

New_PP_Narrative_Analysis_H10

Kristen Johnson

2025-04-24

```
# install necessary packages

#install.packages("lmerTest")

# Load required packages
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 ggplot2    3.5.1      v tibble    3.2.1
## v lubridate  1.9.4      v tidyr     1.3.1
## v purrr      1.0.4
## -- Conflicts ----- tidyverse_conflicts() --
## 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(car)      # For regression diagnostics

## Loading required package: carData
##
## Attaching package: 'car'
##
## The following object is masked from 'package:dplyr':
##
##     recode
##
## The following object is masked from 'package:purrr':
##
##     some

library(lmtest)    # For regression diagnostics

## Loading required package: zoo
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(effectsize) # For effect sizes
library(ggplot2)    # For visualization
library(here)
```

```
## here() starts at /Users/kristenjohnson/KristenWorkingDirectory/Play_Narrative
library(dplyr)
library(lme4)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
##
## The following object is masked from 'package:lme4':
##
##     lmer
##
## The following object is masked from 'package:stats':
##
##     step
```

```
library(extrafont)
```

```
## Registering fonts with R
```

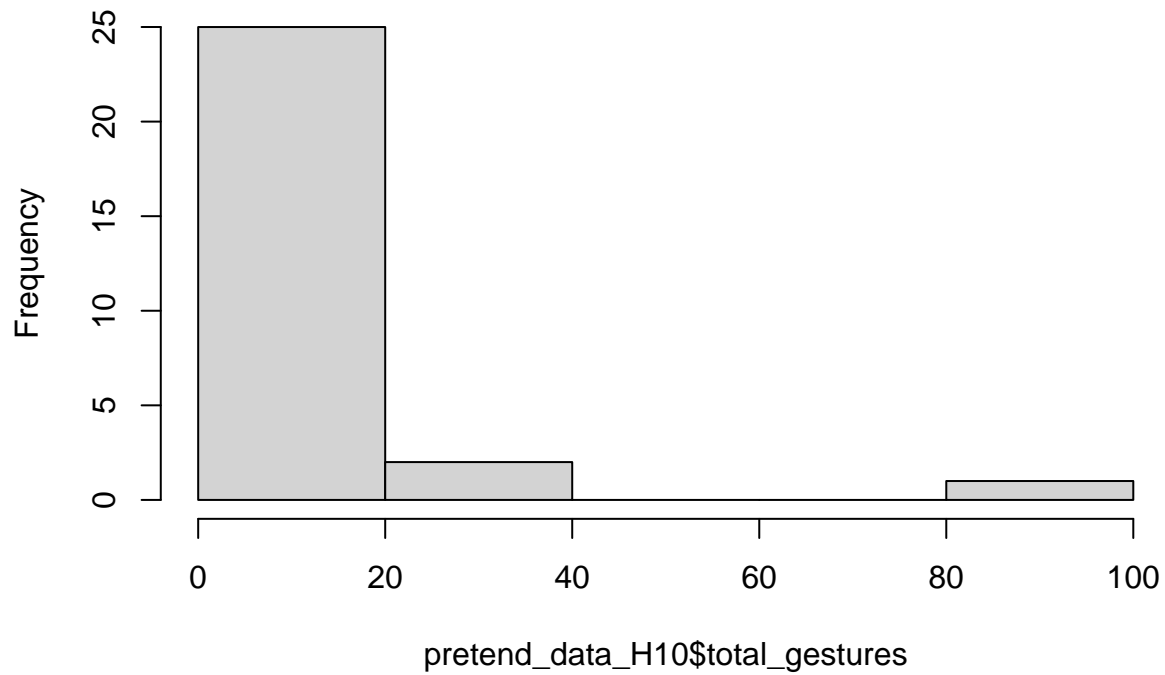
read in, simplify, and check gesture data in pretend play

```
pretend_data_H10 <- read_csv(here("PN_Datasets", "Pretend_Play", "CSVs_of_Combined_Data_PP", "child_summ

## Rows: 28 Columns: 10
## -- Column specification -----
## Delimiter: ","
## chr (1): groupstatus
## dbl (8): participant_id, total_gestures, rep_gestures, total_pretend_episode...
## lgl (1): mlu_missing_flag
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
# Create binary variable for representational gesture presence
pretend_data_H10 <- pretend_data_H10 %>%
  mutate(rep_gesture_present = ifelse(rep_gestures > 0, 1, 0))

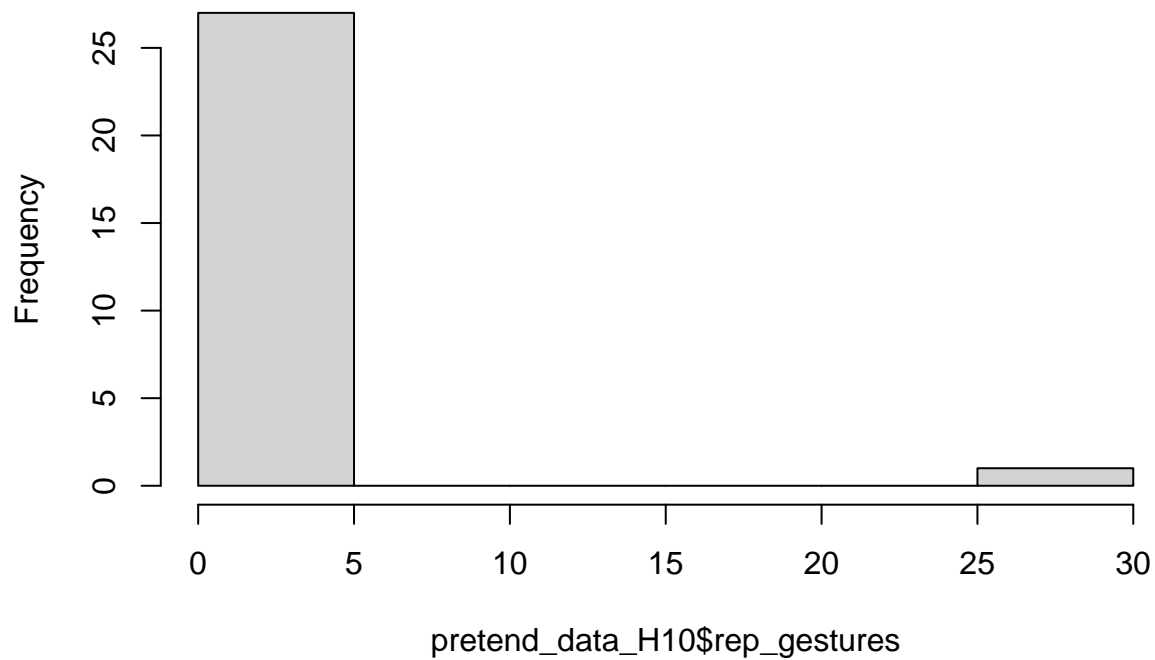
# Check distributions
hist(pretend_data_H10$total_gestures)
```

Histogram of pretend_data_H10\$total_gestures



```
hist(pretend_data_H10$rep_gestures)
```

Histogram of pretend_data_H10\$rep_gestures



```
table(pretend_data_H10$rep_gesture_present, pretend_data_H10$groupstatus)
```

```
##
```

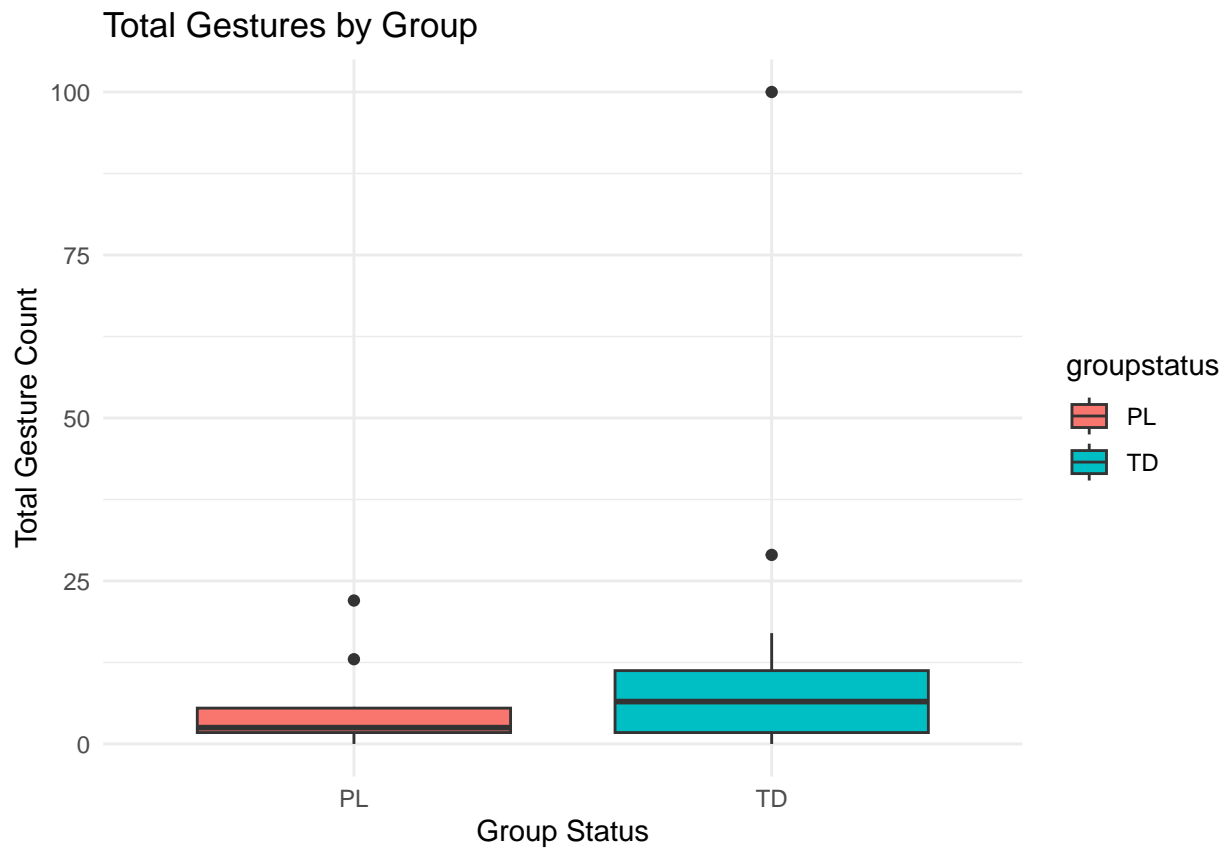
```
##      PL TD
##    0  7  8
##    1  1 12

# Check correlation between total gestures and representational gestures
cor.test(pretend_data_H10$total_gestures, pretend_data_H10$rep_gestures)

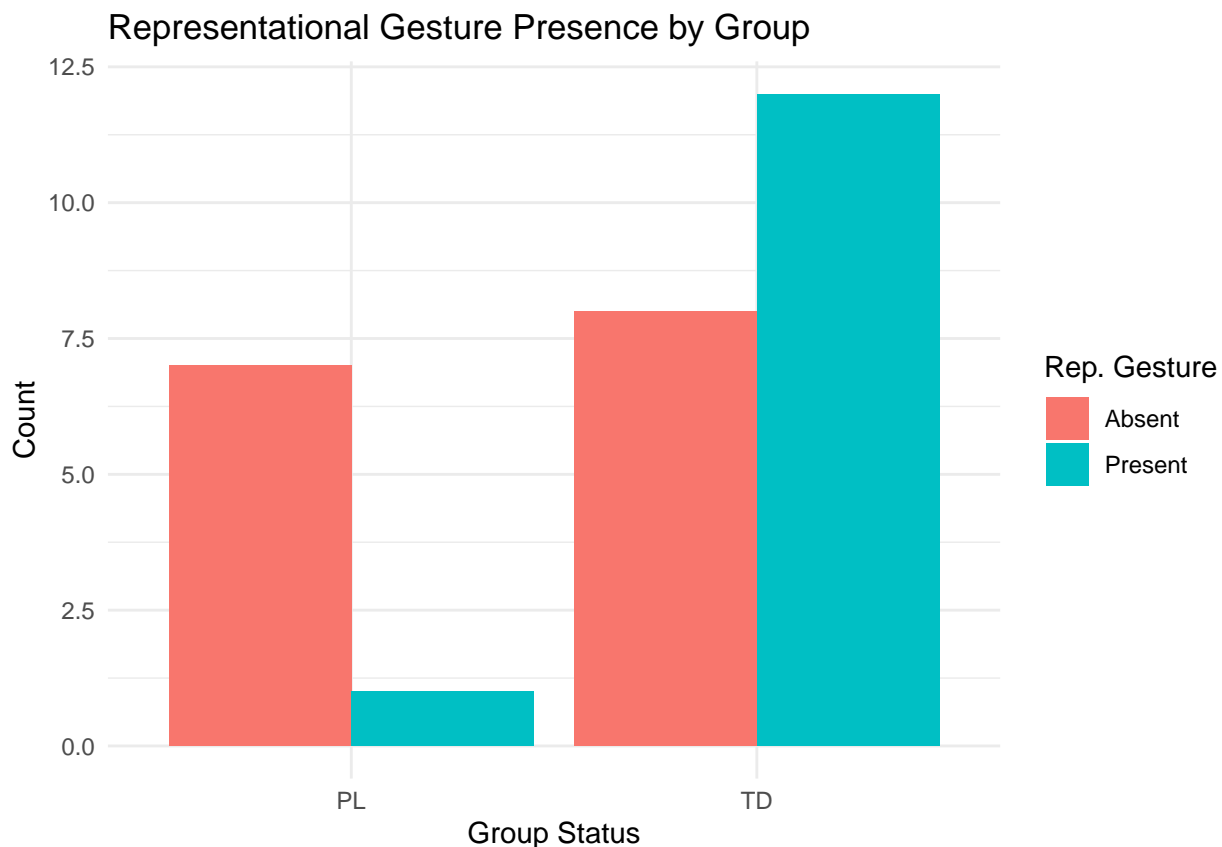
##
## Pearson's product-moment correlation
##
## data: pretend_data_H10$total_gestures and pretend_data_H10$rep_gestures
## t = 12.793, df = 26, p-value = 1.002e-12
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8506549 0.9669106
## sample estimates:
##          cor
## 0.9289276
```

2. Data visualization for gesture in pretend play

```
# Boxplot of total gestures by group
ggplot(pretend_data_H10, aes(x = groupstatus, y = total_gestures, fill = groupstatus)) +
  geom_boxplot() +
  theme_minimal() +
  labs(title = "Total Gestures by Group",
       x = "Group Status",
       y = "Total Gesture Count")
```



```
# Bar chart of representational gesture presence by group
ggplot(pretend_data_H10, aes(x = groupstatus, fill = factor(rep_gesture_present))) +
  geom_bar(position = "dodge") +
  scale_fill_discrete(name = "Rep. Gesture", labels = c("Absent", "Present")) +
  theme_minimal() +
  labs(title = "Representational Gesture Presence by Group",
       x = "Group Status",
       y = "Count")
```



3. Merge with narrative outcome data

```
narrative_data <- read.csv(here("PN_Datasets", "Narrative", "CSVs_of_Combined_Data_N", "Narrative_Resul

# Rename participant column
#narrative_data <- narrative_data %>%
# rename(participant_id = ChildID)

write.csv(narrative_data, here("PN_DataAnalysis", "PP_Narrative_Analysis", "narrative_data.csv"))

# MERGE IT
completely_merged_data_H10 <- left_join(pretend_data_H10, narrative_data, by = "participant_id") %>%
  mutate(groupstatus = if_else(groupstatus == "BI", "PL", groupstatus))

write_csv(completely_merged_data_H10, here("PN_DataAnalysis", "PP_Narrative_Analysis", "completely_merg
```

4. Run regression models

```
# Model 1: Total gestures as predictor
model1 <- lm(max_avg ~ groupstatus + total_gestures + total_pretend_episodes + mlu + groupstatus:total_g
  data = completely_merged_data_H10)

# Model 2: Representational gesture presence as predictor
model2 <- lm(max_avg ~ groupstatus + rep_gesture_present + total_pretend_episodes + mlu + groupstatus:r
```

```

data = completely_merged_data_H10)

summary(model1)

##
## Call:
## lm(formula = max_avg ~ groupstatus + total_gestures + total_pretend_episodes +
##     mlu + groupstatus:total_gestures, data = completely_merged_data_H10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6016 -1.7586 -0.4604  1.0360  4.4275
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      5.2341584   1.7909294    2.923  0.00789 **
## groupstatusTD     -0.0999778   1.2105783   -0.083  0.93493
## total_gestures    -0.0984640   0.1131446   -0.870  0.39356
## total_pretend_episodes  0.0009609   0.0051899    0.185  0.85481
## mlu                0.1783474   0.4074356    0.438  0.66585
## groupstatusTD:total_gestures  0.1012752   0.1134494    0.893  0.38168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.245 on 22 degrees of freedom
## Multiple R-squared:  0.0584, Adjusted R-squared:  -0.1556
## F-statistic: 0.2729 on 5 and 22 DF,  p-value: 0.9231

```

```

summary(model2)

##
## Call:
## lm(formula = max_avg ~ groupstatus + rep_gesture_present + total_pretend_episodes +
##     mlu + groupstatus:rep_gesture_present, data = completely_merged_data_H10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8418 -1.5303 -0.6303  1.3761  4.6169
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      4.6858819   1.7636793    2.657  0.0144 *
## groupstatusTD     0.1152767   1.2773378    0.090  0.9289
## rep_gesture_present -0.8370101   2.5574458   -0.327  0.7465
## total_pretend_episodes  0.0007274   0.0042860    0.170  0.8668
## mlu                0.2093954   0.4356461    0.481  0.6355
## groupstatusTD:rep_gesture_present  1.2717905   2.8202981    0.451  0.6564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.272 on 22 degrees of freedom
## Multiple R-squared:  0.03579, Adjusted R-squared:  -0.1834
## F-statistic: 0.1633 on 5 and 22 DF,  p-value: 0.9734

```

```
# Calculate effect sizes
```

```
eta_squared(model1)
```

```
## # Effect Size for ANOVA (Type I)
```

```
##
```

```
## Parameter | Eta2 (partial) | 95% CI
```

```
## -----
```

```
## groupstatus | 0.02 | [0.00, 1.00]
```

```
## total_gestures | 2.79e-04 | [0.00, 1.00]
```

```
## total_pretend_episodes | 3.88e-04 | [0.00, 1.00]
```

```
## mlu | 5.25e-03 | [0.00, 1.00]
```

```
## groupstatus:total_gestures | 0.03 | [0.00, 1.00]
```

```
##
```

```
## - One-sided CIs: upper bound fixed at [1.00].
```

```
eta_squared(model2)
```

```
## # Effect Size for ANOVA (Type I)
```

```
##
```

```
## Parameter | Eta2 (partial) | 95% CI
```

```
## -----
```

```
## groupstatus | 0.02 | [0.00, 1.00]
```

```
## rep_gesture_present | 2.37e-03 | [0.00, 1.00]
```

```
## total_pretend_episodes | 7.14e-04 | [0.00, 1.00]
```

```
## mlu | 5.36e-03 | [0.00, 1.00]
```

```
## groupstatus:rep_gesture_present | 9.16e-03 | [0.00, 1.00]
```

```
##
```

```
## - One-sided CIs: upper bound fixed at [1.00].
```

Clean the data for plotting

```
# Recode Factors for clarity
```

```
clean_data_H10 <- completely_merged_data_H10 %>%
```

```
  mutate(
```

```
    groupstatus = factor(groupstatus, levels = c("PL", "TD"), labels = c("Perinatal Lesions", "Typically"))
```

```
    rep_gesture_present = factor(rep_gesture_present, levels = c(0, 1), labels = c("Absent", "Present"))
```

```
  )
```

```
# Rename Variables for clarity
```

```
clean_data_H10 <- clean_data_H10 %>%
```

```
  rename(
```

```
    "Narrative Score" = max_avg,
```

```
    "Instances of Pretend Play" = total_pretend_episodes,
```

```
    "Group Status" = groupstatus,
```

```
    "Representative Gesture" = rep_gesture_present,
```

```
    "Total Gestures" = total_gestures,
```

```
    "Mean Length of Utterance" = mlu
```

```
  )
```

```
write_csv(clean_data_H10, here("PN_Datasets", "Pretend_Play", "CSVs_of_Combined_Data_PP", "clean_data_H10.csv"))
```

```
library(ggplot2)
```

```
library(dplyr)
```



```

# Refit the model using the cleaned dataset and new variable names
model1_clean <- lm(`Narrative Score` ~ `Group Status` * `Total Gestures` +
  `Instances of Pretend Play` + `Mean Length of Utterance`,
  data = clean_data_H10)

model2_clean <- lm(`Narrative Score` ~ `Group Status` * `Representative Gesture` +
  `Instances of Pretend Play` + `Mean Length of Utterance`,
  data = clean_data_H10)

```

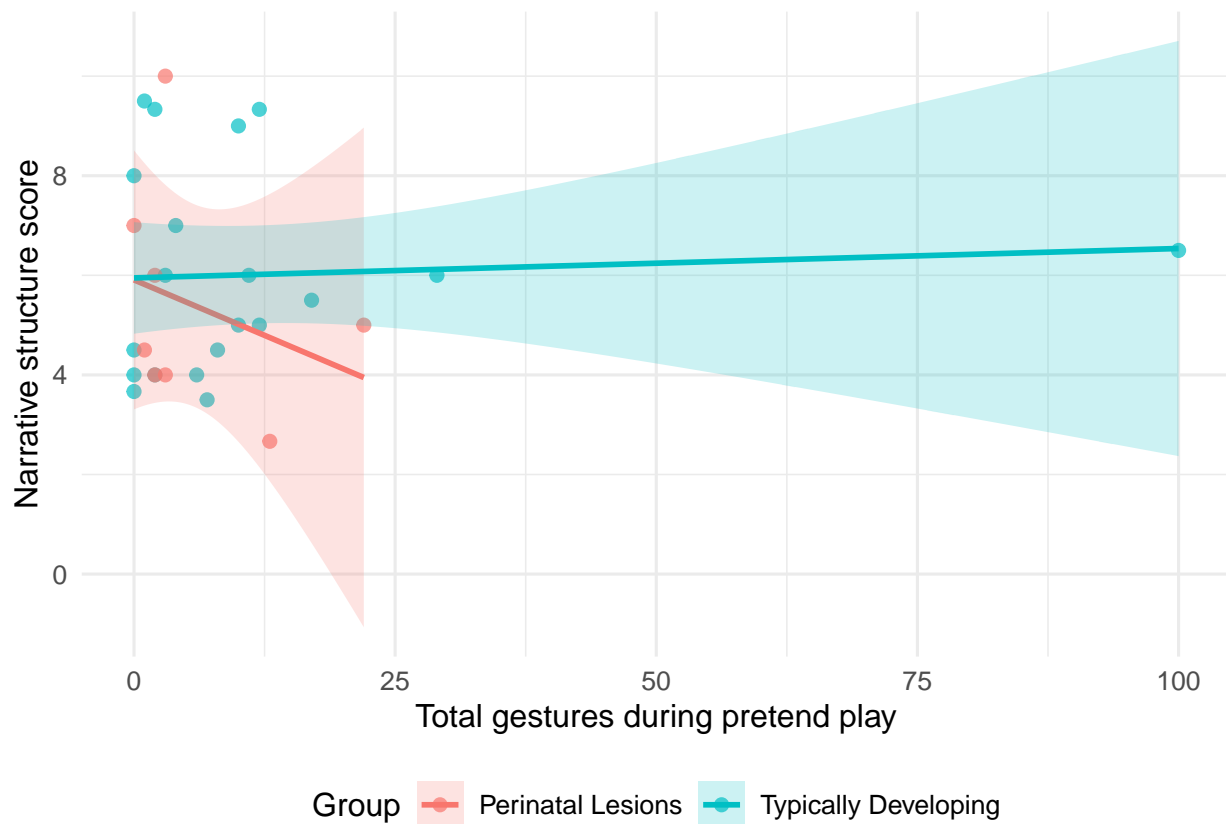
Interaction Plot relationship between total gestures and narrative score, colored by groupstatus.

```

ggplot(clean_data_H10, aes(x = `Total Gestures`, y = `Narrative Score`, color = `Group Status`)) +
  geom_point(alpha = 0.7, size = 2) +
  geom_smooth(method = "lm", se = TRUE, aes(fill = `Group Status`), alpha = 0.2) +
  labs(
    title = NULL,
    x = "Total gestures during pretend play",
    y = "Narrative structure score",
    color = "Group",
    fill = "Group"
  ) +
  theme_minimal(base_size = 12) +
  theme(legend.position = "bottom")

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

```



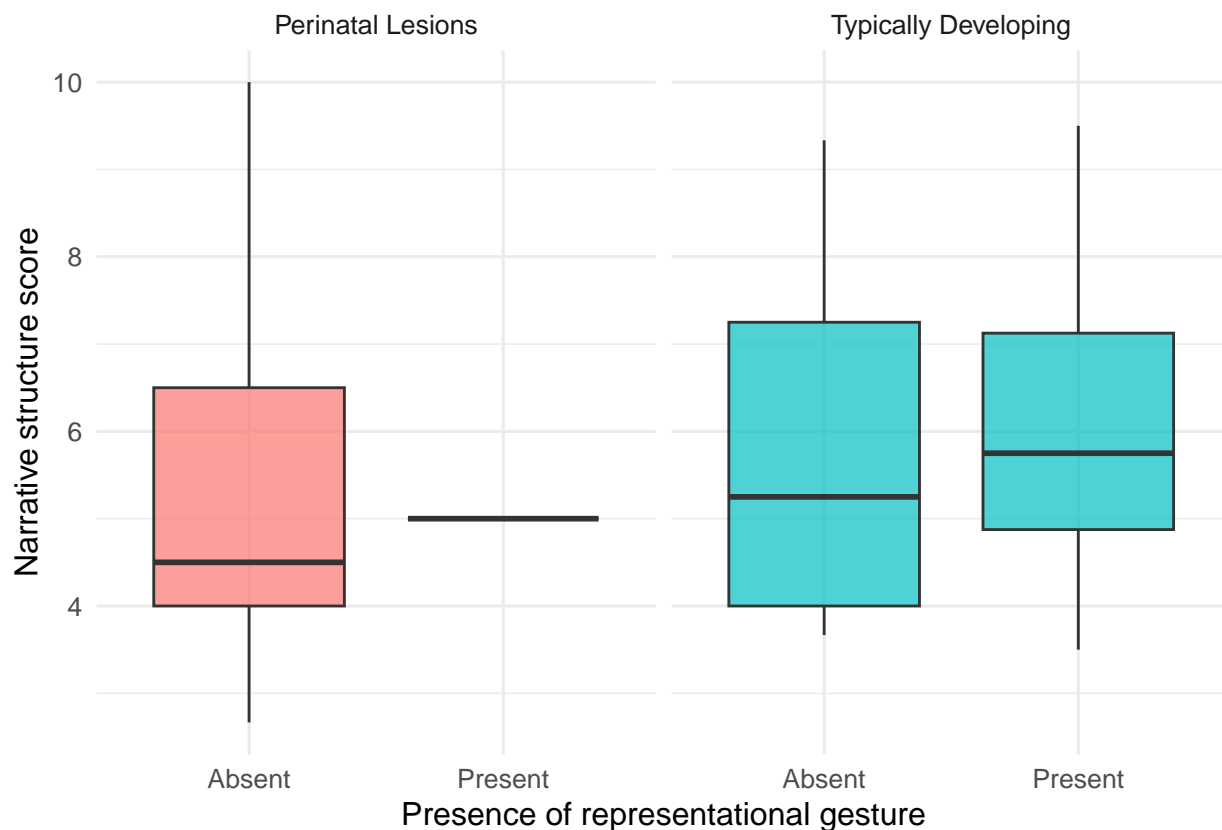
APA Figure Caption

Figure 1

#Interaction between total gestures during pretend play and group status in predicting children's narra

Boxplot of narrative structure score by representational gesture presence and group

```
ggplot(clean_data_H10, aes(x = `Representative Gesture`, y = `Narrative Score`, fill = `Group Status`))
  geom_boxplot(alpha = 0.7) +
  facet_wrap(~ `Group Status`) +
  labs(
    title = NULL,
    x = "Presence of representational gesture",
    y = "Narrative structure score",
    fill = "Group"
  ) +
  theme_minimal(base_size = 12) +
  theme(legend.position = "none")
```



#APA Figure Caption

#Figure 1

#Interaction between total gestures during pretend play and group status in predicting children's narra

Predicted narrative structure scores from regression model

```
#install.packages("ggeffects") # Only run this once
library(ggeffects)

model2_clean <- lm(`Narrative Score` ~ `Group Status` * `Representative Gesture` +
  `Instances of Pretend Play` + `Mean Length of Utterance`,
  data = clean_data_H10)

effect_df <- ggpredict(model2_clean, terms = c("Representative Gesture", "Group Status"))

## Warning: Looks like you are using syntactically invalid variable names, quoted in
## backticks: `Mean Length of Utterance`. This may result in unexpected
## behaviour. Please rename your variables (e.g.,
## `Mean.Length.of.Utterance` instead of `Mean Length of Utterance`) and
## fit the model again.

ggplot(effect_df, aes(x = x, y = predicted, color = group)) +
  geom_line(size = 1.2) +
  geom_point(size = 2) +
  geom_ribbon(aes(ymin = conf.low, ymax = conf.high, fill = group), alpha = 0.2, color = NA) +
  labs(
```

```

title = NULL,
x = "Presence of representational gesture",
y = "Predicted narrative structure score",
color = "Group",
fill = "Group"
) +
theme_minimal(base_size = 12) +
theme(legend.position = "bottom")

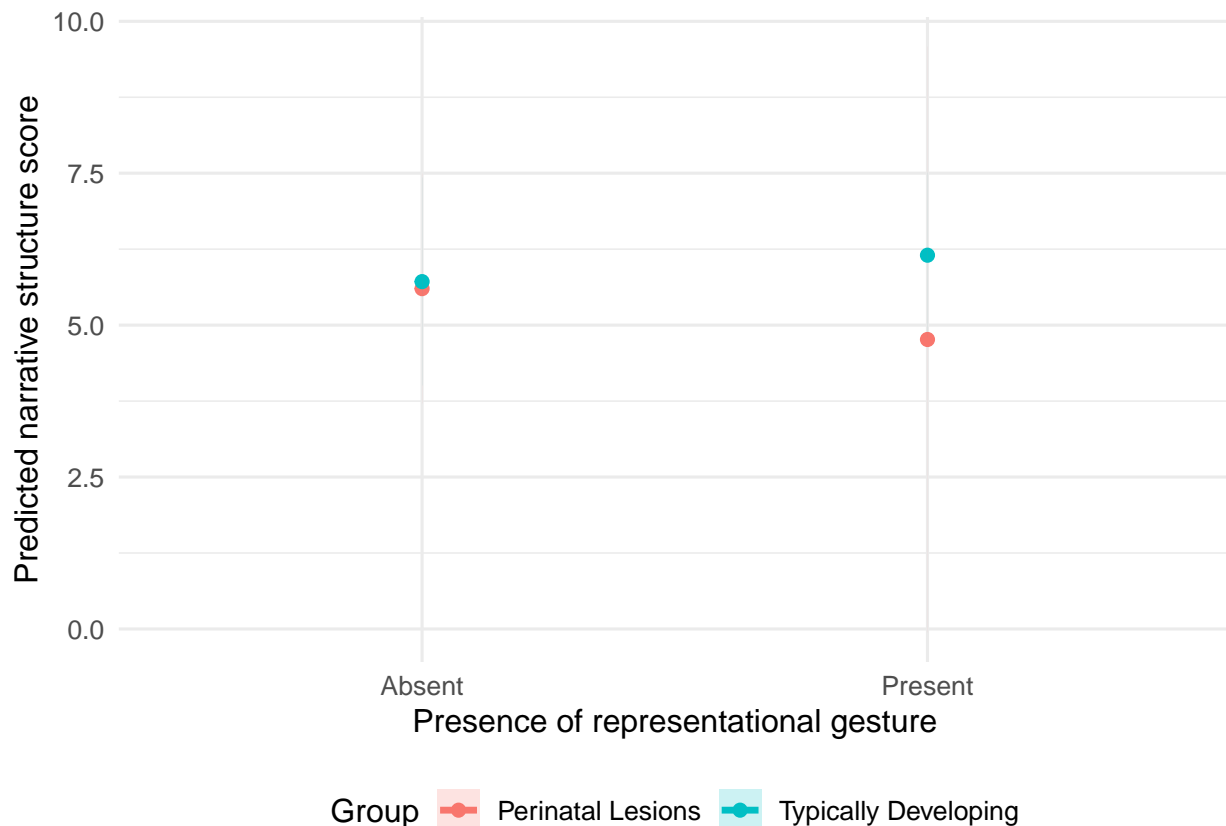
```

```

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.

## `geom_line()`: Each group consists of only one observation.
## i Do you need to adjust the group aesthetic?

```



#APA Figure Caption

#Figure 3

#Predicted narrative structure scores by group and representational gesture presence. Model estimates i