NPB Collaboration Analysis Plan

Sydney Wood

2024-09-17

# Descriptives

Summarize the characteristics of the sample. Especially first generation status, native language, english proficiency, SES and racial composition.

# Confirmatory analysis Multilevel Model

## Creating fake data for analysis

packages <- c("ggplot2", "tidyverse", "apaTables", "colourpicker", "dplyr", "gridExtra", "knitr", "lme4", "reshape2", "stargazer", "gtsummary", "performance", "afex", "jtools", "svglite")  
  
# Install packages not yet installed  
installed\_packages <- packages %in% rownames(installed.packages())  
if (any(installed\_packages == FALSE)) {  
 install.packages(packages[!installed\_packages])  
}  
  
# Packages loading  
invisible(lapply(packages, library, character.only = TRUE))

## Warning: package 'ggplot2' was built under R version 4.3.3

## Warning: package 'tidyr' was built under R version 4.3.3

## Warning: package 'readr' was built under R version 4.3.3

## Warning: package 'dplyr' was built under R version 4.3.3

## Warning: package 'stringr' was built under R version 4.3.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ lubridate 1.9.3 ✔ tibble 3.2.1  
## ✔ purrr 1.0.2 ✔ tidyr 1.3.1  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

## Warning: package 'colourpicker' was built under R version 4.3.2

##   
## Attaching package: 'gridExtra'  
##   
## The following object is masked from 'package:dplyr':  
##   
## combine

## Warning: package 'knitr' was built under R version 4.3.3

## Warning: package 'lme4' was built under R version 4.3.3

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

## Warning in check\_dep\_version(): ABI version mismatch:   
## lme4 was built with Matrix ABI version 1  
## Current Matrix ABI version is 0  
## Please re-install lme4 from source or restore original 'Matrix' package

##   
## Attaching package: 'reshape2'  
##   
## The following object is masked from 'package:tidyr':  
##   
## smiths  
##   
##   
## Please cite as:   
##   
## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.  
## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

## Warning: package 'gtsummary' was built under R version 4.3.3

## Warning: package 'performance' was built under R version 4.3.3

## Warning: package 'afex' was built under R version 4.3.3

## \*\*\*\*\*\*\*\*\*\*\*\*  
## 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

## Warning: package 'jtools' was built under R version 4.3.3

## Warning: package 'svglite' was built under R version 4.3.3

anon\_id <- c(1:300)  
section <- sample(1:12, 300, replace = T)  
TA <- sample(1:3, 300, replace = T)  
Exam1 <- rnorm(300, mean = 80, sd = 5)  
Exam2 <- rnorm(300, mean = 85, sd = 7)  
  
fake\_data <- data.frame(anon\_id, section, TA, Exam1, Exam2)  
  
fake\_data <- fake\_data %>% mutate(TA = case\_when(section == 1 ~ 1, section == 2 ~ 1, section == 3 ~ 1, section == 4 ~ 2, section == 5 ~ 2, section == 6 ~ 2, section == 7 ~ 3, section == 8 ~ 3, section == 9 ~ 3, section == 10 ~ 4, section == 11 ~ 4, section == 12 ~ 4))  
  
fake\_data <- fake\_data %>% mutate(condition = case\_when(section == 1 | section == 5 | section == 11 | section == 8 ~ 1, section == 3 | section == 9 | section == 4 | section == 10 ~ 2, section == 2 | section == 6 | section == 12 | section == 7 ~ 3))  
  
fake\_data\_long <- gather(fake\_data, key = "obs",value = "score" , Exam1, Exam2)

I didn’t model any shared variance in the nested grouping structure so the proposed models will report singularity. This is very unlikely to occur in the real data, even if there are no inherent differences between discussion groups. Therefore the code for fitting the model is commented out

## Confirmatory Analysis

eq0 <- (score ~ obs + (1|anon\_id))  
#mod0 <- lmer(eq0, data = fake\_data\_long)  
#summary(mod0)  
  
  
  
eq\_5 <- (score ~ obs + condition + (1|anon\_id))  
#mod\_5 <- lmer(eq\_5, data = fake\_data\_long)  
#summary(mod\_5)  
  
  
eq1 <- (score ~ obs\*condition + (1|anon\_id))  
#mod1 <- lmer(eq1, data = fake\_data\_long)  
#summary(mod1)  
  
  
eq2 <- (score ~ obs\*condition + (1|anon\_id) + (1|anon\_id:section))  
#mod2 <- lmer(eq2, data = fake\_data\_long)  
#summary(mod2)  
  
eq3 <- (score ~ obs\*condition + (1|anon\_id) + (1|anon\_id:section) + (1|anon\_id:section:TA))  
#mod3 <- lmer(eq3, data = fake\_data\_long)  
#summary(mod3)

#Exploratory analyses

These models and analyses will be created post hoc as we do not have specific hypotheses regarding effect of various demographic or course related variables.