489 Data

Preston Lyons

25/05/2021

## Select the demographic variables and rename them  
  
library(table1)

##   
## Attaching package: 'table1'

## The following objects are masked from 'package:base':  
##   
## units, units<-

library(kableExtra)  
  
df2 <- df %>%  
 dplyr::select("Q6", "Q7", "Q9", "Q10")  
  
df2 <- rename(df2, Age = Q6, Gender = Q7, Education\_Level = Q9, Parent = Q10) %>%  
 mutate(Age = factor(Age,   
 levels = c(1,2,3,4,5,6),  
 labels = c("(18 – 24)", "(25 – 34)",  
 "(35 – 44)", "(45 – 54)", "(55 – 64)",  
 "(65+)")),  
 Gender = factor(Gender,   
 levels = c(1, 2, 3),  
 labels = c("Male", "Female", "Non-Binary")),  
 Education\_Level = factor(Education\_Level,  
 levels = c(1, 2, 3),  
 labels = c("None", "Secondary", "Tertiary")),  
 Parent = factor(Parent,  
 levels = c(1, 2),  
 labels = c("Yes", "No")))  
  
tabnz <- df2 %>%  
 select(Age, Gender, Education\_Level, Parent)  
  
df2\_described <-   
prettyR::describe(df2)

## Description of df2

## Descriptive Statistics  
  
df1$id\_test\_child <- rowMeans(df1[id\_test\_child], na.rm = T)  
df1$id\_mem\_child <- rowMeans(df1[id\_mem\_child], na.rm = T)  
df1$id\_sug\_child <- rowMeans(df1[id\_sug\_child], na.rm = T)  
  
df1$td\_test\_3\_5 <- rowMeans(df1[td\_test\_3\_5], na.rm = T)  
df1$td\_mem\_3\_5 <- rowMeans(df1[td\_mem\_3\_5], na.rm = T)  
df1$td\_sug\_3\_5 <- rowMeans(df1[td\_sug\_3\_5], na.rm = T)  
  
df1$td\_test\_6\_11 <- rowMeans(df1[td\_test\_6\_11], na.rm = T)  
df1$td\_mem\_6\_11 <- rowMeans(df1[td\_mem\_6\_11], na.rm = T)  
df1$td\_sug\_6\_11 <- rowMeans(df1[td\_sug\_6\_11], na.rm = T)  
  
df1$id\_test\_adult <- rowMeans(df1[id\_test\_adult], na.rm = T)  
df1$id\_mem\_adult <- rowMeans(df1[id\_mem\_adult], na.rm = T)  
df1$id\_sug\_adult <- rowMeans(df1[id\_sug\_adult], na.rm = T)  
  
df1$td\_test\_adult <- rowMeans(df1[td\_test\_adult], na.rm = T)  
df1$td\_mem\_adult <- rowMeans(df1[td\_mem\_adult], na.rm = T)  
df1$td\_sug\_adult <- rowMeans(df1[td\_sug\_adult], na.rm = T)  
  
domain\_list <- list("id\_mem\_child",  
 "id\_test\_child",   
 "id\_sug\_child",  
 "td\_mem\_3\_5",  
 "td\_test\_3\_5",  
 "td\_sug\_3\_5",  
 "td\_mem\_6\_11",  
 "td\_test\_6\_11",  
 "td\_sug\_6\_11",  
 "id\_mem\_adult",  
 "id\_test\_adult",  
 "id\_sug\_adult",  
 "td\_mem\_adult",  
 "td\_test\_adult",  
 "td\_sug\_adult")  
  
measure\_names <- c("Memory - CWID",  
 "Ability to Testify - CWID",  
 "Suggestibility - CWID",  
 "Memory - TD 3-5 year olds",  
 "Ability to Testify in Court - TD 3-5 year olds",  
 "Suggestibility - TD 3-5 year olds",  
 "Memory - TD 6-11 year olds",  
 "Ability to Testify in Court - TD 6-11 year olds",  
 "Suggestibility - TD 6-11 year olds",  
 "Memory - AWID",  
 "Ability to Testify - AWID",  
 "Suggestibility - AWID",  
 "Memory - TD Adult",  
 "Ability to Testify - TD Adult",  
 "Suggestibility - TD Adult"  
 )  
  
domain\_out <-   
lapply(domain\_list, function(x){  
 data.frame(Mean = round(mean(df1[[x]], na.rm = T), 2),  
 SD = round(sd(df1[[x]], na.rm = T), 2)  
 )  
 }) %>%  
 do.call(rbind, .) %>%  
 cbind(measure\_names, .)  
  
apa\_table(domain\_out, caption = "Means and standard deviations of measures in the study")

(#tab:data4)

*Means and standard deviations of measures in the study*

|  |  |  |
| --- | --- | --- |
| measure\_names | Mean | SD |
| Memory - CWID | 2.98 | 1.32 |
| Ability to Testify - CWID | 3.41 | 0.99 |
| Suggestibility - CWID | 4.07 | 1.03 |
| Memory - TD 3-5 year olds | 2.66 | 1.40 |
| Ability to Testify in Court - TD 3-5 year olds | 3.18 | 0.93 |
| Suggestibility - TD 3-5 year olds | 4.06 | 1.03 |
| Memory - TD 6-11 year olds | 3.59 | 1.33 |
| Ability to Testify in Court - TD 6-11 year olds | 3.82 | 0.87 |
| Suggestibility - TD 6-11 year olds | 4.02 | 0.90 |
| Memory - AWID | 3.67 | 1.30 |
| Ability to Testify - AWID | 3.85 | 0.97 |
| Suggestibility - AWID | 3.71 | 1.02 |
| Memory - TD Adult | 5.24 | 1.37 |
| Ability to Testify - TD Adult | 4.92 | 0.64 |
| Suggestibility - TD Adult | 3.08 | 1.12 |

## LEMR (Linear Effects Mixed Regression)  
df1$id <- paste0("id\_", 1:nrow(df1))  
library(lme4)  
library(lmerTest)  
library(tidyverse)  
library(scales)

##   
## Attaching package: 'scales'

## The following objects are masked from 'package:psych':  
##   
## alpha, rescale

## The following object is masked from 'package:purrr':  
##   
## discard

## The following object is masked from 'package:readr':  
##   
## col\_factor

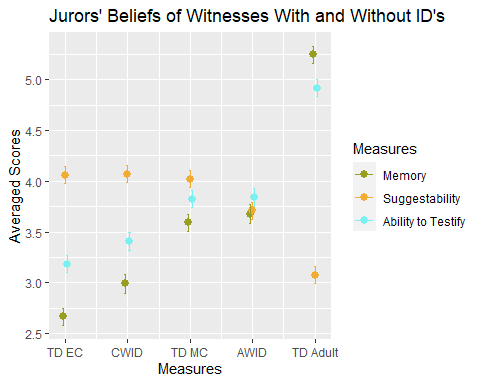
df1\_long <- df1[c("id\_test\_child", "id\_mem\_child", "id\_sug\_child", "td\_test\_3\_5", "td\_mem\_3\_5", "td\_sug\_3\_5", "td\_test\_6\_11", "td\_mem\_6\_11", "td\_sug\_6\_11", "id\_test\_adult", "id\_mem\_adult", "id\_sug\_adult", "td\_test\_adult", "td\_mem\_adult", "td\_sug\_adult", "id")] %>%  
 pivot\_longer(., -id) %>%  
 separate(., "name", into = c("type", "cat", "age", "age2")) %>%  
 mutate(., target = paste0(type, age)) %>%  
 mutate(., target = factor(target, levels = c("td3", "idchild", "td6", "idadult", "tdadult")))

## Warning: Expected 4 pieces. Missing pieces filled with `NA` in 6102 rows [1, 2,  
## 3, 10, 11, 12, 13, 14, 15, 16, 17, 18, 25, 26, 27, 28, 29, 30, 31, 32, ...].

lmer\_out <- lmer("value ~ target \* cat + (1|id)", data = df1\_long)  
  
sjPlot::plot\_model(lmer\_out, type = "int") +  
 labs(title = "Jurors' Beliefs of Witnesses With and Without ID's",  
 x = "Measures",  
 y = "Averaged Scores") +  
 labs(color = "Measures") +  
 scale\_color\_manual(values = c("#989E21", "#F0AD32", "#7AF0F0"), labels = c("Memory", "Suggestability", "Ability to Testify")) +  
 scale\_x\_continuous(labels=c("td3" = "TD EC",  
 "idchild" = "CWID",  
 "td6" = "TD MC",  
 "idadult" = "AWID",  
 "tdadult" = "TD Adult"))

## Scale for 'colour' is already present. Adding another scale for 'colour',  
## which will replace the existing scale.

## Scale for 'x' is already present. Adding another scale for 'x', which will  
## replace the existing scale.



report::report(lmer\_out)

## We fitted a linear mixed model (estimated using REML and nloptwrap optimizer) to predict value with target and cat (formula: value ~ target \* cat). The model included id as random effect (formula: ~1 | id). The model's total explanatory power is substantial (conditional R2 = 0.39) and the part related to the fixed effects alone (marginal R2) is of 0.27. The model's intercept, corresponding to target = td3 and cat = mem, is at 2.66 (95% CI [2.58, 2.75], t(9417) = 62.07, p < .001). Within this model:  
##   
## - The effect of target [idchild] is statistically significant and positive (beta = 0.33, 95% CI [0.21, 0.44], t(9417) = 5.53, p < .001; Std. beta = 0.25, 95% CI [0.16, 0.34])  
## - The effect of target [td6] is statistically significant and positive (beta = 0.93, 95% CI [0.82, 1.04], t(9417) = 16.80, p < .001; Std. beta = 0.72, 95% CI [0.64, 0.81])  
## - The effect of target [idadult] is statistically significant and positive (beta = 1.01, 95% CI [0.90, 1.13], t(9417) = 16.96, p < .001; Std. beta = 0.79, 95% CI [0.70, 0.88])  
## - The effect of target [tdadult] is statistically significant and positive (beta = 2.58, 95% CI [2.47, 2.69], t(9417) = 46.70, p < .001; Std. beta = 2.00, 95% CI [1.92, 2.09])  
## - The effect of cat [sug] is statistically significant and positive (beta = 1.40, 95% CI [1.29, 1.50], t(9417) = 25.28, p < .001; Std. beta = 1.08, 95% CI [1.00, 1.17])  
## - The effect of cat [test] is statistically significant and positive (beta = 0.52, 95% CI [0.41, 0.63], t(9417) = 9.41, p < .001; Std. beta = 0.40, 95% CI [0.32, 0.49])  
## - The interaction effect of cat [sug] on target [idchild] is statistically significant and negative (beta = -0.32, 95% CI [-0.48, -0.16], t(9417) = -3.87, p < .001; Std. beta = -0.24, 95% CI [-0.37, -0.12])  
## - The interaction effect of cat [sug] on target [td6] is statistically significant and negative (beta = -0.97, 95% CI [-1.12, -0.82], t(9417) = -12.45, p < .001; Std. beta = -0.75, 95% CI [-0.87, -0.63])  
## - The interaction effect of cat [sug] on target [idadult] is statistically significant and negative (beta = -1.37, 95% CI [-1.53, -1.20], t(9417) = -16.60, p < .001; Std. beta = -1.06, 95% CI [-1.19, -0.94])  
## - The interaction effect of cat [sug] on target [tdadult] is statistically significant and negative (beta = -3.56, 95% CI [-3.72, -3.41], t(9417) = -45.87, p < .001; Std. beta = -2.77, 95% CI [-2.89, -2.65])  
## - The interaction effect of cat [test] on target [idchild] is statistically non-significant and negative (beta = -0.10, 95% CI [-0.26, 0.06], t(9417) = -1.28, p = 0.202; Std. beta = -0.08, 95% CI [-0.20, 0.04])  
## - The interaction effect of cat [test] on target [td6] is statistically significant and negative (beta = -0.29, 95% CI [-0.44, -0.14], t(9417) = -3.74, p < .001; Std. beta = -0.23, 95% CI [-0.34, -0.11])  
## - The interaction effect of cat [test] on target [idadult] is statistically significant and negative (beta = -0.36, 95% CI [-0.52, -0.20], t(9417) = -4.36, p < .001; Std. beta = -0.28, 95% CI [-0.40, -0.15])  
## - The interaction effect of cat [test] on target [tdadult] is statistically significant and negative (beta = -0.85, 95% CI [-1.00, -0.70], t(9417) = -10.92, p < .001; Std. beta = -0.66, 95% CI [-0.78, -0.54])  
##   
## Standardized parameters were obtained by fitting the model on a standardized version of the dataset. 95% Confidence Intervals (CIs) and p-values were computed using the Wald approximation.

# Try to get estimated marginal means from the emmeans package  
  
emdf <- emmeans::emmeans(lmer\_out, specs = pairwise ~ target:cat)

## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.  
## To enable adjustments, add the argument 'pbkrtest.limit = 9434' (or larger)  
## [or, globally, 'set emm\_options(pbkrtest.limit = 9434)' or larger];  
## but be warned that this may result in large computation time and memory use.

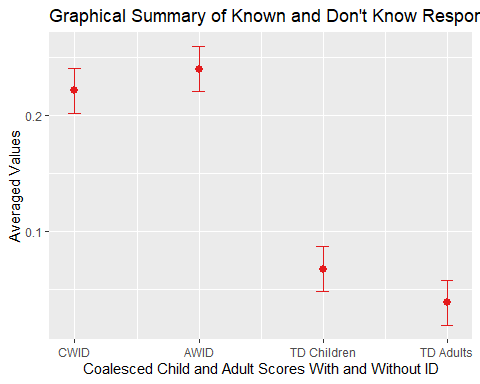
## Note: D.f. calculations have been disabled because the number of observations exceeds 3000.  
## To enable adjustments, add the argument 'lmerTest.limit = 9434' (or larger)  
## [or, globally, 'set emm\_options(lmerTest.limit = 9434)' or larger];  
## but be warned that this may result in large computation time and memory use.

emdf1 <- emdf$contrasts %>%  
 data.frame() %>%  
 dplyr::select(., contrast, estimate, p.value) %>%  
 mutate(p.value = round(p.value, 3))  
  
report::report(emdf1)

## The data contains 105 observations of the following variables:  
## - contrast: 105 entries, such as idadult mem - idadult sug (0.95%%); idadult mem - idadult test (0.95%%); idadult mem - idchild sug (0.95%%) and 102 others (0 missing)  
## - estimate: n = 105, Mean = -0.24, SD = 0.94, Median = -0.23, MAD = 0.91, range: [-2.58, 2.17], Skewness = 0.11, Kurtosis = 2.51e-03, 0% missing  
## - p.value: n = 105, Mean = 0.11, SD = 0.29, Median = 0.00, MAD = 0.00, range: [0, 1], Skewness = 2.55, Kurtosis = 4.86, 0% missing

## Analyse the DK responses  
  
df1$dn\_id <- rowMeans(df1[paste0(c(id\_test\_child, id\_mem\_child, id\_sug\_child), "\_bin")])  
df1$dn\_td <- rowMeans(df1[paste0(c(td\_test\_3\_5, td\_mem\_3\_5, td\_sug\_3\_5, td\_test\_6\_11, td\_mem\_6\_11, td\_sug\_6\_11), "\_bin")])  
df1$dn\_id\_ad <- rowMeans(df1[paste0(c(id\_test\_adult, id\_mem\_adult, id\_sug\_adult), "\_bin")])  
df1$dn\_td\_ad <- rowMeans(df1[paste0(c(td\_test\_adult, td\_mem\_adult, td\_sug\_adult), "\_bin")])  
  
dn\_long <-   
 df1 %>%  
 select(., dn\_id, dn\_td, dn\_id\_ad, dn\_td\_ad, id) %>%  
 pivot\_longer(., -id)  
  
lmer\_out <- lmer(value ~ name + (1|id), dn\_long)  
  
lm\_out <- lm(value ~ name , dn\_long)  
  
sjPlot::plot\_model(lmer\_out, type = "pred")$name +  
 labs(title = "Graphical Summary of Known and Don't Know Responses",  
 x = "Coalesced Child and Adult Scores With and Without ID",  
 y = "Averaged Values") +  
 scale\_x\_continuous(labels=c("dn\_id" = "CWID",  
 "dn\_id\_ad" = "AWID",  
 "dn\_td" = "TD Children",  
 "dn\_td\_ad" = "TD Adults"))+  
 aes(color = "green") +  
 theme(legend.position = "none")

## Scale for 'x' is already present. Adding another scale for 'x', which will  
## replace the existing scale.



summary(lmer\_out)

## Linear mixed model fit by REML. t-tests use Satterthwaite's method [  
## lmerModLmerTest]  
## Formula: value ~ name + (1 | id)  
## Data: dn\_long  
##   
## REML criterion at convergence: -357.6  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -2.2526 -0.6309 0.0338 0.3925 3.5368   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## id (Intercept) 0.03182 0.1784   
## Residual 0.03454 0.1858   
## Number of obs: 2712, groups: id, 678  
##   
## Fixed effects:  
## Estimate Std. Error df t value Pr(>|t|)   
## (Intercept) 2.210e-01 9.893e-03 1.603e+03 22.336 <2e-16 \*\*\*  
## namedn\_id\_ad 1.864e-02 1.009e-02 2.031e+03 1.846 0.065 .   
## namedn\_td -1.533e-01 1.009e-02 2.031e+03 -15.190 <2e-16 \*\*\*  
## namedn\_td\_ad -1.825e-01 1.009e-02 2.031e+03 -18.079 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Correlation of Fixed Effects:  
## (Intr) nmdn\_d\_ nmdn\_t  
## namedn\_id\_d -0.510   
## namedn\_td -0.510 0.500   
## namedn\_td\_d -0.510 0.500 0.500

summary(lm\_out)

##   
## Call:  
## lm(formula = value ~ name, data = dn\_long)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.23961 -0.22097 -0.03915 0.03312 0.93235   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.220971 0.009893 22.336 <2e-16 \*\*\*  
## namedn\_id\_ad 0.018638 0.013991 1.332 0.183   
## namedn\_td -0.153325 0.013991 -10.959 <2e-16 \*\*\*  
## namedn\_td\_ad -0.182489 0.013991 -13.043 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2576 on 2708 degrees of freedom  
## Multiple R-squared: 0.1078, Adjusted R-squared: 0.1068   
## F-statistic: 109 on 3 and 2708 DF, p-value: < 2.2e-16

#poisson distribution 489  
  
emmeans::emmeans(lmer\_out, pairwise ~ name)

## $emmeans  
## name emmean SE df lower.CL upper.CL  
## dn\_id 0.2210 0.00989 1603 0.2016 0.2404  
## dn\_id\_ad 0.2396 0.00989 1603 0.2202 0.2590  
## dn\_td 0.0676 0.00989 1603 0.0482 0.0871  
## dn\_td\_ad 0.0385 0.00989 1603 0.0191 0.0579  
##   
## Degrees-of-freedom method: kenward-roger   
## Confidence level used: 0.95   
##   
## $contrasts  
## contrast estimate SE df t.ratio p.value  
## dn\_id - dn\_id\_ad -0.0186 0.0101 2031 -1.846 0.2519   
## dn\_id - dn\_td 0.1533 0.0101 2031 15.190 <.0001   
## dn\_id - dn\_td\_ad 0.1825 0.0101 2031 18.079 <.0001   
## dn\_id\_ad - dn\_td 0.1720 0.0101 2031 17.036 <.0001   
## dn\_id\_ad - dn\_td\_ad 0.2011 0.0101 2031 19.925 <.0001   
## dn\_td - dn\_td\_ad 0.0292 0.0101 2031 2.889 0.0204   
##   
## Degrees-of-freedom method: kenward-roger   
## P value adjustment: tukey method for comparing a family of 4 estimates