Data_Analysis.R

maximelebourgeois

2025-06-02

```
library(ggpmisc)
```

```
## Loading required package: ggpp
```

```
## Loading required package: ggplot2
```

```
## Registered S3 methods overwritten by 'ggpp':
##   method                  from
##   heightDetails.titleGrob ggplot2
##   widthDetails.titleGrob  ggplot2
```

```
##
## Attaching package: 'ggpp'
```

```
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
library(readxl)
library(tibble)
library(tinytex)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x ggpp::annotate() masks ggplot2::annotate()
## x dplyr::filter()  masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(afex)
```

```
## Loading required package: lme4
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
## ************
## Welcome to afex. For support visit: http://afex.singmann.science/
## - Functions for ANOVAs: aov_car(), aov_ez(), and aov_4()
## - Methods for calculating p-values with mixed(): 'S', 'KR', 'LRT', and 'PB'
## - 'afex_aov' and 'mixed' objects can be passed to emmeans() for follow-up tests
## - Get and set global package options with: afex_options()
## - Set sum-to-zero contrasts globally: set_sum_contrasts()
## - For example analyses see: browseVignettes("afex")
## ************
##
## Attaching package: 'afex'
##
## The following object is masked from 'package:lme4':
##
##     lmer
```

```
library(emmeans)
```

```
## Welcome to emmeans.
## Caution: You lose important information if you filter this package's results.
## See '? untidy'
```

```
library(ggbeeswarm)
library(cowplot)
```

```
##
## Attaching package: 'cowplot'
##
## The following object is masked from 'package:lubridate':
##
##     stamp
```

```
library(ggplot2)
library(psych)
```

```
##
## Attaching package: 'psych'
##
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```r
theme_set(theme_bw(base_size = 15) + theme(legend.position = "bottom"))
setwd("/Users/maximelebourgeois/Desktop/experience_of_trust_final/Results")
data <- read_excel("Cleaned_Data_R.xlsx")
glimpse(data)
```

```
## Rows: 145
## Columns: 26
## $ Prolific_ID                <chr> "572f526c3c27e7000e0b8aaa", "5b6db242d2eae0~
## $ Condition                  <chr> "DG", "DG", "TG", "DG", "TG", "DG", "TG", "~
## $ List                       <chr> "Low start", "Low end", "Low middle", "Low ~
## $ Expected_Amount            <dbl> 4, 12, 6, 2, 12, 4, 8, 10, 6, 5, 12, 3, 2, ~
## $ Average_Amounts_Received   <dbl> 24, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24,~
## $ Average_Amounts_Before     <dbl> 36, 36, 36, 36, 36, 36, 36, 36, 36, 36, 36,~
## $ Average_Amounts_Sent_Back  <dbl> NA, NA, 10.6, NA, 8.6, NA, 17.7, 16.4, 14.6~
## $ Average_Amounts_at_the_end <dbl> 36.0, 36.0, 25.4, 36.0, 27.4, 36.0, 18.3, 1~
## $ Average_Received           <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8~
## $ Bonus                      <chr> "0.72", "0.72", "0.51", "0.72", "0.55000000~
## $ PANAS_Pre                  <dbl> 25, 16, 15, 12, 20, 13, 22, 23, 10, 22, 16,~
## $ PANAS_Post                 <dbl> 29, 19, 12, 12, 20, 16, 20, 22, 13, 31, 17,~
## $ STAIS_Pre                  <dbl> 12, 17, 14, 11, 20, 20, 12, 15, 17, 21, 4, ~
## $ STAIS_Post                 <dbl> 11, 13, 13, 11, 15, 18, 11, 13, 20, 12, 5, ~
## $ SSVS                       <dbl> 2.9, 1.8, 2.3, 2.8, 3.0, 2.7, 3.0, 3.4, 2.1~
## $ DV_PANAS                   <dbl> 4, 3, -3, 0, 0, 3, -2, -1, 3, 9, 1, -1, 0, ~
## $ DV_STAIS                   <dbl> -1, -4, -1, 0, -5, -2, -1, -2, 3, -9, 1, 1,~
## $ IV_Dif._Expected           <dbl> 4, -4, 2, 6, -4, 4, 0, -2, 2, 3, -4, 5, 6, ~
## $ IV_Average                 <dbl> 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8~
## $ Fluent_languages           <chr> "English", "English", "English", "Bulgarian~
## $ Age                        <dbl> 35, 33, 26, 32, 27, 32, 31, 20, 33, 31, 33,~
## $ Sex                        <chr> "Male", "Female", "Female", "Female", "Male~
## $ Ethnicity_simplified       <chr> "White", "White", "Black", "White", "White"~
## $ Country_of_birth           <chr> "United Kingdom", "United Kingdom", "United~
## $ Country_of_residence       <chr> "United Kingdom", "United Kingdom", "United~
## $ Nationality                <chr> "United Kingdom", "United Kingdom", "United~
```

```r
## I - Data Cleaning

# 1) Factorisation of the categorical variable
data$Prolific_ID <- factor(data$Prolific_ID)
data$Condition <- factor(data$Condition, levels = c("DG", "TG"), labels = c("Dictator Game", "Trust Game
data$Sex <- factor(data$Sex)
data$Ethnicity_simplified <- factor(data$Ethnicity_simplified)
data$Nationality <- factor(data$Nationality)

# 2) Ensure DV variables are numeric
data$PANAS_Pre <- as.numeric(data$PANAS_Pre)
data$PANAS_Post <- as.numeric(data$PANAS_Post)
data$STAIS_Pre <- as.numeric(data$STAIS_Pre)
data$STAIS_Post <- as.numeric(data$STAIS_Post)
data$DV_PANAS <- as.numeric(data$DV_PANAS)
data$DV_STAIS <- as.numeric(data$DV_STAIS)
data$IV_Dif._Expected <- as.numeric(data$IV_Dif._Expected)
data$Age <- as.numeric(data$Age)
```

```r
## II - Data Analysis

# 1) Demographics Data Analysis

data_condition <- data %>%
  group_by(Condition) %>%
  summarise(n = n()) %>%
  mutate(proportion = n / sum(n))
data_condition
```

```
## # A tibble: 2 x 3
##   Condition         n proportion
##   <fct>         <int>      <dbl>
## 1 Dictator Game    82      0.566
## 2 Trust Game       63      0.434
```

```r
data_sex <- data %>%
  group_by(Sex) %>%
  summarise(n = n()) %>%
  mutate(proportion = n / sum(n))
data_sex
```

```
## # A tibble: 2 x 3
##   Sex        n proportion
##   <fct>  <int>      <dbl>
## 1 Female    73      0.503
## 2 Male      72      0.497
```

```r
data_ethnicity <- data %>%
  group_by(Ethnicity_simplified) %>%
  summarise(n = n()) %>%
  mutate(proportion = n / sum(n))
data_ethnicity
```

```
## # A tibble: 5 x 3
##   Ethnicity_simplified     n proportion
##   <fct>                <int>      <dbl>
## 1 Asian                    4    0.0276
## 2 Black                   16    0.110
## 3 Mixed                    3    0.0207
## 4 Other                    1    0.00690
## 5 White                  121    0.834
```

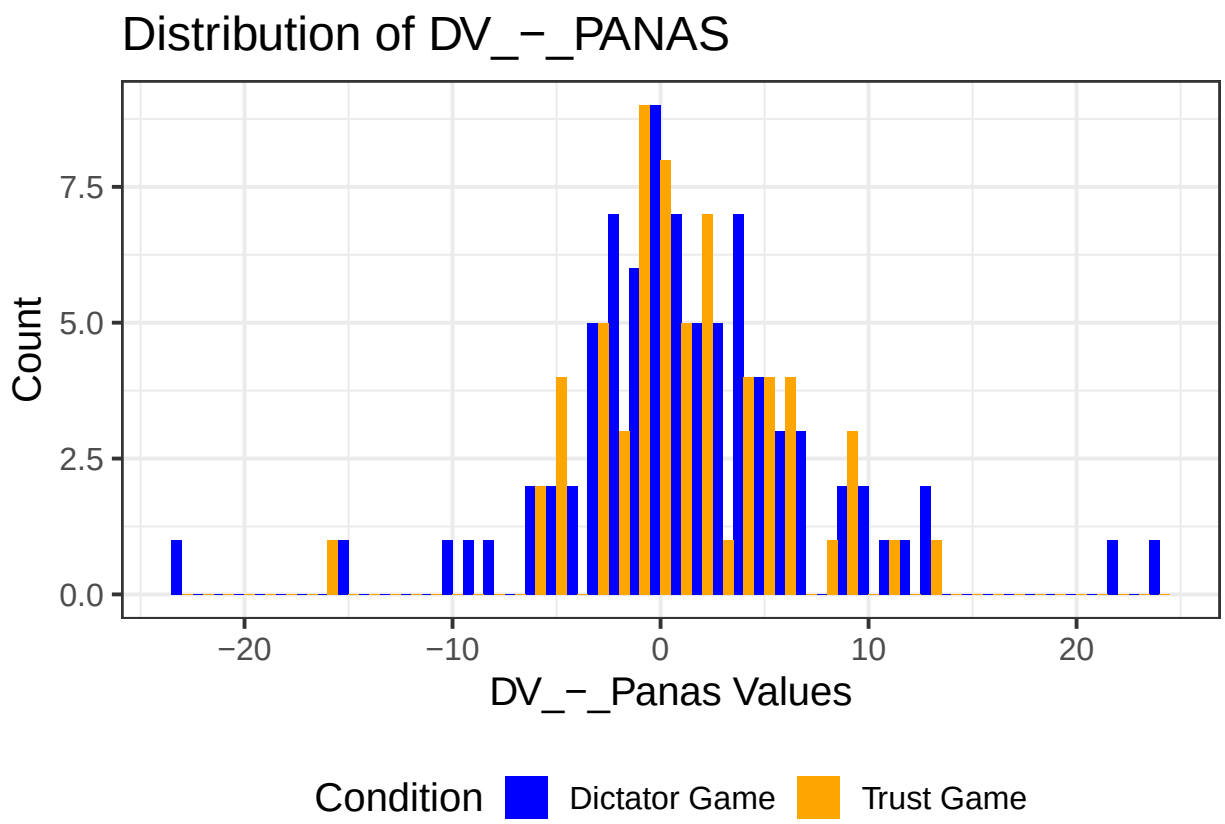
```r
data_age <- data %>%
  summarise(sd = sd(Age),
            mean = mean(Age),
            min = min(Age),
            max = max(Age),
            )
data_age
```

```
## # A tibble: 1 x 4
##      sd  mean   min   max
##   <dbl> <dbl> <dbl> <dbl>
## 1  4.46  28.7    18    35
```

```
data_nationality <- data %>%
  group_by(Nationality) %>%
  summarise(n = n()) %>%
  mutate(proportion = n / sum(n))
data_nationality
```

```
## # A tibble: 4 x 3
##   Nationality        n proportion
##   <fct>          <int>      <dbl>
## 1 Albania            1    0.00690
## 2 Bulgaria           3    0.0207
## 3 France             1    0.00690
## 4 United Kingdom   140    0.966
```

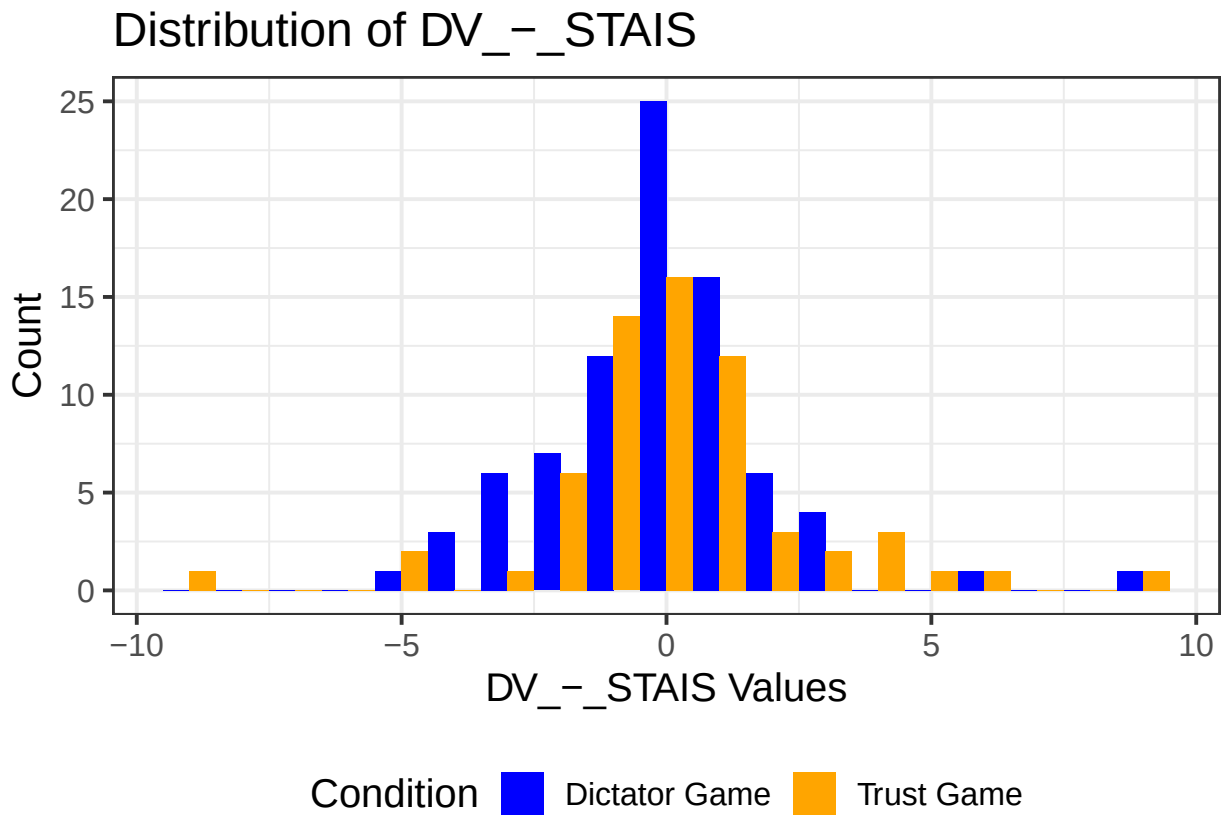
```
# 2) Plot for DV_-_PANAS
h_PANAS <- ggplot(data, aes(x = DV_PANAS, fill = Condition)) +
  geom_histogram(position = "dodge", binwidth = 1) +
  labs(title = "Distribution of DV_-_PANAS", x = "DV_-_Panas Values", y = "Count") +
  scale_fill_manual(values = c("Dictator Game" = "blue", "Trust Game" = "orange"))
h_PANAS
```



```r
summary_stats_PANAS <- data %>%
  group_by(Condition) %>%
  summarise(
    Mean_PANAS_Pre = mean(PANAS_Pre, na.rm = TRUE),
    Mean_PANAS_Post = mean(PANAS_Post, na.rm = TRUE),
    Mean_Dif_PANAS = Mean_PANAS_Post - Mean_PANAS_Pre,
    Median_PANAS_Pre = median(PANAS_Pre, na.rm = TRUE),
    Median_PANAS_Post = mean(PANAS_Post, na.rm = TRUE),
    Median_Dif_PANAS = Median_PANAS_Post - Median_PANAS_Pre,
    SD_Dif_PANAS = sd(DV_PANAS, na.rm = TRUE),
    .groups = "drop"
  )
glimpse(summary_stats_PANAS)
```

```
## Rows: 2
## Columns: 8
## $ Condition         <fct> Dictator Game, Trust Game
## $ Mean_PANAS_Pre    <dbl> 15.78049, 16.07937
## $ Mean_PANAS_Post   <dbl> 17.26829, 17.07937
## $ Mean_Dif_PANAS    <dbl> 1.487805, 1.000000
## $ Median_PANAS_Pre  <dbl> 15.5, 18.0
## $ Median_PANAS_Post <dbl> 17.26829, 17.07937
## $ Median_Dif_PANAS  <dbl> 1.7682927, -0.9206349
## $ SD_Dif_PANAS      <dbl> 6.680068, 4.714425
```

```r
# 3) Plot for DV_-_STAIS
h_STAIS <- ggplot(data, aes(x = DV_STAIS, fill = Condition)) +
  geom_histogram(position = "dodge", binwidth = 1) +
  labs(title = "Distribution of DV_-_STAIS", x = "DV_-_STAIS Values", y = "Count") +
  scale_fill_manual(values = c("Dictator Game" = "blue", "Trust Game" = "orange"))
h_STAIS
```




```
summary_stats_STAIS <- data %>%
  group_by(Condition) %>%
  summarise(
    Mean_STAIS_Pre = mean(STAIS_Pre, na.rm = TRUE),
    Mean_STAIS_Post = mean(STAIS_Post, na.rm = TRUE),
    Mean_Dif_STAIS = Mean_STAIS_Post - Mean_STAIS_Pre,
    Median_STAIS_Pre = median(STAIS_Pre, na.rm = TRUE),
    Median_STAIS_Post = mean(STAIS_Post, na.rm = TRUE),
    Median_Dif_STAIS = Median_STAIS_Post - Median_STAIS_Pre,
    SD_Dif_STAIS = sd(DV_STAIS, na.rm = TRUE),
    .groups = "drop"
  )
glimpse(summary_stats_STAIS)
```

```
## Rows: 2
## Columns: 8
## $ Condition         <fct> Dictator Game, Trust Game
## $ Mean_STAIS_Pre    <dbl> 8.414634, 8.666667
## $ Mean_STAIS_Post   <dbl> 8.341463, 8.793651
## $ Mean_Dif_STAIS    <dbl> -0.07317073, 0.12698413
## $ Median_STAIS_Pre  <dbl> 8, 8
## $ Median_STAIS_Post <dbl> 8.341463, 8.793651
## $ Median_Dif_STAIS  <dbl> 0.3414634, 0.7936508
## $ SD_Dif_STAIS      <dbl> 2.106899, 2.568369
```

```r
### III - Statistical Analysis

#  1°) ANOVA Condition Analysis

res_ANOVA_PANAS <- aov_car(DV_PANAS ~ Condition + Error (Prolific_ID), data)
```

```
## Warning: More than one observation per design cell, aggregating data using ‘fun_aggregate = mean‘.
## To turn off this warning, pass ‘fun_aggregate = mean‘ explicitly.
```

```
## Contrasts set to contr.sum for the following variables: Condition
```

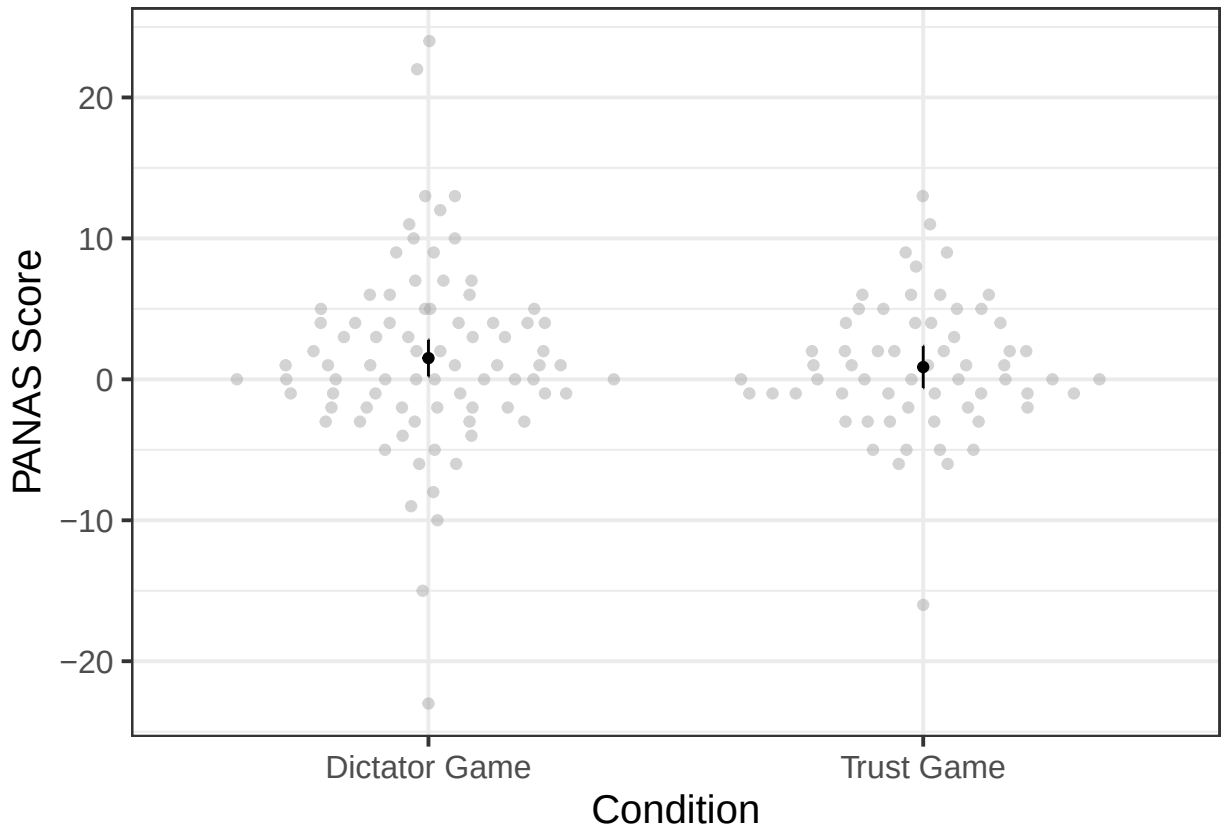
```r
res_ANOVA_PANAS
```

```
## Anova Table (Type 3 tests)
##
## Response: DV_PANAS
##      Effect     df   MSE    F  ges p.value
## 1 Condition 1, 141 34.87 0.41 .003    .525
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’+’ 0.1 ’ ’ 1
```

```r
Diff_ANOVA_PANAS = emmeans(res_ANOVA_PANAS,"Condition")
pairs(Diff_ANOVA_PANAS,adjust="holm")
```

```
##  contrast                   estimate    SE  df t.ratio p.value
##  Dictator Game - Trust Game    0.635 0.997 141   0.637  0.5249
```

```r
plot_ANOVA_PANAS <- afex_plot(res_ANOVA_PANAS,"Condition", data_geom = ggbeeswarm::geom_quasirandom)+ la
plot_ANOVA_PANAS
```




```
res_ANOVA_STAIS <- aov_car(DV_STAIS ~ Condition + Error (Prolific_ID), data)
```

```
## Warning: More than one observation per design cell, aggregating data using ‘fun_aggregate = mean‘.
## To turn off this warning, pass ‘fun_aggregate = mean‘ explicitly.
```

```
## Contrasts set to contr.sum for the following variables: Condition
```

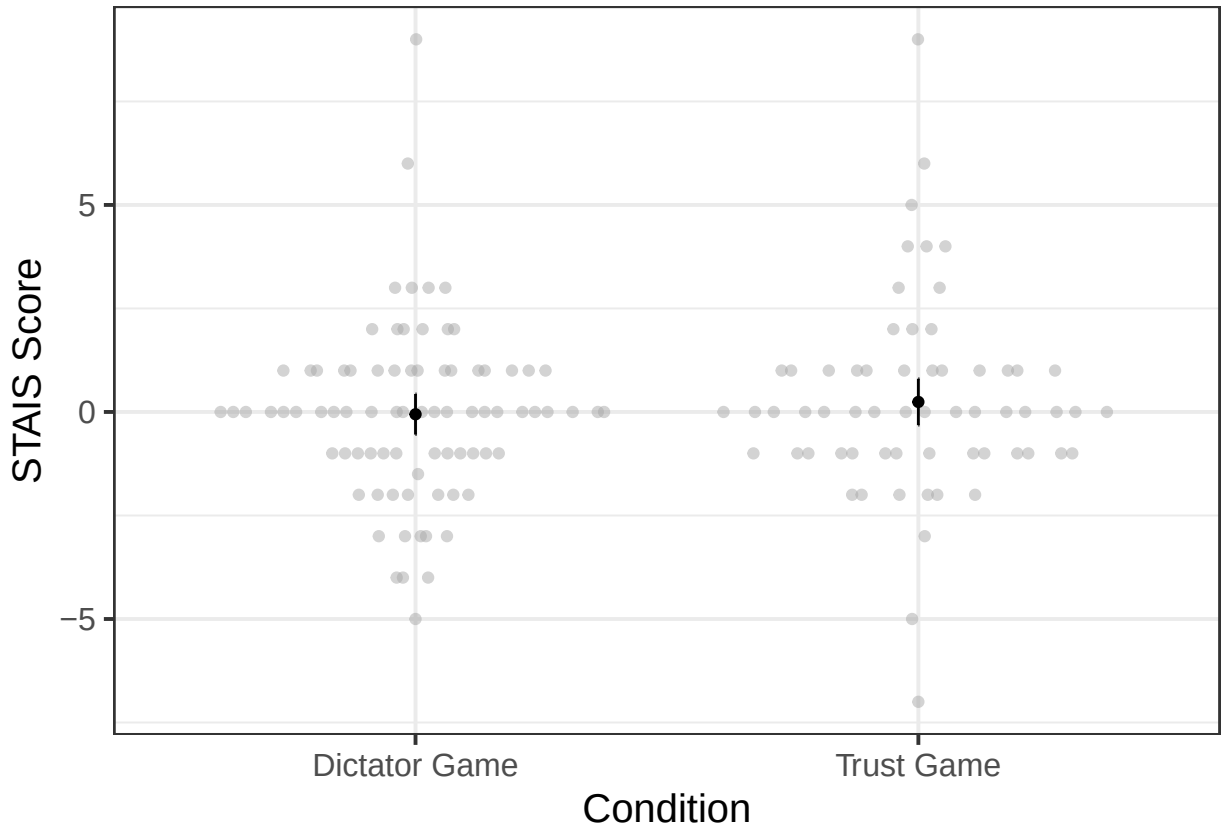
```
res_ANOVA_STAIS
```

```
## Anova Table (Type 3 tests)
##
## Response: DV_STAIS
##      Effect     df  MSE    F  ges p.value
## 1 Condition 1, 141 4.98 0.62 .004    .431
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’+’ 0.1 ’ ’ 1
```

```
Diff_ANOVA_STAIS = emmeans(res_ANOVA_STAIS,"Condition")
pairs(Diff_ANOVA_STAIS, adjust="holm")
```

```
##  contrast                  estimate    SE  df t.ratio p.value
##  Dictator Game - Trust Game   -0.297 0.377 141  -0.790  0.4309
```

```r
plot_ANOVA_STAIS <- afex_plot(res_ANOVA_STAIS,"Condition", data_geom = ggbeeswarm::geom_quasirandom)+ la
plot_ANOVA_STAIS
```



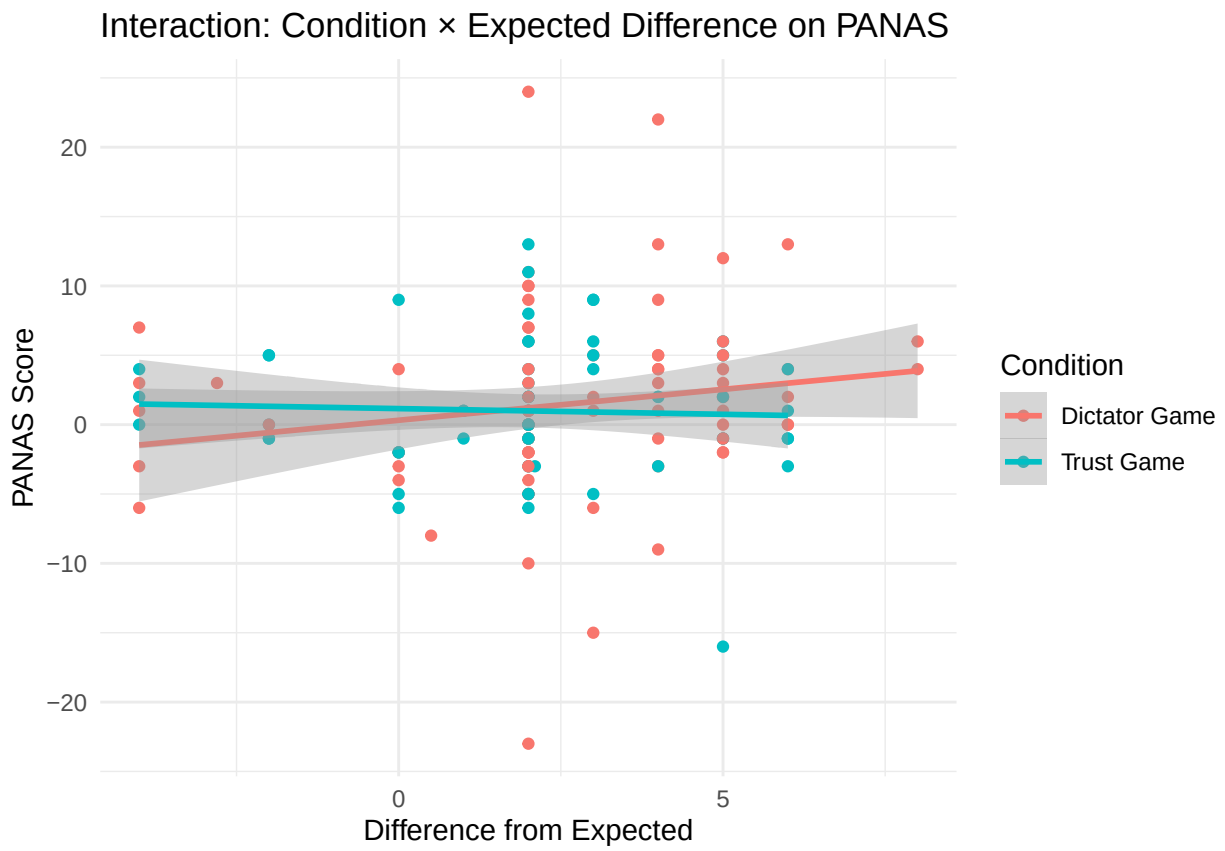

```r
# 2°) LR Diff Expectation

res_LR_PANAS <- lm(DV_PANAS ~ Condition * IV_Dif._Expected, data = data)
summary(res_LR_PANAS)
```

```
##
## Call:
## lm(formula = DV_PANAS ~ Condition * IV_Dif._Expected, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -24.2069  -3.1568  -0.6524   2.9020  22.7931
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                            0.3157     0.9369   0.337   0.7366
## ConditionTrust Game                    0.8411     1.3342   0.630   0.5295
## IV_Dif._Expected                       0.4456     0.2565   1.737   0.0845 .
## ConditionTrust Game:IV_Dif._Expected  -0.5278     0.4034  -1.308   0.1928
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 5.886 on 141 degrees of freedom
## Multiple R-squared:  0.02309,    Adjusted R-squared:  0.002303
## F-statistic: 1.111 on 3 and 141 DF,  p-value: 0.3469
```

```
plot_LR_PANAS <- ggplot(data, aes(x = IV_Dif._Expected, y = DV_PANAS, color = Condition)) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  theme_minimal() +
  labs(title = "Interaction: Condition × Expected Difference on PANAS",
       x = "Difference from Expected", y = "PANAS Score")
plot_LR_PANAS
```

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




```
res_LR_STAIS <- lm(DV_STAIS ~ Condition * IV_Dif._Expected, data = data)
summary(res_LR_STAIS)
```

```
##
## Call:
## lm(formula = DV_STAIS ~ Condition * IV_Dif._Expected, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3330 -1.1446  0.0043  1.0438  9.2322
```

```
##
## Coefficients:
##                                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)                        -0.20544    0.36819  -0.558    0.578
## ConditionTrust Game                -0.02676    0.52431  -0.051    0.959
## IV_Dif._Expected                    0.05028    0.10080   0.499    0.619
## ConditionTrust Game:IV_Dif._Expected  0.13813    0.15853   0.871    0.385
##
## Residual standard error: 2.313 on 141 degrees of freedom
## Multiple R-squared:  0.02006,    Adjusted R-squared:  -0.0007863
## F-statistic: 0.9623 on 3 and 141 DF,  p-value: 0.4125
```

```r
plot_LR_STAIS <- ggplot(data, aes(x = IV_Dif._Expected, y = DV_STAIS, color = Condition)) +
  geom_point() +
  geom_smooth(method = "lm", se = TRUE) +
  theme_minimal() +
  labs(title = "Interaction: Condition × Expected Difference on PANAS",
       x = "Difference from Expected", y = "STAIS Score")
plot_LR_STAIS
```

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