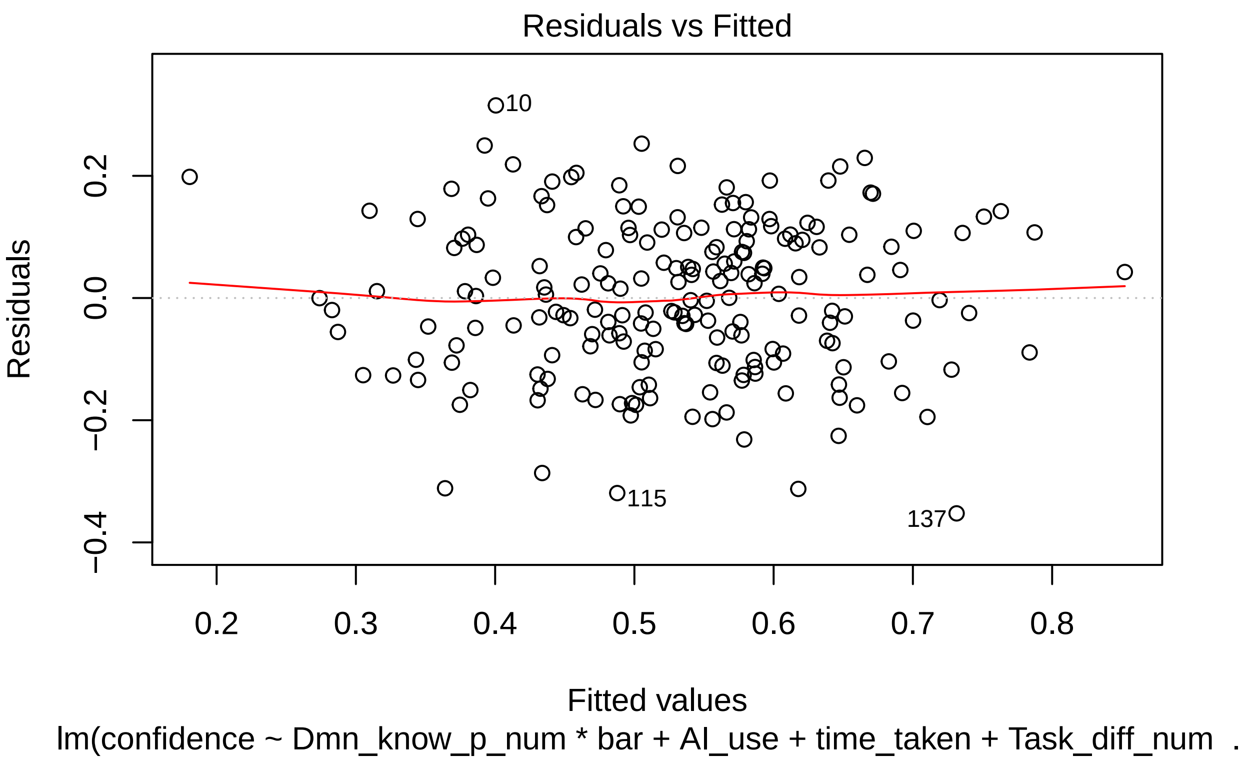
```
lm.6.p.conf <- lm(confidence ~ Dmn_know_p_num*bar + AI_use +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = plants_person_AI)

summary(lm.6.p.conf)
```

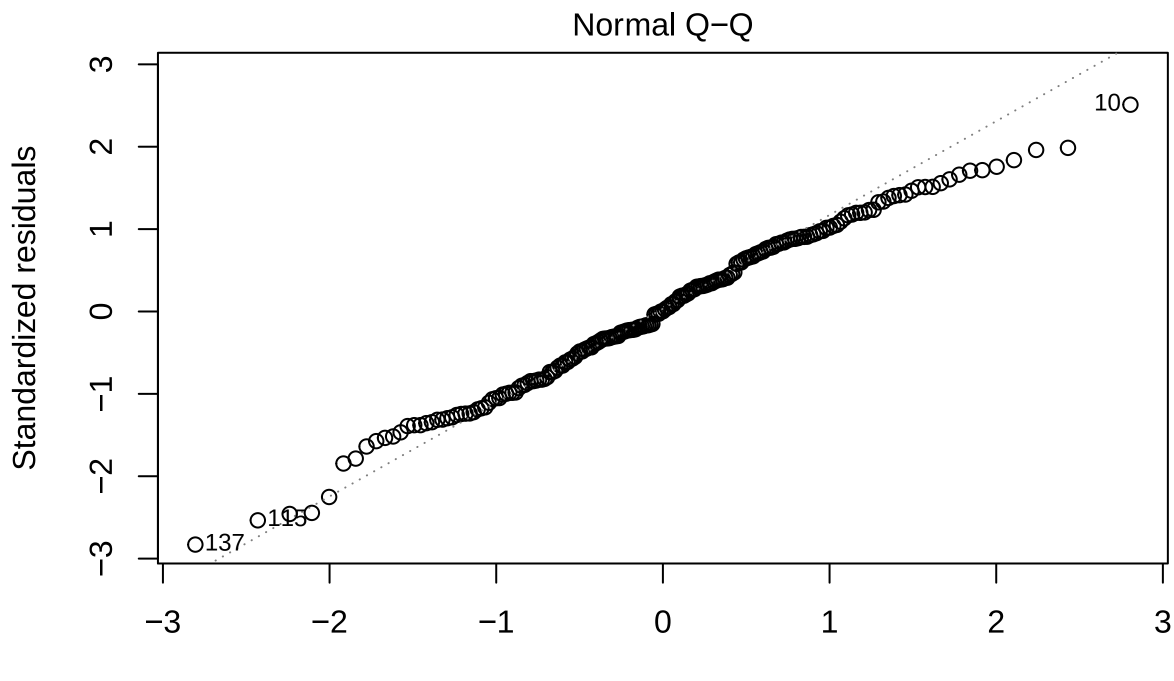
```
##
## Call:
## lm(formula = confidence ~ Dmn_know_p_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = plants_person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35245 -0.09227  0.00032  0.10144  0.31525
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.510022   0.119327   4.274 3.05e-05 ***
## Dmn_know_p_num      0.188807   0.059090   3.195  0.00164 **
## bar                 0.048500   0.026550   1.827  0.06933 .
## AI_use              0.547938   0.061584   8.897 4.89e-16 ***
```

```
## time_taken          0.001233   0.001367   0.902  0.36835
## Task_diff_num      -0.007834   0.009441  -0.830  0.40771
## AI_trust_num       -0.016128   0.010101  -1.597  0.11202
## atn_ch             -0.038573   0.022402  -1.722  0.08676 .
## log(age)           -0.071047   0.032048  -2.217  0.02784 *
## male_num            0.010547   0.019044   0.554  0.58036
## college             0.023398   0.020688   1.131  0.25951
## Dmn_know_p_num:bar -0.225066   0.097228  -2.315  0.02171 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1299 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.4358, Adjusted R-squared:  0.4026
## F-statistic: 13.13 on 11 and 187 DF,  p-value: < 2.2e-16
```

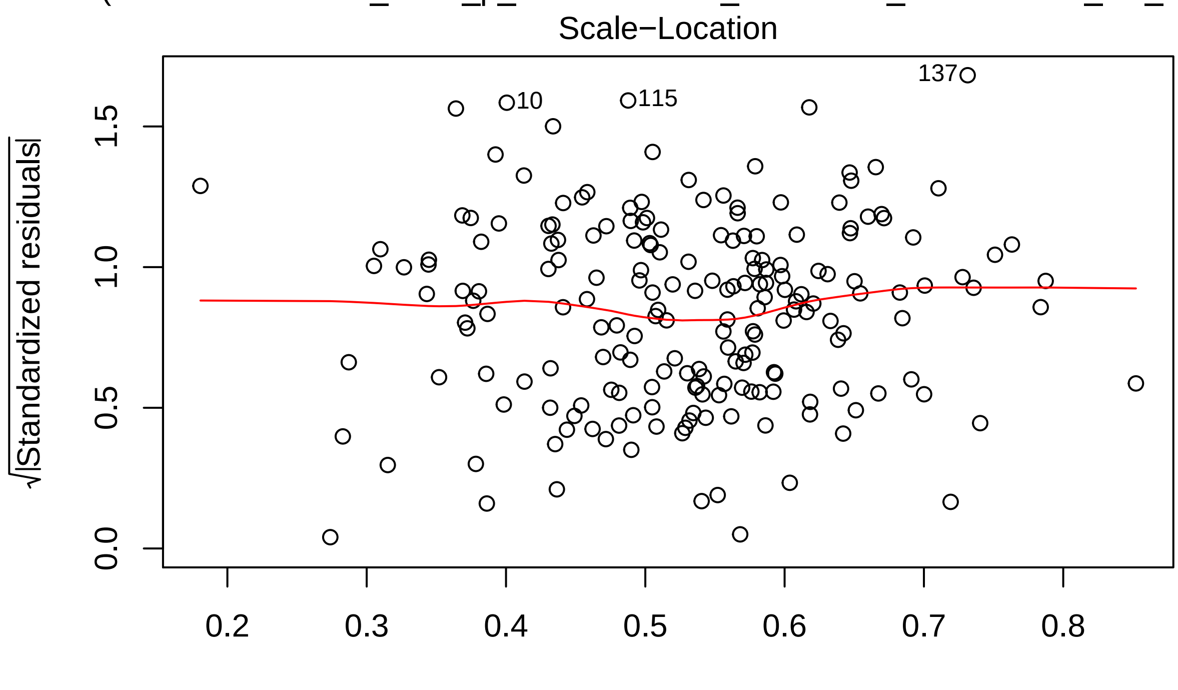
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.6.p.conf)
```

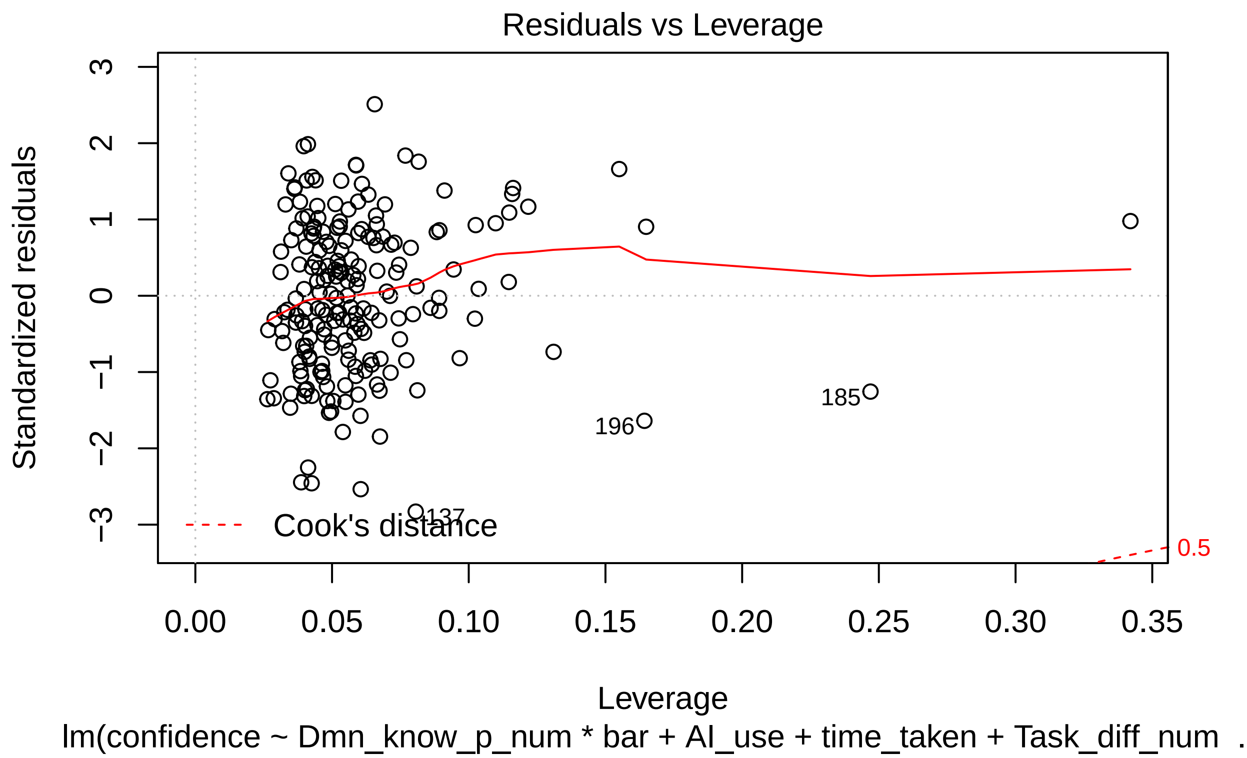




lm(confidence ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num .



lm(confidence ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num .



LINEAR MODELS ON OVERCONFIDENCE

Effect of AI on over confidence

Results of the simple linear regression indicate a negative significant relationship between AI recommendations and over confidence. it reduces it. (F(1,400) = 6.94, p <0.01, R2 = 0.01).

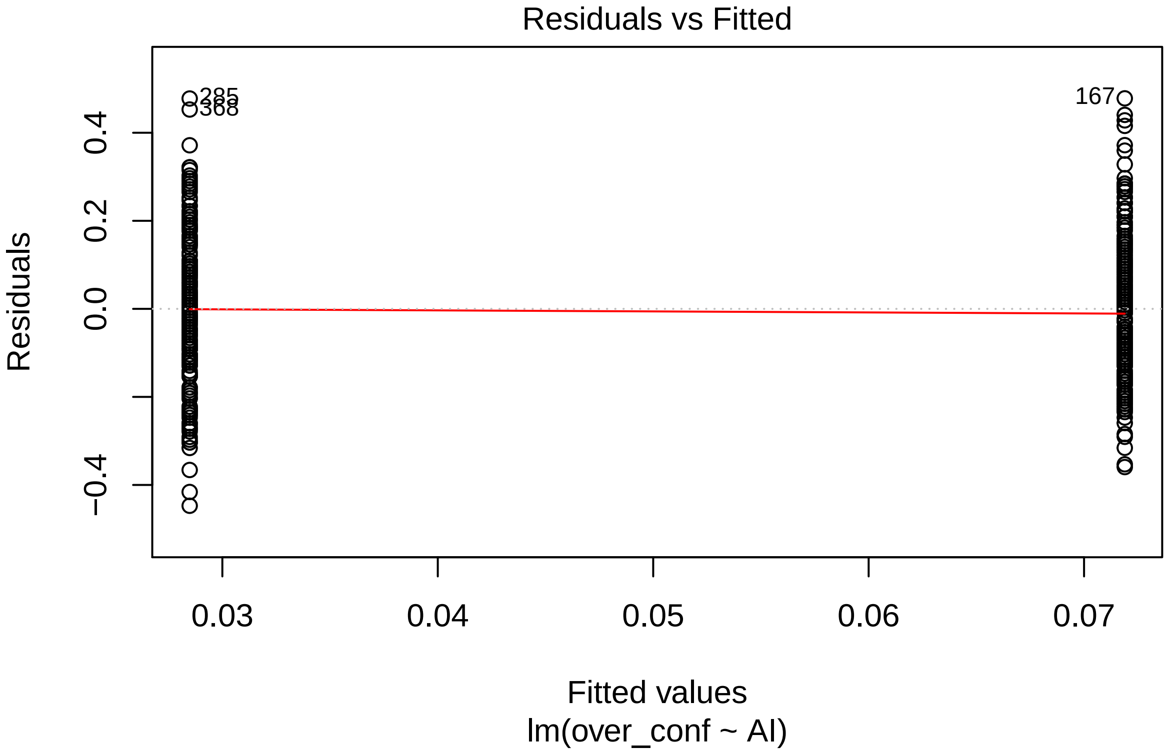
There is only so much variance that one predictor could explain. That's why r^2 is small below.

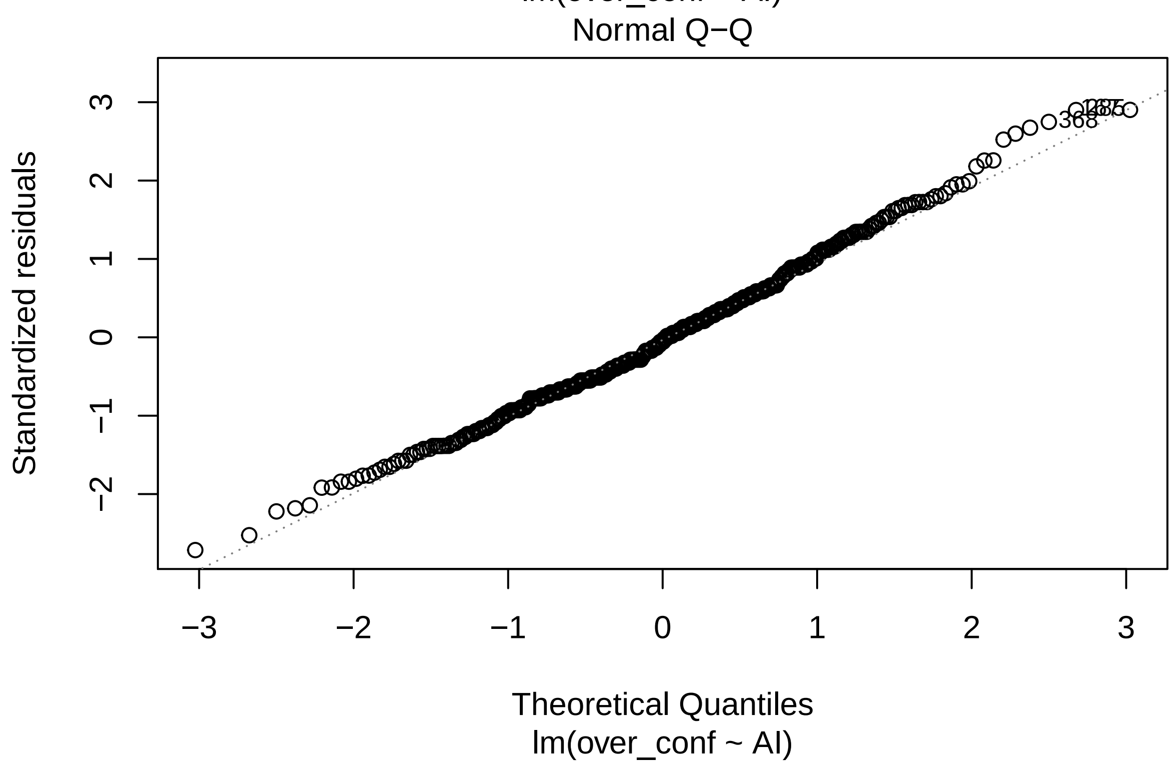
```
lm.1.overconf <- lm(over_conf ~ AI, data = person) #linear model
summary(lm.1.overconf)
```

```
##
## Call:
## lm(formula = over_conf ~ AI, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44723 -0.11416 -0.00939  0.10302  0.47811
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.07189    0.01165   6.171 1.66e-09 ***
## AI          -0.04341    0.01647  -2.635  0.00875 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1652 on 400 degrees of freedom
## Multiple R-squared:  0.01706,    Adjusted R-squared:  0.0146
## F-statistic: 6.942 on 1 and 400 DF,  p-value: 0.008746
```

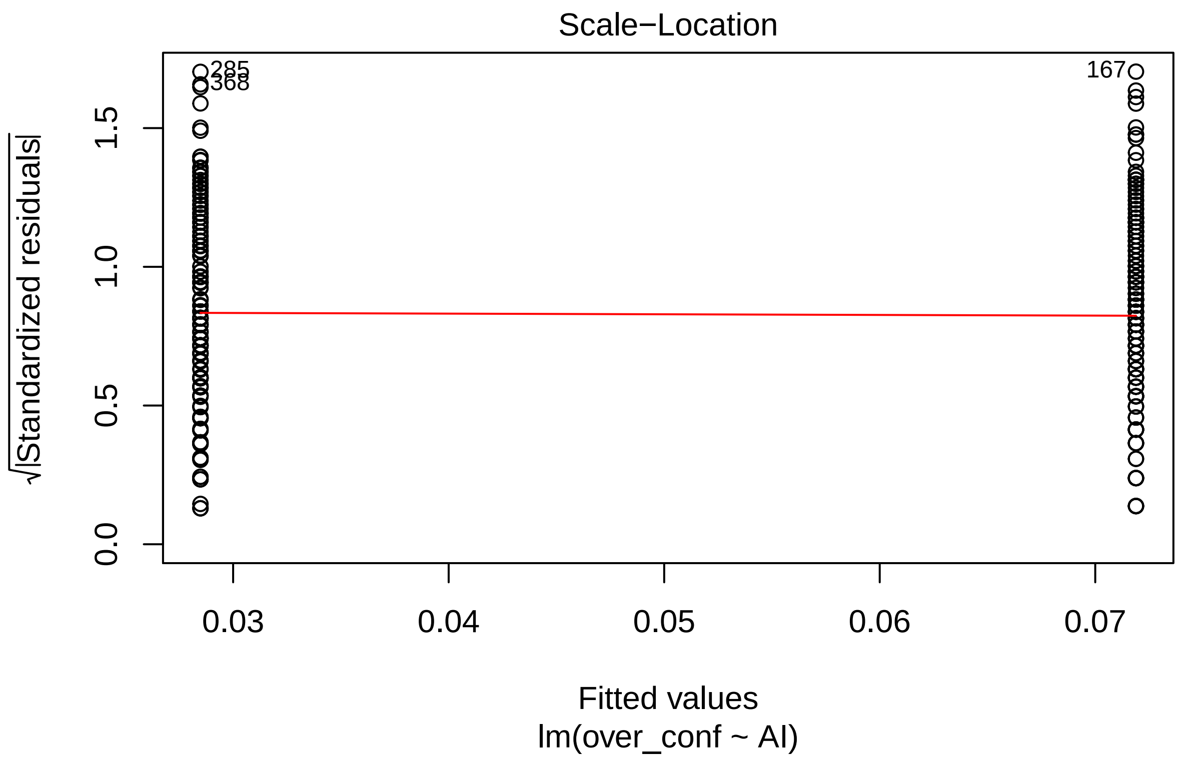
Given the predictor variable is binary, we see a pattern in the residuals vs fitted plot. Otherwise, the model is confeptable. The patter in the Q-Q plot is confeptable given the binary predictor variable, but it does deviate from the line at the edges.

```r
plot(lm.1.overconf)
```






```
## hat values (leverages) are all = 0.004975124
##  and there are no factor predictors; no plot no. 5
```



Effect of Uncertainty Information on over confidence

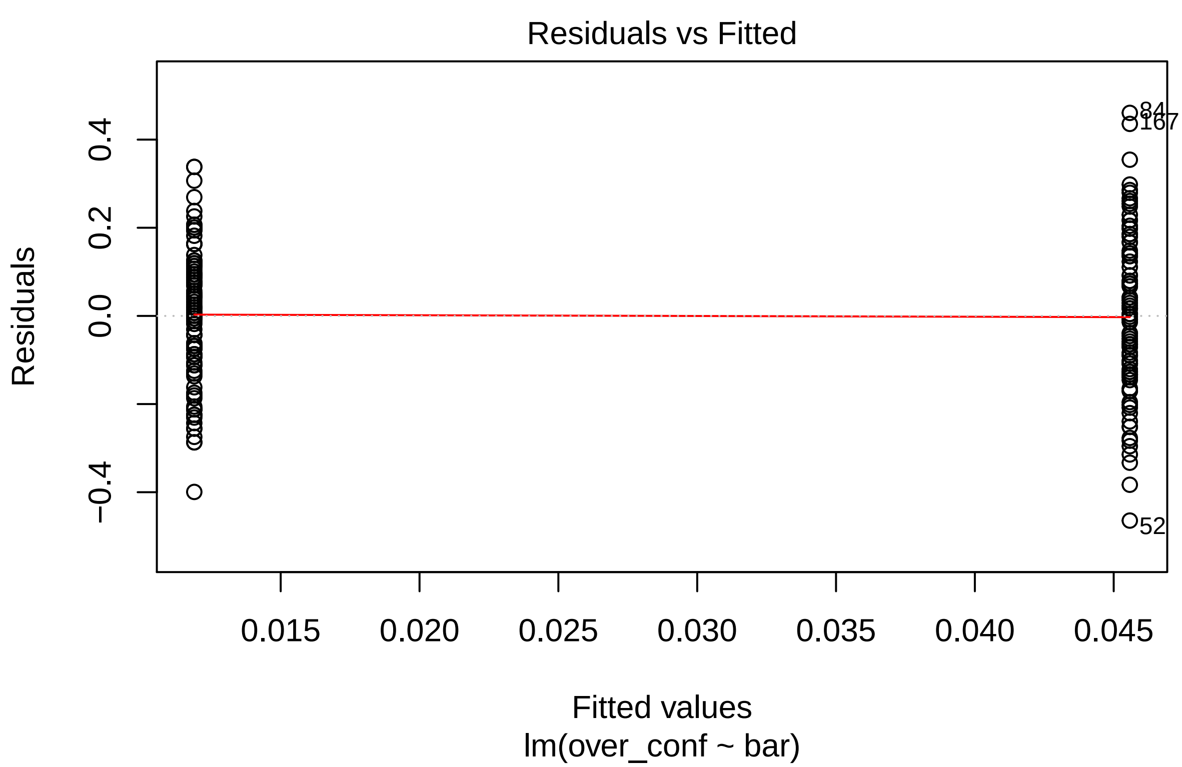
Results of the simple linear regression indicate an insignificant relationship between Unceraintty information and over confidence.

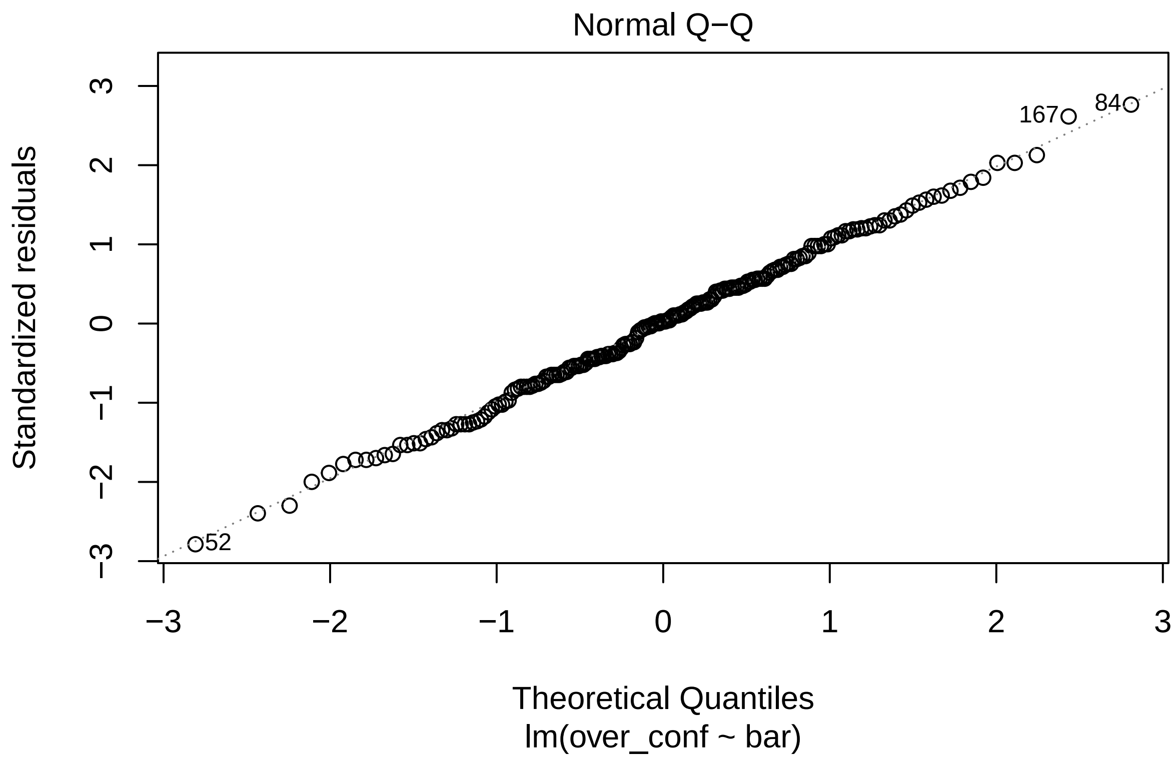
```r
lm.2.overconf <- lm(over_conf ~ bar, data = person_AI)

summary(lm.2.overconf)
```

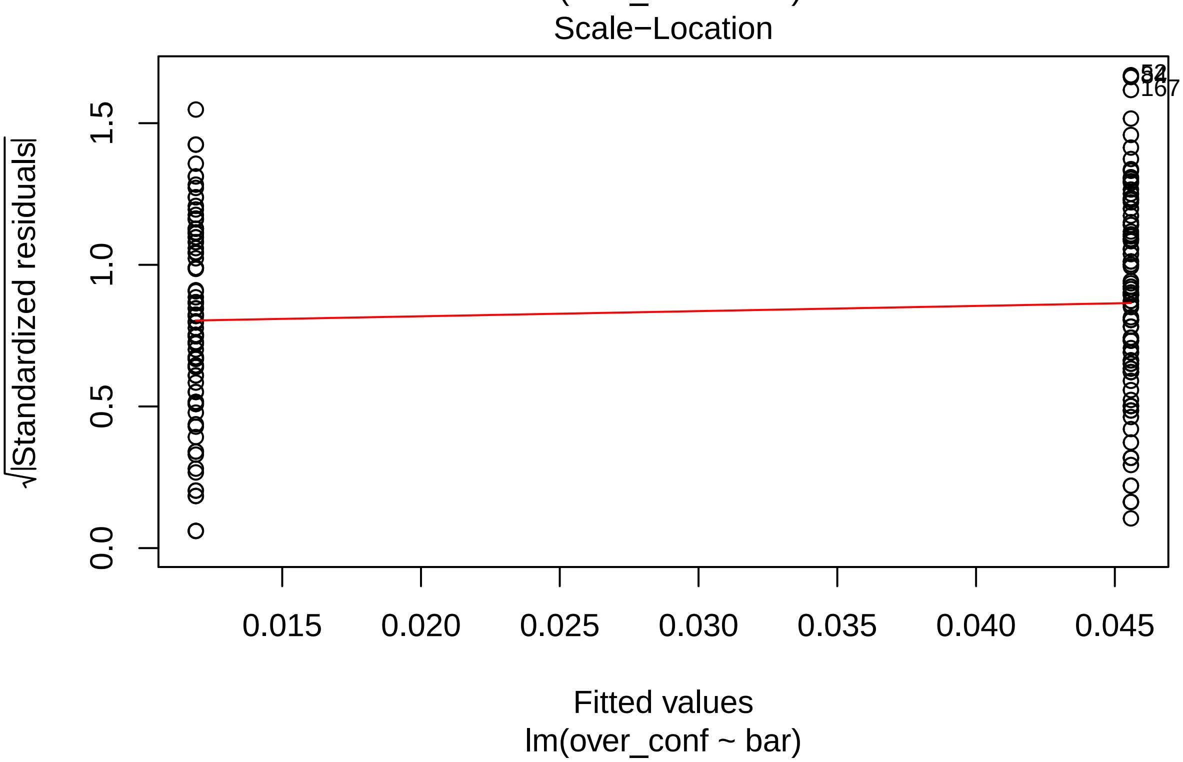
```
##
## Call:
## lm(formula = over_conf ~ bar, data = person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46433 -0.10808  0.00442  0.11311  0.46067
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.04558    0.01683   2.708  0.00736 **
## bar         -0.03369    0.02363  -1.426  0.15543
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1675 on 199 degrees of freedom
## Multiple R-squared:  0.01012,    Adjusted R-squared:  0.005141
## F-statistic: 2.033 on 1 and 199 DF,  p-value: 0.1554
```

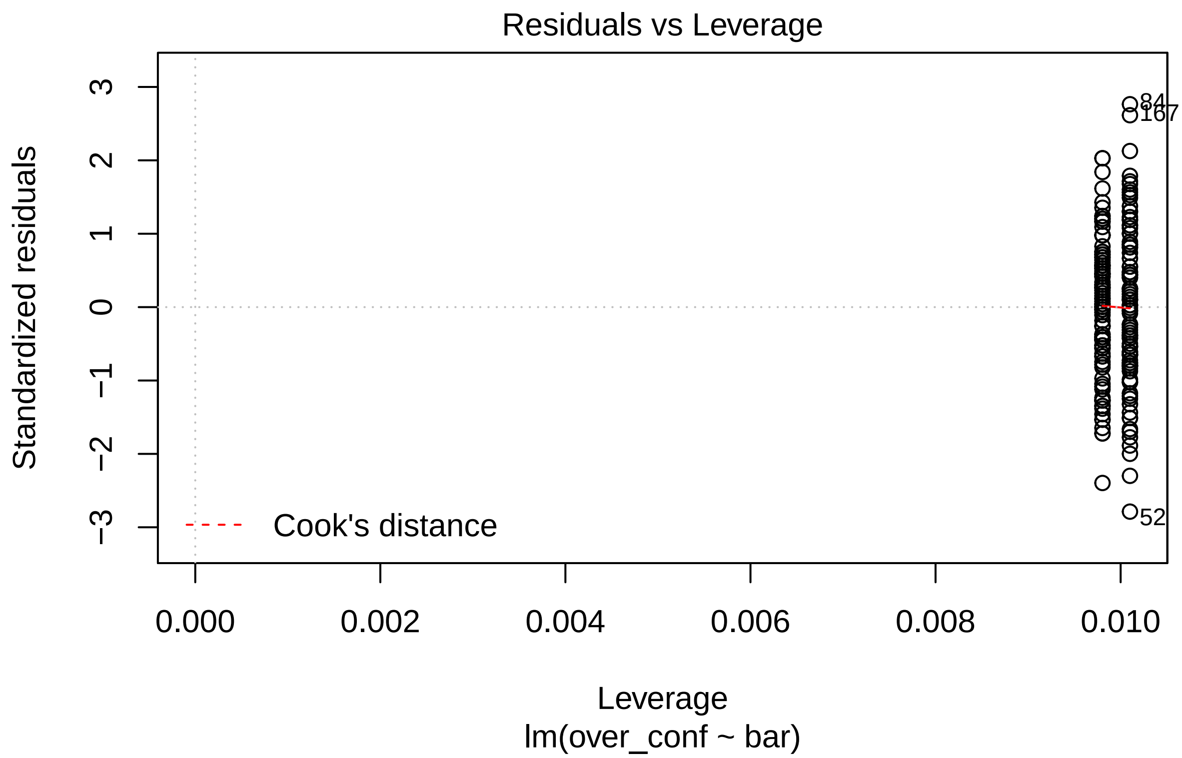
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.2.overconf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(over_conf ~ bar)
167
84
52


Scale−Location
√|Standardized residuals|
Fitted values
lm(over_conf ~ bar)
52
84
167



Effect of AI on over confidence with other predictor variables

AI recommendations, time taken, task difficulty, and attention check are all significant. More time taken by the participants increases their over confidence, however the provision of AI recommendations makes participant rationalize. Rest of the significant predictor variables all negatively affect over confidence. F(8, 389) = 4.87, p < 0.001, R^2 = 0.07

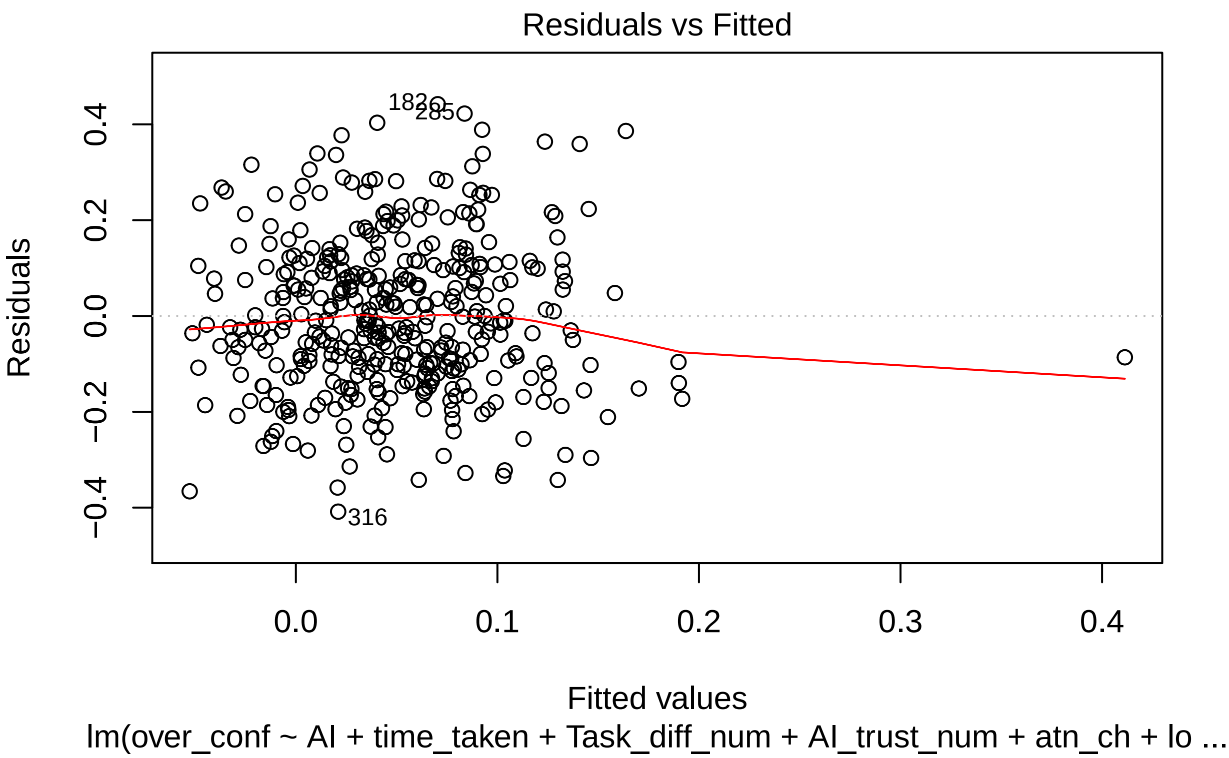
```r
lm.3.overconf <- lm(over_conf ~ AI + time_taken + Task_diff_num +
                  AI_trust_num + atn_ch + log(age) + male_num + college,
                data = person)

summary(lm.3.overconf)
```

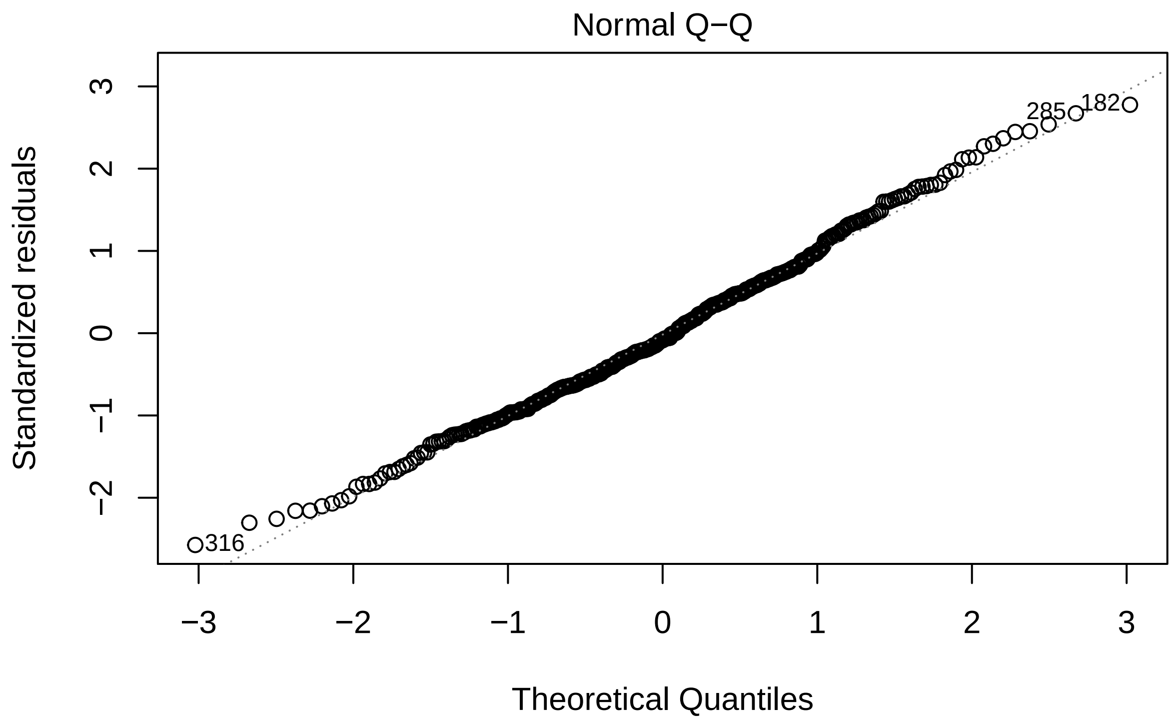
```
##
## Call:
## lm(formula = over_conf ~ AI + time_taken + Task_diff_num + AI_trust_num +
##     atn_ch + log(age) + male_num + college, data = person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.40849 -0.10751 -0.01345  0.10304  0.44211
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.207823   0.099170   2.096  0.03676 *
## AI            -0.036652   0.016372  -2.239  0.02574 *
## time_taken     0.001330   0.000582   2.286  0.02279 *
## Task_diff_num -0.022704   0.008049  -2.821  0.00503 **
## AI_trust_num   0.013187   0.008030   1.642  0.10136
## atn_ch        -0.049827   0.018558  -2.685  0.00756 **
```

```
## log(age)       -0.032891   0.027410  -1.200  0.23088
## male_num        0.025754   0.016354   1.575  0.11613
## college         0.008685   0.017697   0.491  0.62386
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1603 on 389 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.09104,    Adjusted R-squared:  0.07235
## F-statistic:  4.87 on 8 and 389 DF,  p-value: 9.654e-06
```

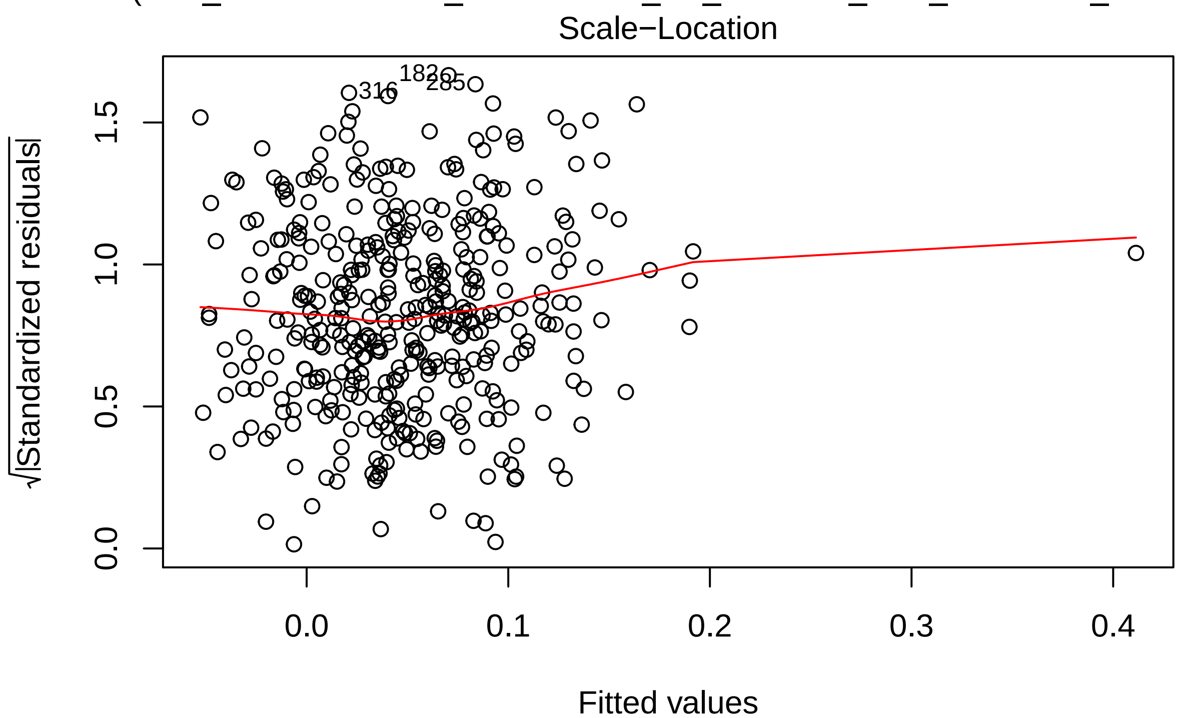
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.3.overconf)
```

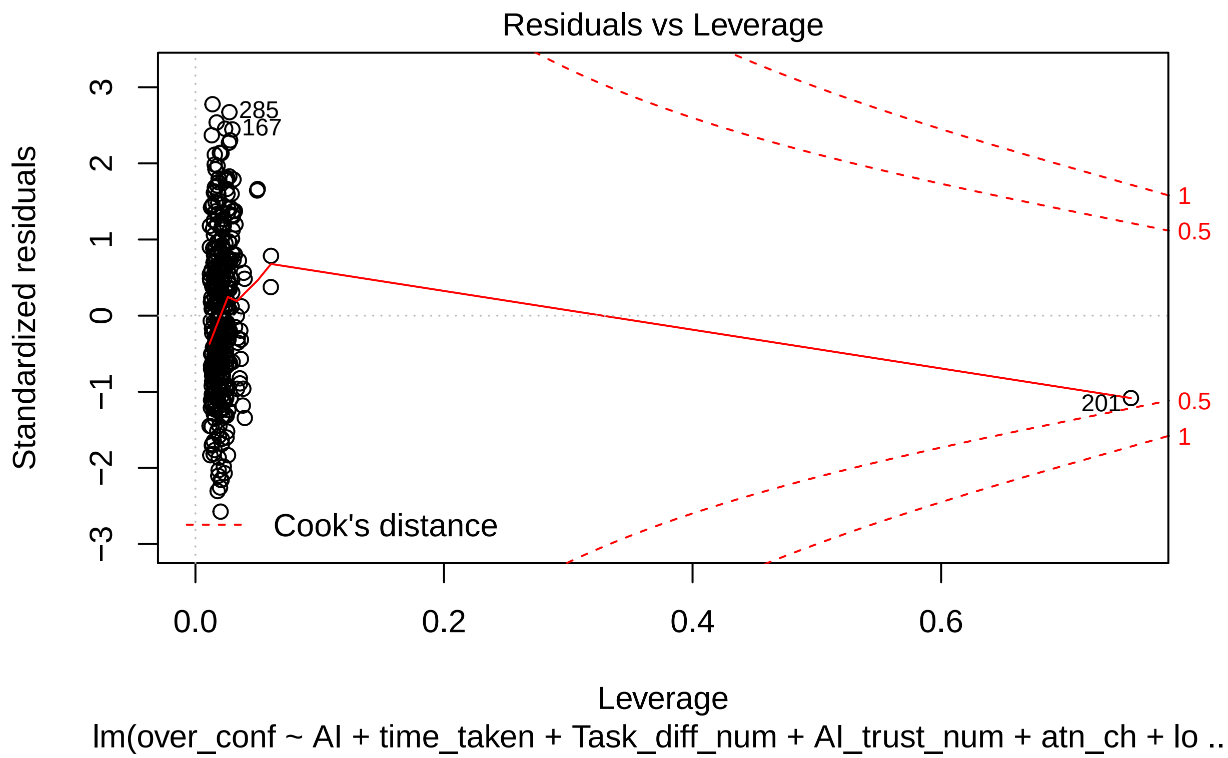




lm(over_conf ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + lo ...



lm(over_conf ~ AI + time_taken + Task_diff_num + AI_trust_num + atn_ch + lo ...



Effect of Uncertainty Information on over confidence with other predictor variables

Uncertainty information, perceived AI usefulness rating, and task difficulty are significantly affecting the participants over confidence. Provision of uncertainty information brings down their over confidence but when the participants AI usefulness increased, so did their confidence. So whenever they found the AI to be increasingly useful, they were more confident than accurate. $F(9, 189) = 8.59$, $p < 0.001$, $R^2 = 0.26$.

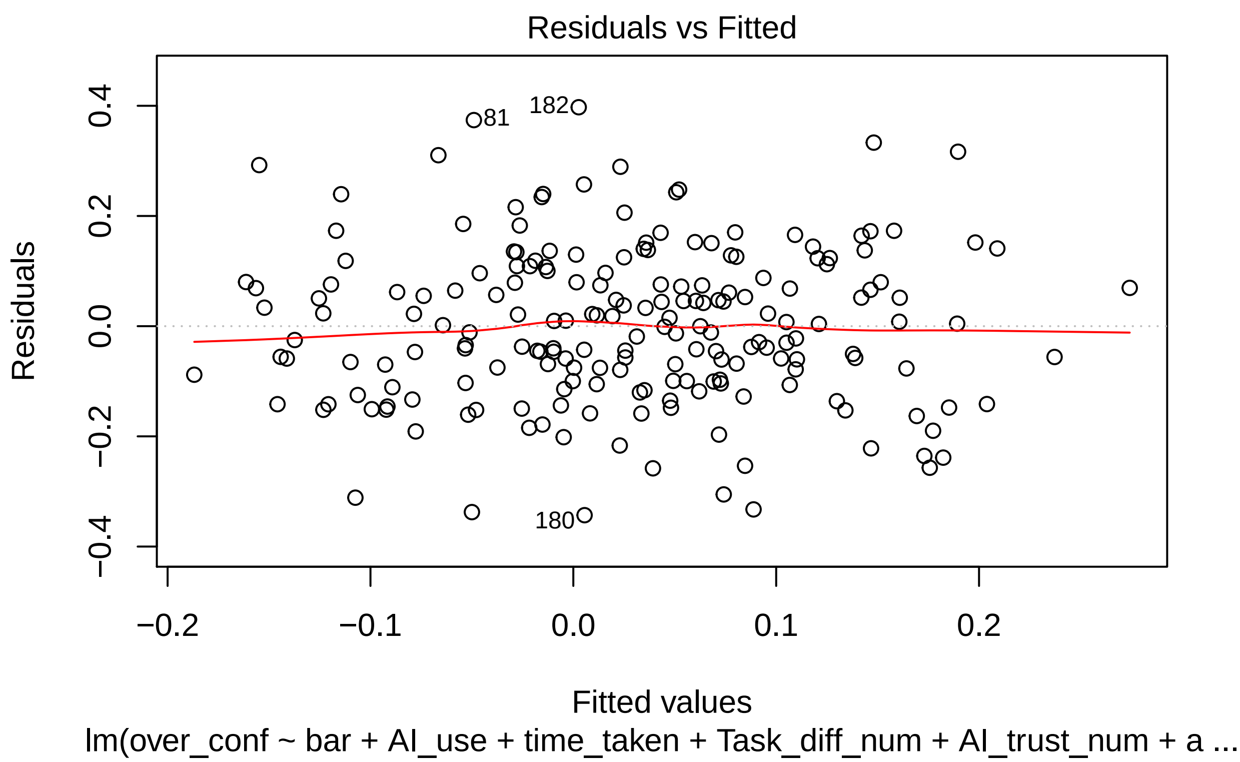
```
lm.4.overconf <- lm(over_conf ~ bar + AI_use + time_taken + Task_diff_num +
                    AI_trust_num + atn_ch + log(age) + male_num + college,
                data = person_AI)

summary(lm.4.overconf)
```

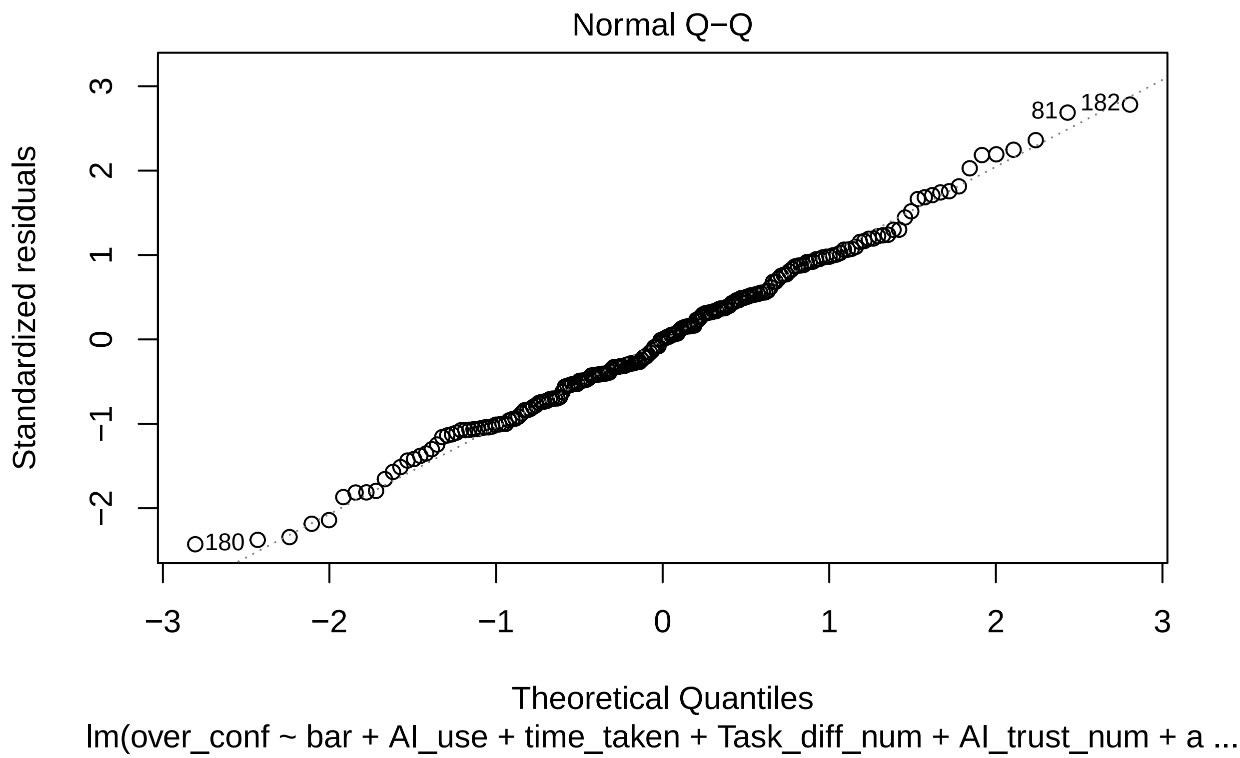
```
##
## Call:
## lm(formula = over_conf ~ bar + AI_use + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.34305 -0.10009 -0.00012  0.09639  0.39735
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.093616   0.132632   0.706  0.48116
## bar           -0.062835   0.021509  -2.921  0.00391 **
## AI_use         0.465409   0.069102   6.735 1.92e-10 ***
## time_taken     0.002433   0.001868   1.302  0.19439
## Task_diff_num -0.028780   0.010303  -2.793  0.00575 **
## AI_trust_num  -0.011672   0.011216  -1.041  0.29935
```

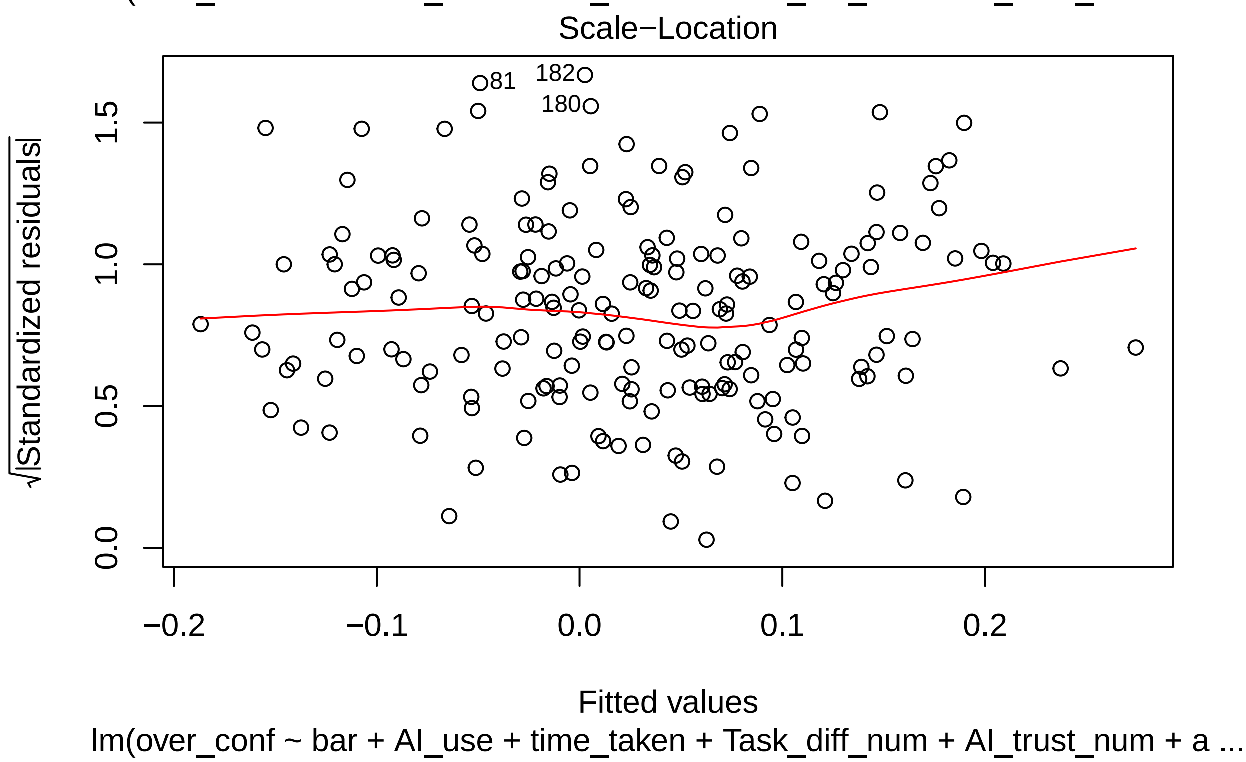
```
## atn_ch        -0.047198   0.024647  -1.915  0.05701 .
## log(age)      -0.050411   0.035299  -1.428  0.15492
## male_num       0.023191   0.021035   1.103  0.27164
## college        0.013236   0.022912   0.578  0.56415
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1451 on 189 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.2903, Adjusted R-squared:  0.2565
## F-statistic: 8.591 on 9 and 189 DF,  p-value: 9.318e-11
```

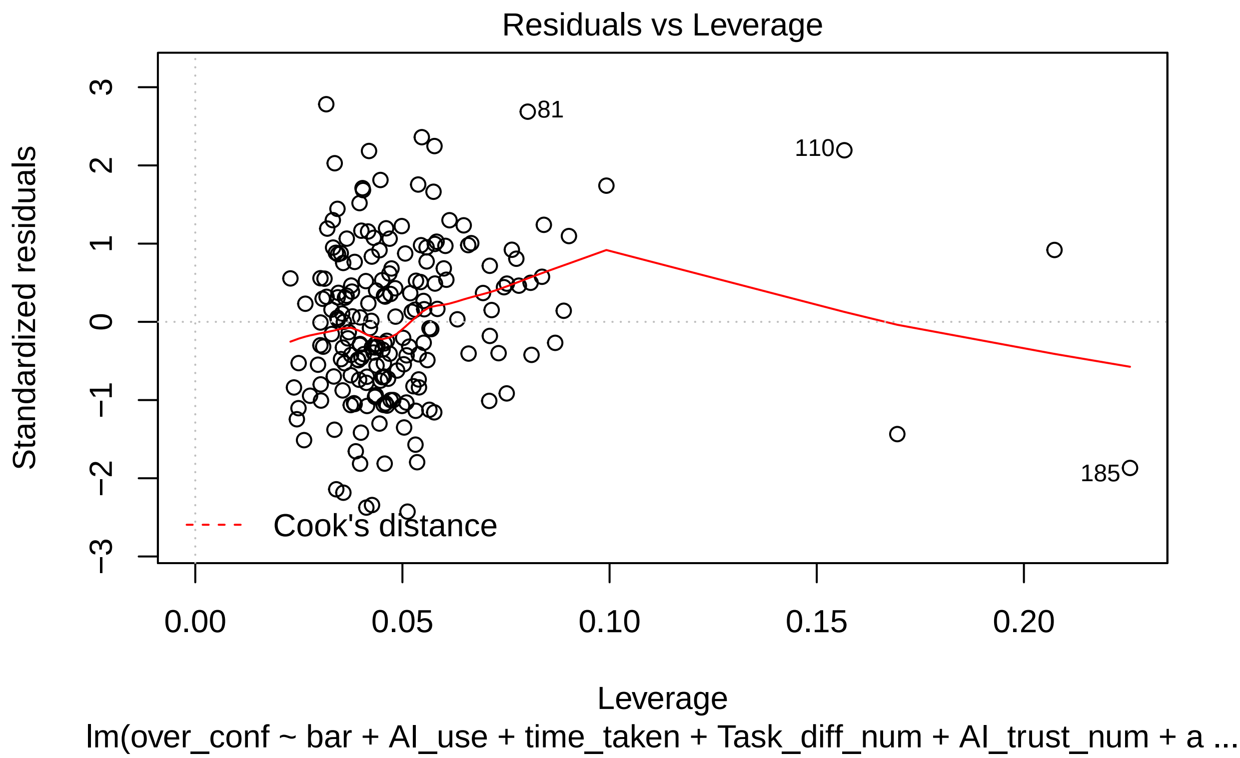
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.4.overconf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(over_conf ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + a ...
180
81
182


Scale−Location
√|Standardized residuals|
Fitted values
lm(over_conf ~ bar + AI_use + time_taken + Task_diff_num + AI_trust_num + a ...
81
182
180



Effect of AI recommendations on over confidence with animal domain knowledge

Domain knowledge, Task difficulty rating, age, and gender are all significantly affecting the over confidence of the participants.

F(10, 387) = 5.606, p < 0.001, R^2 0.10.

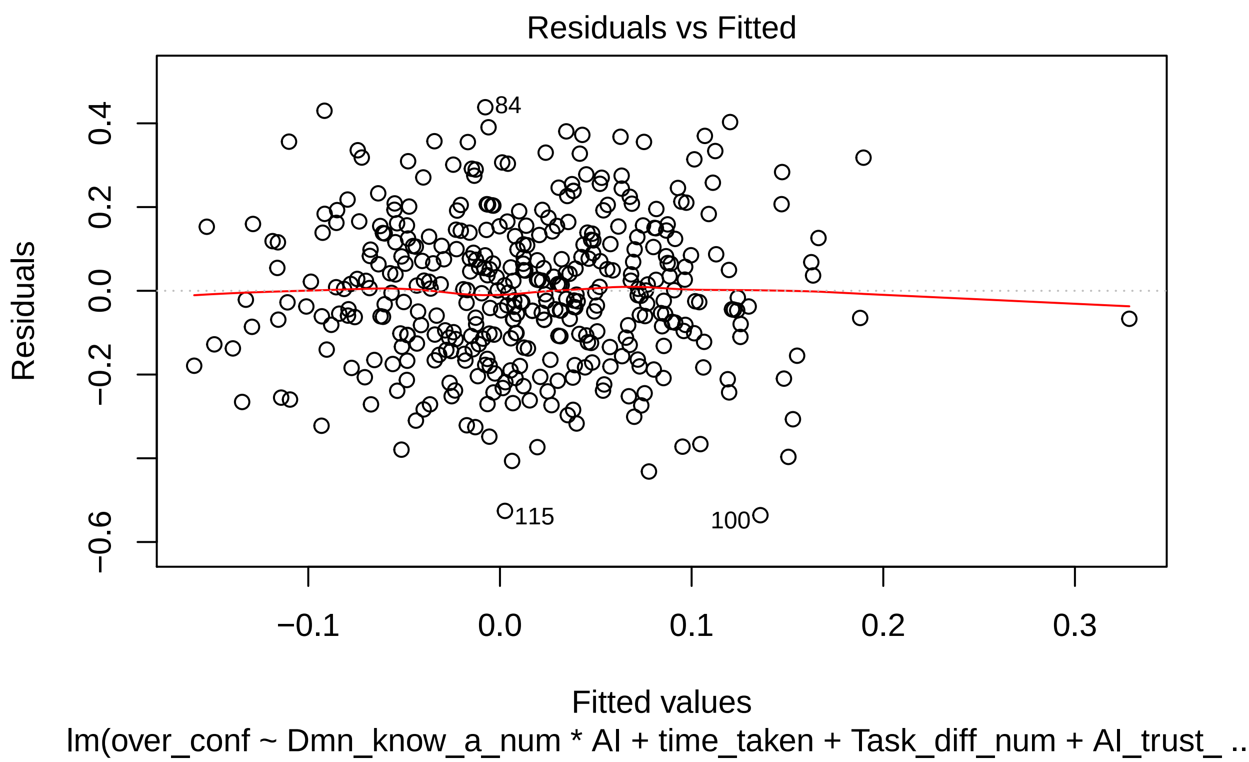
```
lm.5.a.overconf <- lm(over_conf ~ Dmn_know_a_num*AI +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = animals_person)

summary(lm.5.a.overconf)
```

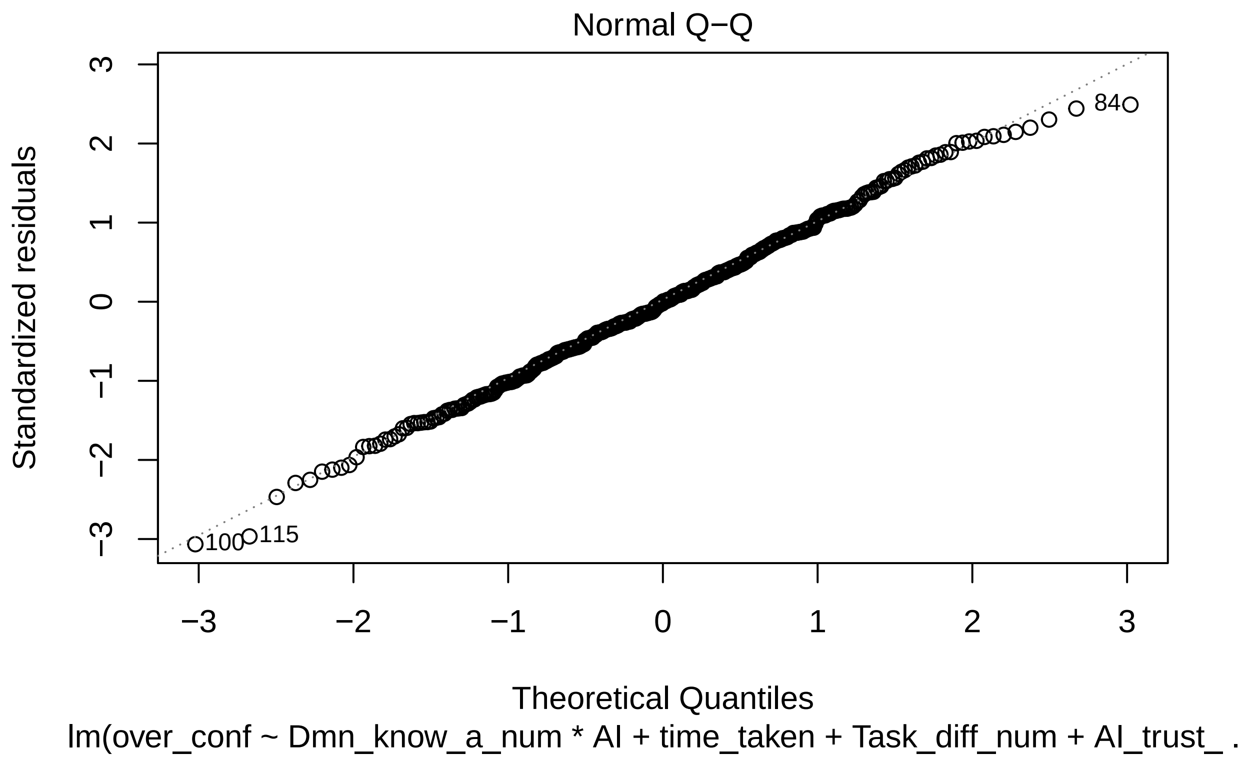
```
##
## Call:
## lm(formula = over_conf ~ Dmn_know_a_num * AI + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = animals_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53590 -0.11304 -0.00143  0.12342  0.43842
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.0769604  0.1159283   0.664   0.5072
## Dmn_know_a_num     0.2499023  0.0547266   4.566 6.68e-06 ***
## AI                 0.0724314  0.0407346   1.778   0.0762 .
## time_taken         0.0004785  0.0002947   1.624   0.1053
## Task_diff_num     -0.0194922  0.0090476  -2.154   0.0318 *
## AI_trust_num       0.0159935  0.0089497   1.787   0.0747 .
## atn_ch            -0.0254473  0.0210334  -1.210   0.2271
```

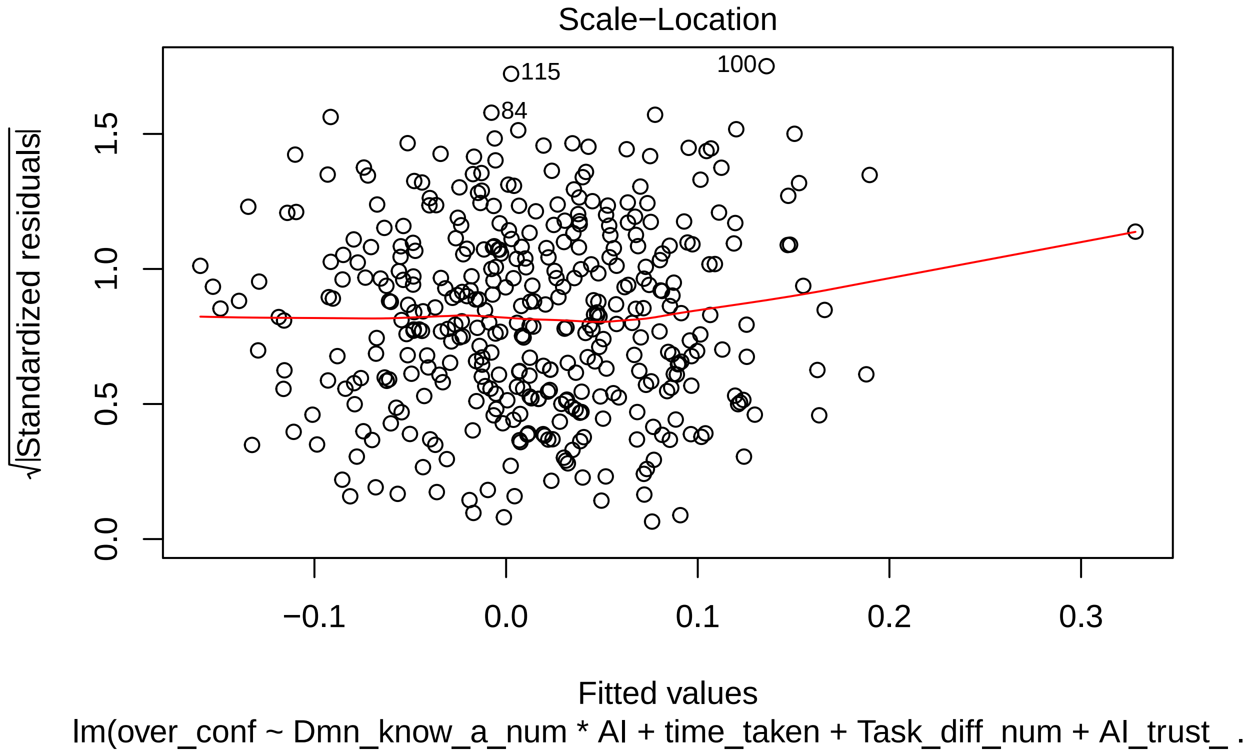
```
## log(age)           -0.0606656  0.0304576  -1.992   0.0471 *
## male_num            0.0460924  0.0182654   2.523   0.0120 *
## college             0.0292256  0.0197822   1.477   0.1404
## Dmn_know_a_num:AI  -0.1298575  0.0764049  -1.700   0.0900 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1791 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.1265, Adjusted R-squared:  0.104
## F-statistic: 5.606 on 10 and 387 DF,  p-value: 7.821e-08
```

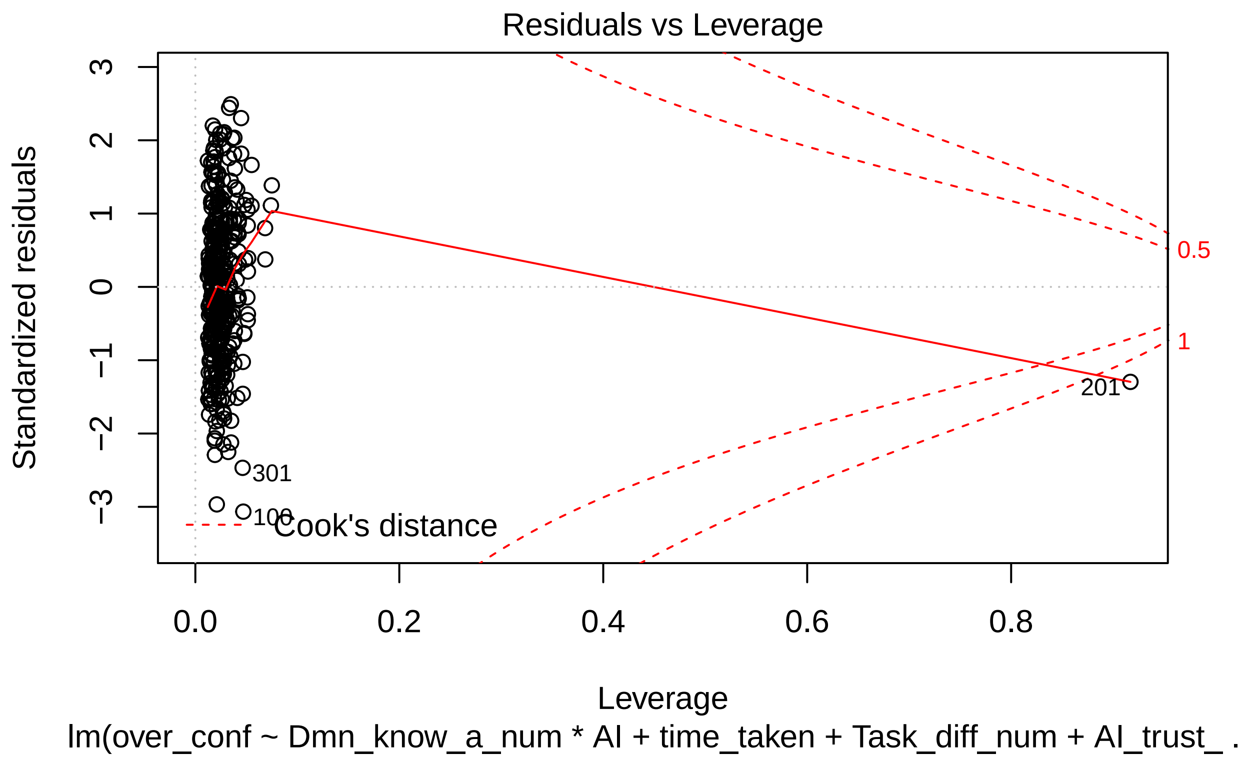
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.5.a.overconf)
```




Normal Q−Q
Standardized residuals
3
2
1
0
−1
−2
−3
84
100
115
−3
−2
−1
0
1
2
3
Theoretical Quantiles
lm(over_conf ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust_ ..


Scale−Location
√|Standardized residuals|
1.5
1.0
0.5
0.0
115
100
84
−0.1
0.0
0.1
0.2
0.3
Fitted values
lm(over_conf ~ Dmn_know_a_num * AI + time_taken + Task_diff_num + AI_trust_ ..



Effect of AI recommendations on over confidence with plants domain knowledge

Domain knowledge, task diffculty rating, log(age), and gender were all significantly affecting the over confidence of the participants. F(10, 387) = 4.08, p < 0.001, R^2 = 0.07.

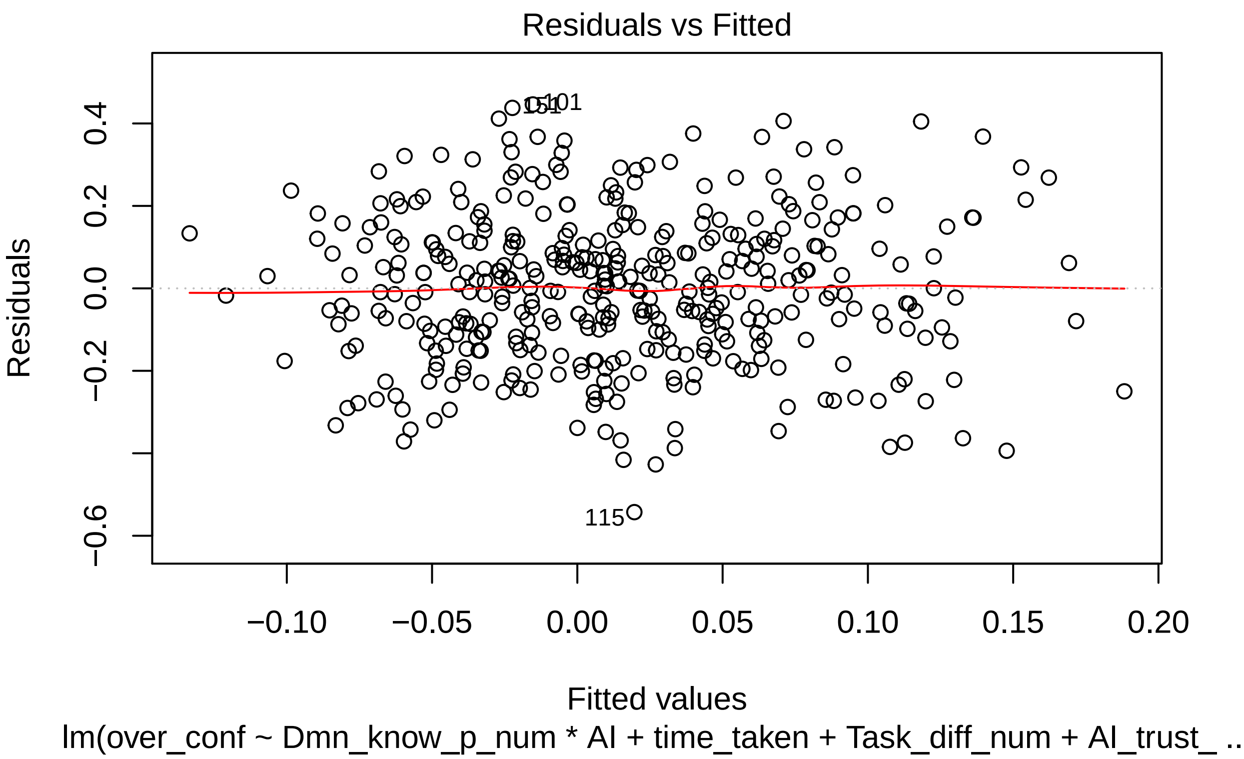
```
lm.5.p.overconf <- lm(over_conf ~ Dmn_know_p_num*AI +
                 time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                 male_num + college, data = plants_person)

summary(lm.5.p.overconf)
```

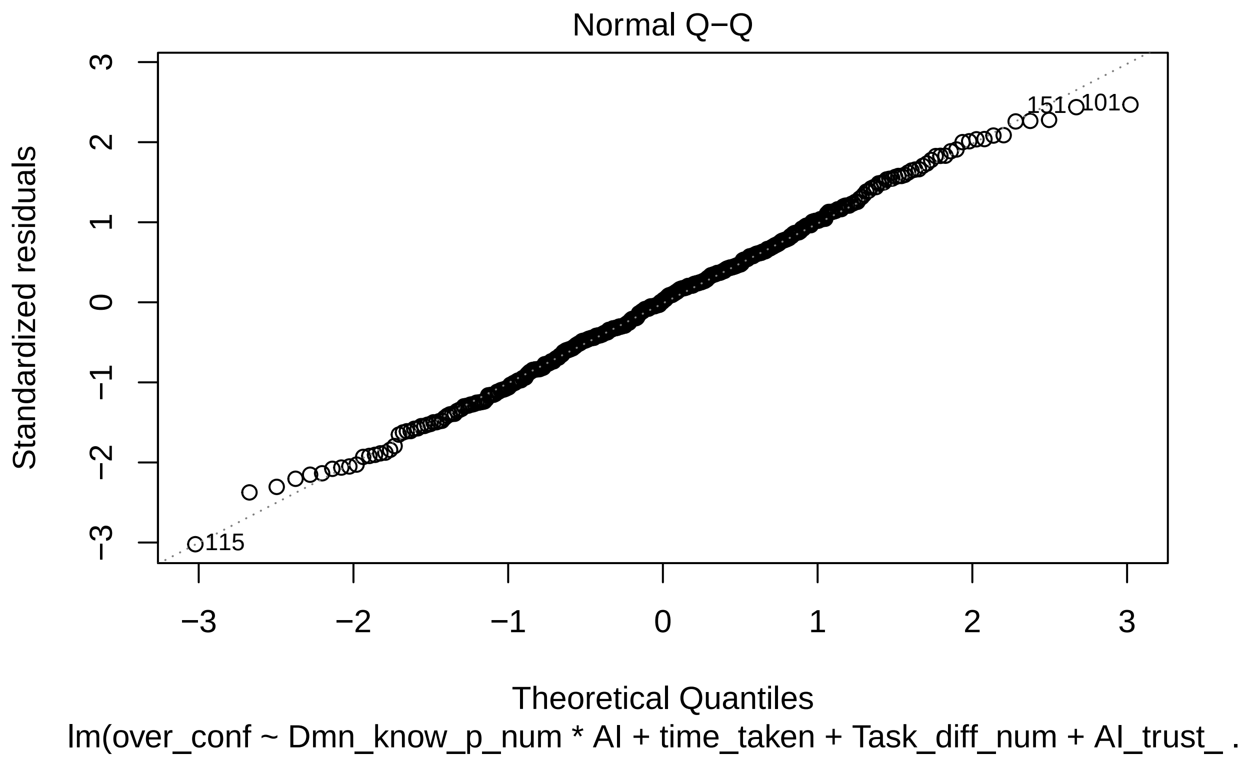
```
##
## Call:
## lm(formula = over_conf ~ Dmn_know_p_num * AI + time_taken + Task_diff_num +
##     AI_trust_num + atn_ch + log(age) + male_num + college, data = plants_person)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.54270 -0.12277  0.00323  0.11663  0.44607
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.2189342  0.1129760   1.938  0.05336 .
## Dmn_know_p_num     0.2011180  0.0638524   3.150  0.00176 **
## AI                 0.0177578  0.0248569   0.714  0.47541
## time_taken         0.0002063  0.0012424   0.166  0.86818
## Task_diff_num     -0.0189442  0.0092827  -2.041  0.04195 *
## AI_trust_num       0.0106612  0.0092607   1.151  0.25035
## atn_ch            -0.0281405  0.0213745  -1.317  0.18877
## log(age)          -0.0721084  0.0312885  -2.305  0.02172 *
```

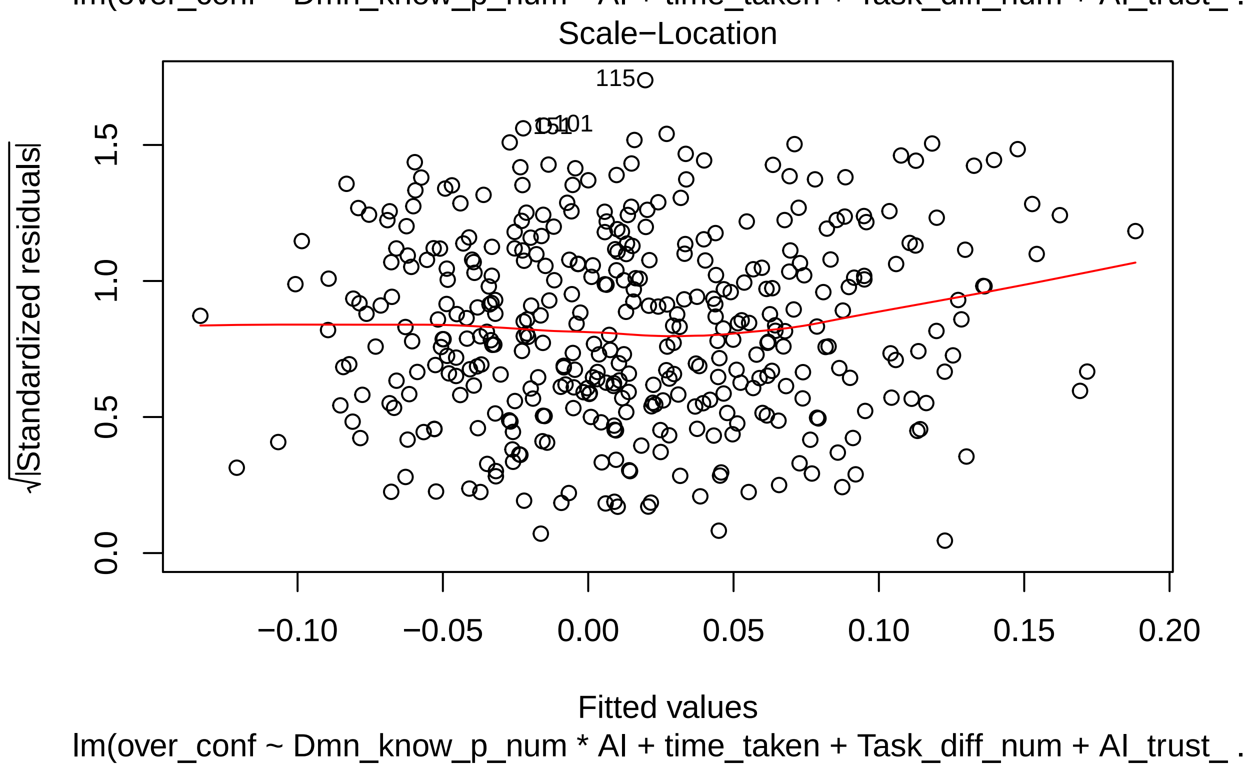
```
## male_num           0.0445323  0.0185664   2.399  0.01693 *
## college            0.0366436  0.0202873   1.806  0.07166 .
## Dmn_know_p_num:AI -0.0559754  0.0865056  -0.647  0.51797
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1822 on 387 degrees of freedom
##   (4 observations deleted due to missingness)
## Multiple R-squared:  0.09544,    Adjusted R-squared:  0.07207
## F-statistic: 4.083 on 10 and 387 DF,  p-value: 2.326e-05
```

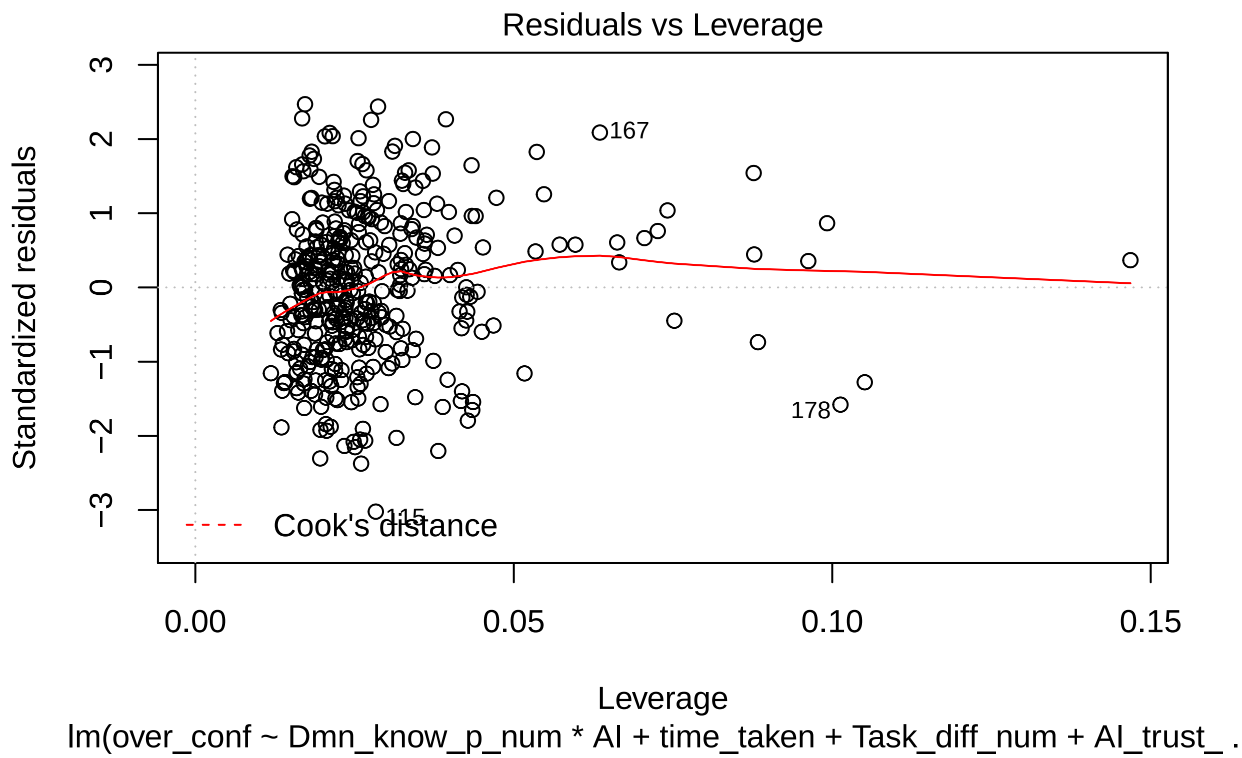
Both jackknife and Q-Q plot are acceptable. good fit.

```r
plot(lm.5.p.overconf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(over_conf ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust_ ..
151
101
115


Scale−Location
√|Standardized residuals|
Fitted values
lm(over_conf ~ Dmn_know_p_num * AI + time_taken + Task_diff_num + AI_trust_ ..
115
151
101



Effect of Uncertainty Information on over confidence with animal domain knowledge

Domain knowledge, perceived AI usefulness rating, task difficulty rating, log(age), education level, and the interaction are all significantly affecting the participants' over confidence. However, task difficulty rating and the interaction term is negatively affecting their over confidence.

F(11, 187) = 6.03, p < 0.001, R^2 = 0.22

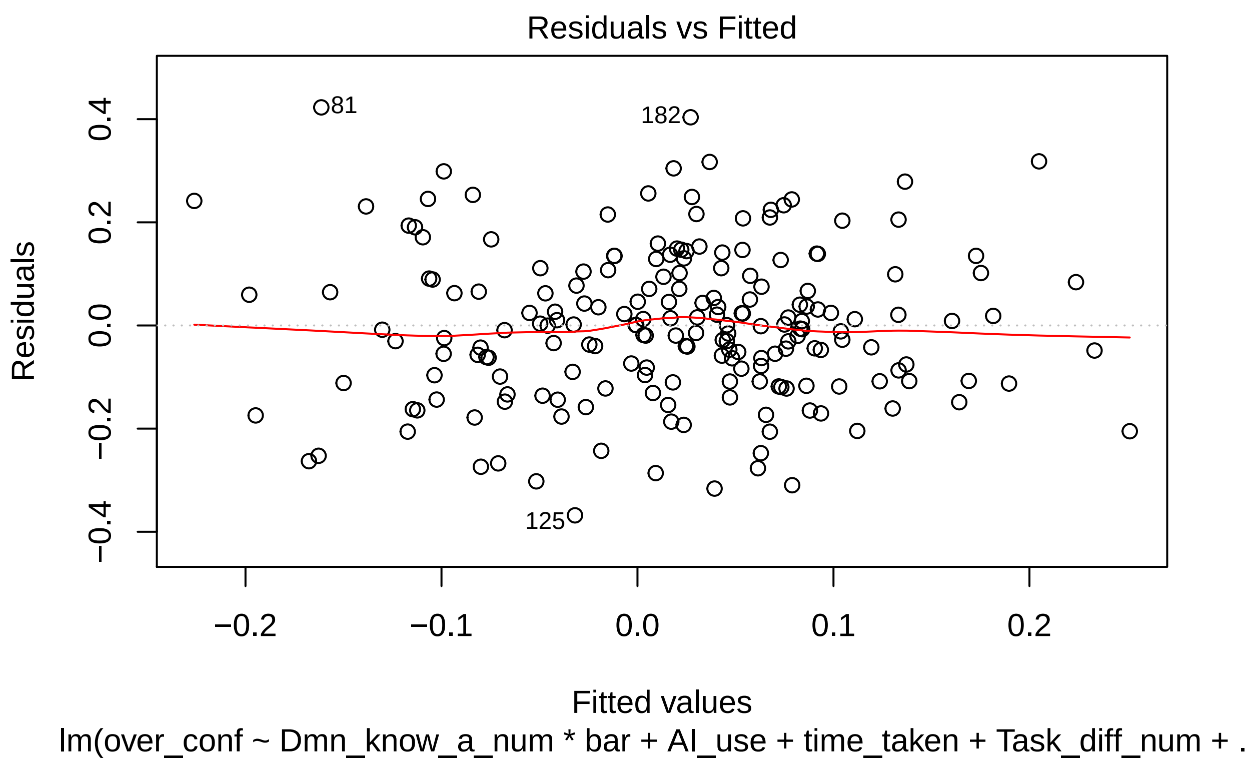
```r
lm.6.a.overconf <- lm(over_conf ~ Dmn_know_a_num*bar + AI_use +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = animals_person_AI)

summary(lm.6.a.overconf)
```

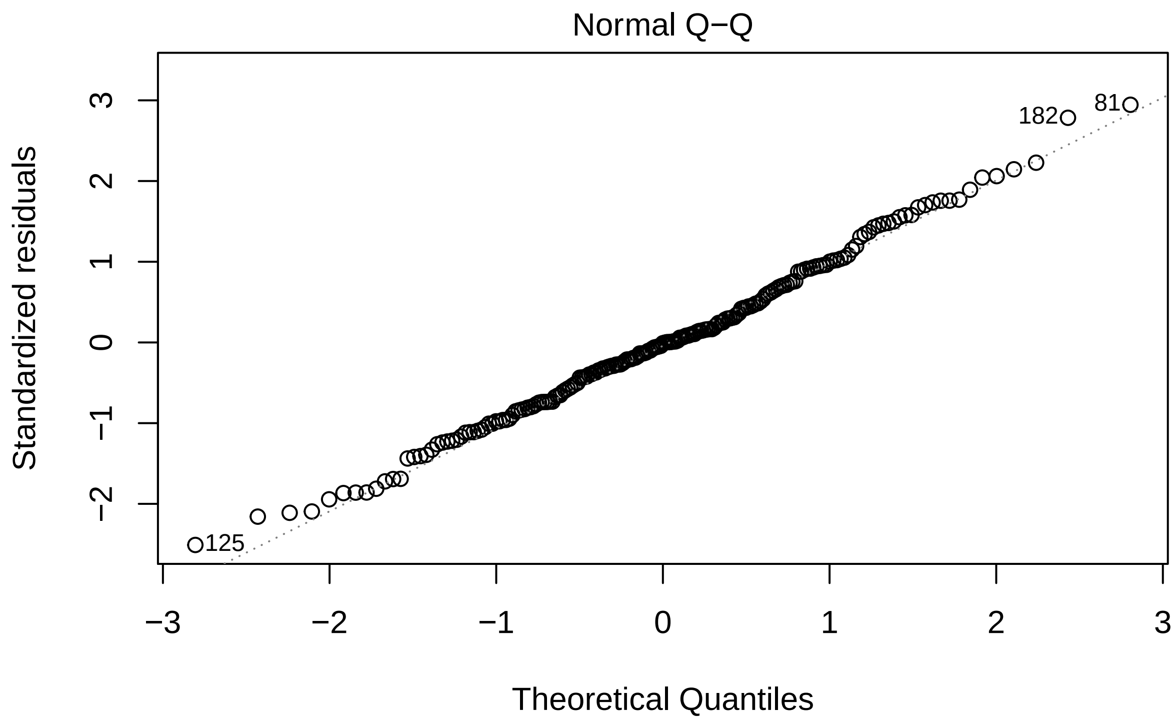
```
##
## Call:
## lm(formula = over_conf ~ Dmn_know_a_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = animals_person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36812 -0.10770 -0.00140  0.09537  0.42286
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.0418908  0.1418810   0.295 0.768128
## Dmn_know_a_num        0.2218033  0.0647082   3.428 0.000748 ***
## bar                   0.0728020  0.0503233   1.447 0.149659
## AI_use                0.3338547  0.0674864   4.947 1.67e-06 ***
## time_taken            0.0007850  0.0021741   0.361 0.718455
```

```
## Task_diff_num      -0.0271913  0.0107732  -2.524 0.012435 *
## AI_trust_num       -0.0006921  0.0113772  -0.061 0.951557
## atn_ch             -0.0145036  0.0259809  -0.558 0.577349
## log(age)           -0.0777586  0.0367992  -2.113 0.035923 *
## male_num            0.0359821  0.0218247   1.649 0.100891
## college             0.0616772  0.0239777   2.572 0.010880 *
## Dmn_know_a_num:bar -0.2163741  0.0937262  -2.309 0.022061 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1505 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.2617, Adjusted R-squared:  0.2183
## F-statistic: 6.027 on 11 and 187 DF,  p-value: 2.116e-08
```

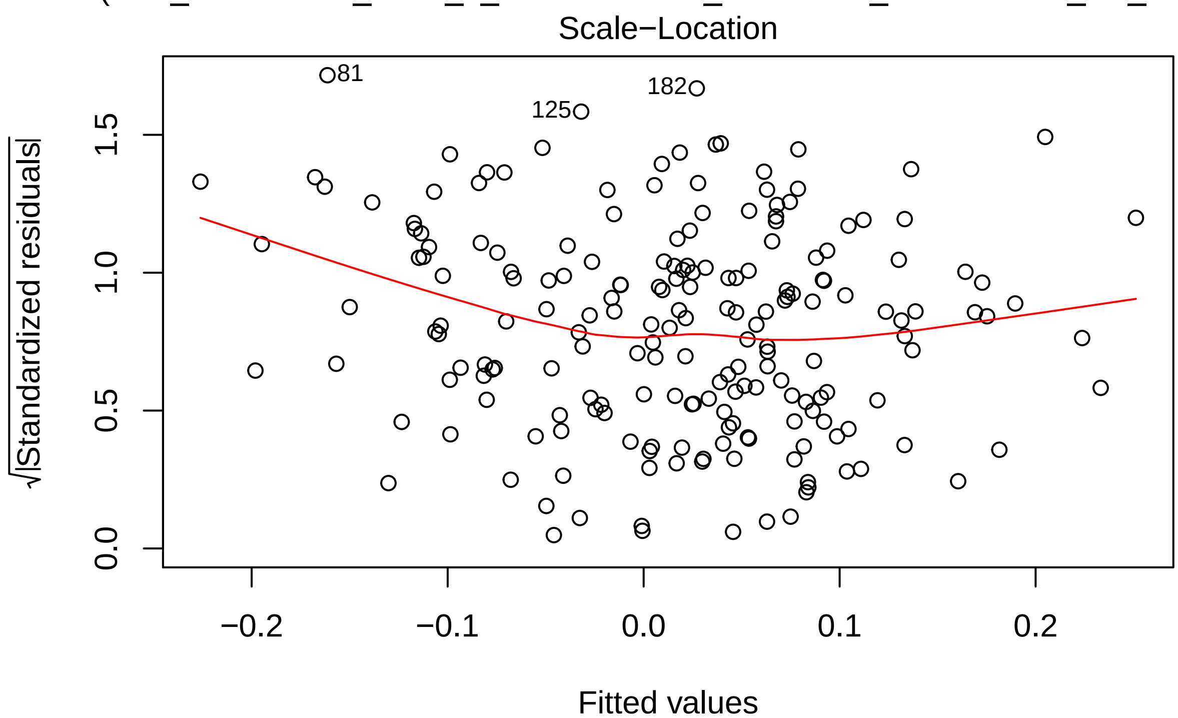
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.6.a.overconf)
```

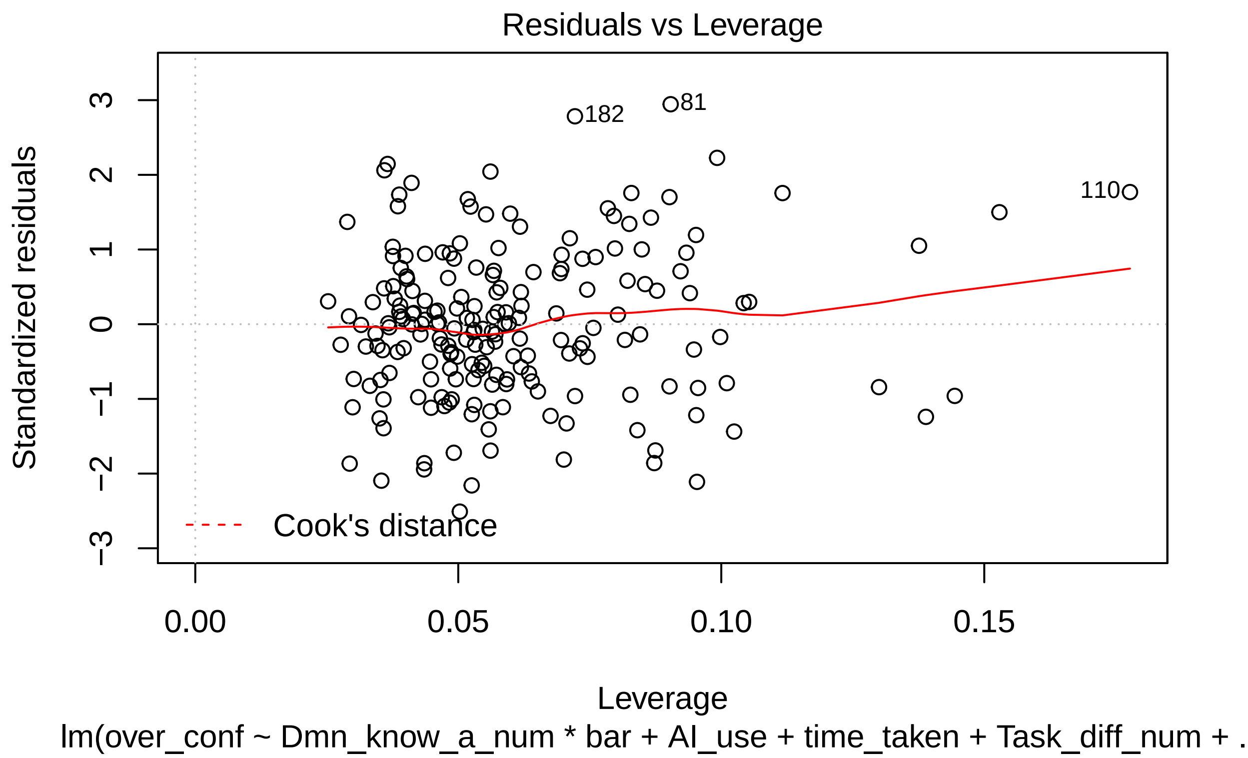




lm(over_conf ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num + .



lm(over_conf ~ Dmn_know_a_num * bar + AI_use + time_taken + Task_diff_num + .



Effect of Uncertainty Information on over confidence with plant domain knowledge

Perceived AI usefulness rating, task difficulty, log(age), gender, and education all significantly affect the participants' over confidence.

$F(11, 187) = 6.84$, $p < 0.001$, $R\hat{}2 = 0.24$

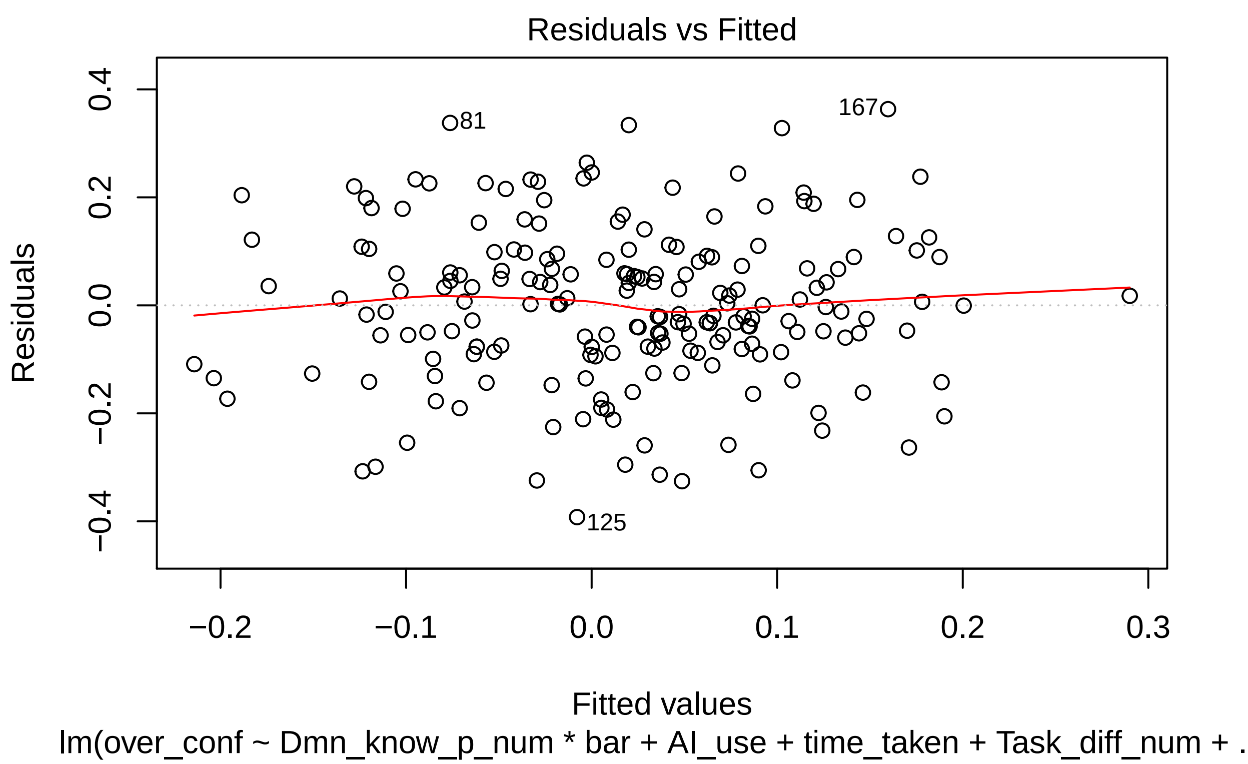
```
lm.6.p.overconf <- lm(over_conf ~ Dmn_know_p_num*bar + AI_use +
                  time_taken + Task_diff_num + AI_trust_num + atn_ch + log(age) +
                  male_num + college, data = plants_person_AI)

summary(lm.6.p.overconf)
```

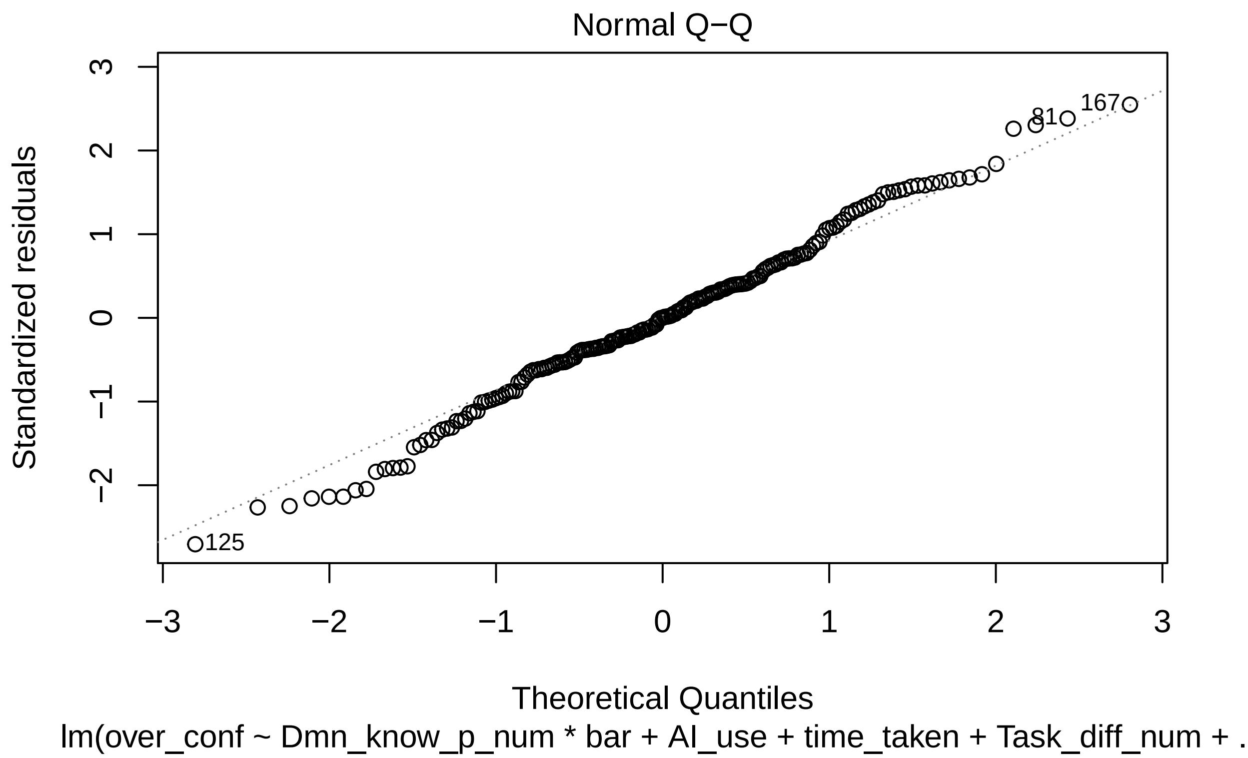
```
##
## Call:
## lm(formula = over_conf ~ Dmn_know_p_num * bar + AI_use + time_taken +
##     Task_diff_num + AI_trust_num + atn_ch + log(age) + male_num +
##     college, data = plants_person_AI)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.39215 -0.08249  0.00011  0.08944  0.36334
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)         0.097179   0.135907   0.715   0.4755
## Dmn_know_p_num      0.121792   0.067301   1.810   0.0720 .
## bar                -0.009176   0.030239  -0.303   0.7619
## AI_use              0.401133   0.070141   5.719 4.19e-08 ***
## time_taken          0.001219   0.001557   0.783   0.4346
## Task_diff_num      -0.023751   0.010753  -2.209   0.0284 *
```

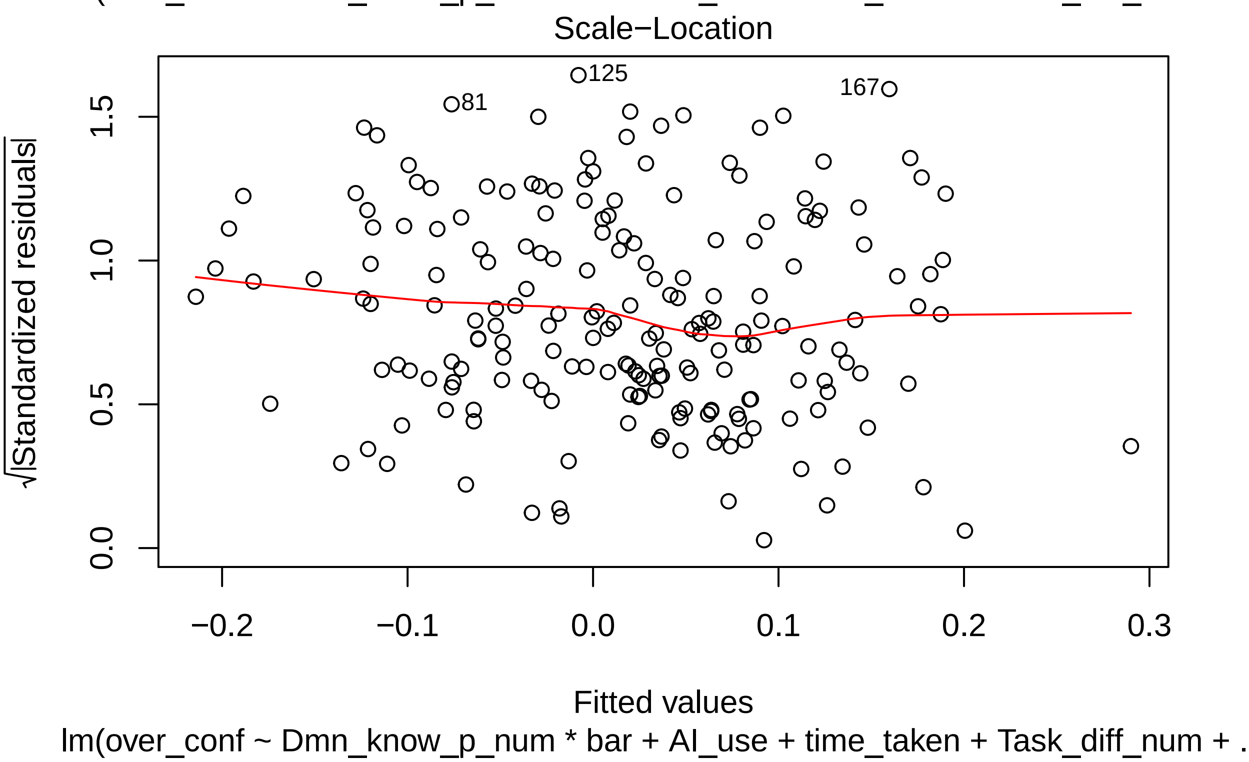
```
## AI_trust_num       -0.009603   0.011504  -0.835   0.4049
## atn_ch             -0.017005   0.025515  -0.666   0.5059
## log(age)           -0.072510   0.036501  -1.987   0.0484 *
## male_num            0.047984   0.021690   2.212   0.0282 *
## college             0.055353   0.023563   2.349   0.0199 *
## Dmn_know_p_num:bar -0.160851   0.110737  -1.453   0.1480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.148 on 187 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.2868, Adjusted R-squared:  0.2448
## F-statistic: 6.836 on 11 and 187 DF,  p-value: 1.242e-09
```

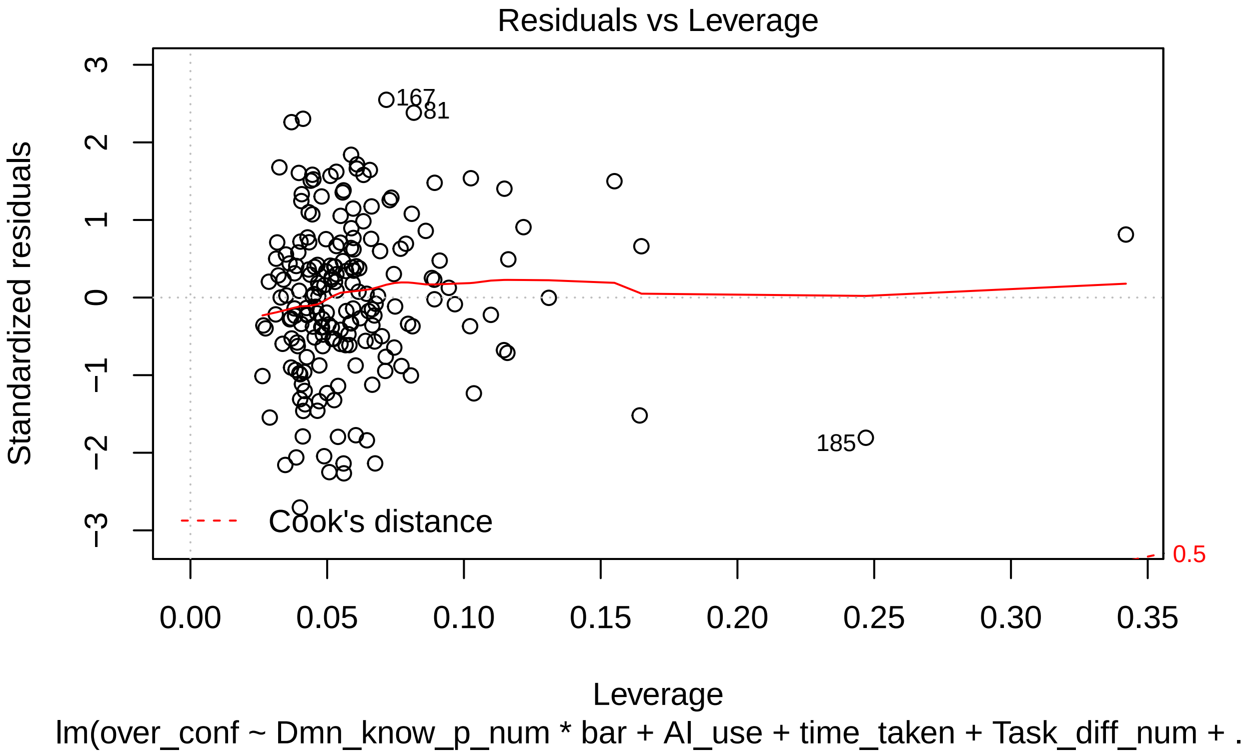
Both jackknife and Q-Q plot are acceptable. good fit.

```
plot(lm.6.p.overconf)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(over_conf ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num + .
125
81
167


Scale−Location
√|Standardized residuals|
Fitted values
lm(over_conf ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num + .
125
81
167


Residuals vs Leverage
Standardized residuals
Leverage
lm(over_conf ~ Dmn_know_p_num * bar + AI_use + time_taken + Task_diff_num + .
Cook's distance
0.5
167
81
185
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
−3
−2
−1
0
1
2
3