Grey matter volume in ASD subgroups and correlations with language abilities

Setup

```
packages <- c("here","readxl","dplyr","ggpubr","ggplot2","psych","mgcv","caret",
         "MASS")

lapply(packages, library, character.only = TRUE)

# load functions
#source("code/pplot.R")
source("code/diff_volume.R")
```

Demographic information in ASD subgroups and TD

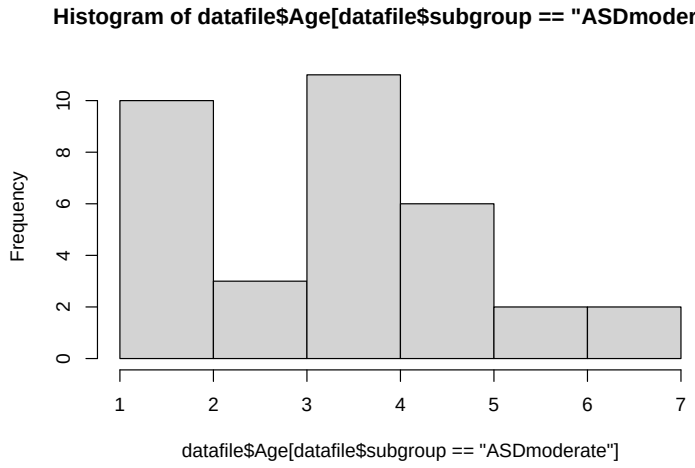
```
# load data
datafile <- read.csv("data/datafile_clean.csv", header = T)

table(datafile$subgroup, datafile$Sex)
```

```
##
##                 1  2
##   ASDmoderate 32  2
##   ASDsevere   25  9
##   TD          32  5
```

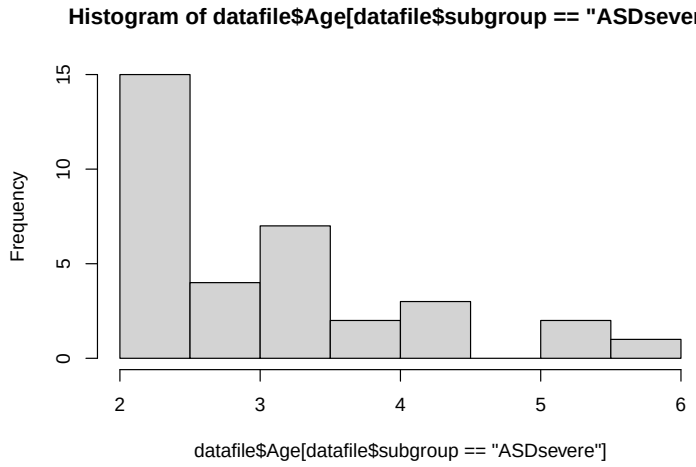
```
table(datafile$Dx)
```

```
##
## ASD  TD
##  68  37
```

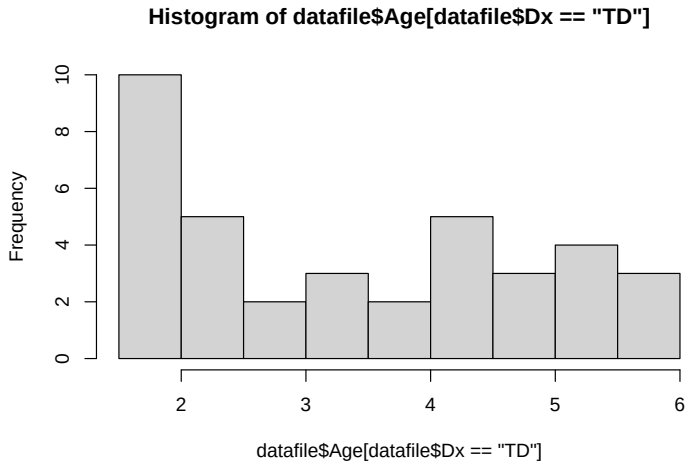
```
# age distribution in ASD and TD
hist(datafile$Age[datafile$subgroup == "ASDmoderate"])
```




```
hist(datafile$Age[datafile$subgroup == "ASDsevere"])
```




```
hist(datafile$Age[datafile$Dx == "TD"])
```




```
table(datafile$Sex[datafile$subgroup != "TD"])
```

```
##
##  1  2
## 57 11
```

```
# summarize demographic information
demo_all <- as.data.frame(as.matrix(0, 8,12))

datafile$subgroup <- factor(datafile$subgroup, levels = c("TD", "ASDmoderate",
                                "ASDsevere"))
mean_all <- describeBy(datafile[,c(3,5,6:11:16)], group ="subgroup")

demo_all[1:12, 1:6] <-
    cbind(paste0(round(mean_all[[1]]$mean[c(1,3:13)],2),"±",
            round(mean_all[[1]]$sd[c(1,3:13)],2)),
          paste0(round(mean_all[[1]]$min[c(1,3:13)],2),"-",
                 round(mean_all[[1]]$max[c(1,3:13)],2)),
          paste0(round(mean_all[[2]]$mean[c(1,3:13)],2),"±",
                 round(mean_all[[2]]$sd[c(1,3:13)],2)),
          paste0(round(mean_all[[2]]$min[c(1,3:13)],2),"-",
                 round(mean_all[[2]]$max[c(1,3:13)],2)),
          paste0(round(mean_all[[3]]$mean[c(1,3:13)],3),"±",
                 round(mean_all[[3]]$sd[c(1,3:13)],3)),
          paste0(round(mean_all[[3]]$min[c(1,3:13)],3),"-",
                 round(mean_all[[3]]$max[c(1,3:13)],3)))

rownames(demo_all) <- colnames(datafile[,c(3,6:11:16)])

for(tt in colnames(datafile)[c(3,6:11:16)]) {

    if (all(is.na(datafile[datafile$subgroup == "TD", tt]))) {
        myt3 <- t.test(datafile[datafile$subgroup == "ASDmoderate", tt],
              datafile[datafile$subgroup == "ASDsevere", tt],na.action = T)

        demo_all[tt, 7:12] <- c(NA,NA,NA,NA,round(myt3$statistic,2),
```

```
                    round(myt3$p.value,3))

    } else {
        myt1 <- t.test(datafile[datafile$subgroup == "TD", tt],
              datafile[datafile$subgroup == "ASDmoderate", tt],na.action = T)
        myt2 <- t.test(datafile[datafile$subgroup == "TD", tt],
              datafile[datafile$subgroup == "ASDsevere", tt],na.action = T)
        myt3 <- t.test(datafile[datafile$subgroup == "ASDmoderate", tt],
              datafile[datafile$subgroup == "ASDsevere", tt],na.action = T)

        demo_all[tt, 7:12] <- c(round(myt1$statistic,2),
            round(myt1$p.value,3),
            round(myt2$statistic,2),round(myt2$p.value,3),
            round(myt3$statistic,2),round(myt3$p.value,3))
    }


}

colnames(demo_all) <- c("TD_mean","TD_range",
            "ASD_ASDmoderate_mean","ASD_ASDmoderate_range",
            "ASD_ASDsevere_mean","ASD_ASDsevere_range",
            "TDvsASDsevere_t","TDvsASDmoderate_p",
            "TDvsASDsevere_t","TDvsASDsevere_p",
            "ASDmodvsASDsev_t","ASDmodvsASDsev_p")

#View(demo_all)


# gender
median(datafile$Gesell_Lang[datafile$Dx == "ASD"])
```

```
## [1] 44.4
```

```
k <- table(datafile$subgroup, datafile$Sex)
tk <- chisq.test(k)

datafile$subgroup <- as.character(datafile$subgroup)

k1 <- table(datafile$subgroup[datafile$subgroup !="ASDsevere"],
       datafile$Sex[datafile$subgroup !="ASDsevere"])
tk1 <- chisq.test(k1)

k2 <- table(datafile$subgroup[datafile$subgroup !="ASDmoderate"],
       datafile$Sex[datafile$subgroup !="ASDmoderate"])
tk2 <- chisq.test(k2)

k3 <- table(datafile$subgroup[datafile$subgroup !="TD"],
       datafile$Sex[datafile$subgroup !="TD"])
tk3 <- chisq.test(k3)

demo_all["Sex", 1:12] <- c(paste0(k[1,1], "/",k[1,2]), " ",
              paste0(k[2,1], "/",k[2,2]), " ",
```

```
              paste0(k[3,1], "/",k[3,2]), " ",
              round(tk1$statistic,3), round(tk1$p.value,3),
              round(tk2$statistic,3), round(tk2$p.value,3),
              round(tk3$statistic,3), round(tk3$p.value,3))

dim(demo_all)
```

```
## [1] 13 12
```

```
#View(demo_all)

demo_all_new <- demo_all[c(1,13,3:4,6,5,2,7,12,8:11), ]
#View(demo_all_new)

# correcting for multiple comparisons using FDR
row.names(demo_all_new)
```

```
##  [1] "Age"            "Sex"            "Gesell_GrMot"   "Gesell_FineMot"
##  [5] "Gesell_Social"  "Gesell_Lang"    "Gesell_Adpt"    "Gesell_Total"
##  [9] "ABC"            "ADOS_SA"        "ADOS_RRB"       "ADOS_Total"
## [13] "CARS"
```

```
p.adjust(c(demo_all_new$TDvsASDmoderate_p[9],demo_all_new$TDvsASDsevere_p[9],
     demo_all_new$ASDmodvsASDsev_p[9]), method = "fdr")
```

```
## [1] 0.000 0.000 0.051
```

```
p.adjust(c(demo_all_new$TDvsASDmoderate_p[13],demo_all_new$TDvsASDsevere_p[13],
     demo_all_new$ASDmodvsASDsev_p[13]), method = "fdr")
```

```
## [1] 0.000 0.000 0.002
```

```
knitr::kable(demo_all)
```

	TD_mean_range	ASD_ASDmoderate_range	ASD_ASDsevere_range	TDvsASDmoderate_t	TDvsASDmoderate_p	TDvsASDsevere_t	TDvsASDsevere_p	ASDmodvsASDsev_t	ASDmodvsASDsev_p
Age	3.39±1.49 (1.5-5.87)	3.31±1.33 (1.52-6.53)	3.039±0.98 (1.57-5.57)	0.25	0.807	1.19	0.239	0.96	0.341
Gesell_Adpt	93.79±5.39 (80-107)	66.18±8.76 (40.9-82)	55.068±13.64 (31-93.7)	15.82	0	15.79	0	4.06	0
Gesell_GrMot	97.42±6.11 (80-107)	77.76±8.74 (57.8-91)	71.791±8.52 (52-87.5)	10.94	0	14.67	0	2.88	0.005
Gesell_FineMot	96.73±6.34 (80-108)	73.31±9.39 (54.1-90.4)	64.247±10.29 (45.2-84.1)	12.77	0	16.58	0	3.81	0
Gesell_Lang	91.47±7.57 (76-101)	55.07±7.76 (44.6-70.7)	37.112±5.96 (26.6-44.2)	22.55	0	39.86	0	10.73	0
Gesell_Social	93.76±8.52 (80-103)	60.46±7.43 (31-70.2)	48.156±7.79 (27.9-64.2)	22.65	0	29.86	0	6.68	0
Gesell_Total	94.31±8.22 (76-100.8)	66.72±5.64 (49.1-79.4)	55.254±7.91 (39.3-68.74)	23.27	0	27.85	0	7.38	0

	TD_mean_range	ASD_mean_range	ASDmoderate_mean_range	ASDsevere_mean_range	TDvsASD_t	TDvsASD_p	ASDmodvsASDsev_t	ASDmodvsASDsev_p
ADOS_Total	NA±NA NA-Inf	12.41±3.67 7-20	13.04±3.25 9-22	NA	NA	NA	NA	-0.67	0.505
ADOS_RRB	NA±NA NA-Inf	1.24±0.91 0-4	1.52±1.04 0-4	NA	NA	NA	NA	-1.04	0.306
ADOS_Social	NA±NA NA-Inf	13.66±4.08 8-23	14.56±3.90 9-24	NA	NA	NA	NA	-0.83	0.41
CARS	14.33±1.315 15	31.92±2.03 30-38	35±4.967 25-45	-22.69	0	-19.11	0	-3.28	0.002
ABC	25.3±8.856 43	66.65±14.28 36-107	74.971±15.806 0	-14.66	0	-13.51	0	-1.99	0.051
Sex	32/5	32/2	25/9	0.461	0.497	1.15	0.284	3.904	0.048

```
write.csv(demo_all_new,"results/demographic_info_subgroups.csv", row.names = T)

writexl::write_xlsx(demo_all_new,"results/demographic_info_subgroups.xlsx" ,
            col_names = T)

# ASD subjects who had ADOS or CARS
subj_info_ASD <- datafile[datafile$Dx == "ASD", ]
dim(subj_info_ASD)
```

```
## [1] 68 20
```

```
length(which(!is.na(datafile[datafile$subgroup == "ASDmoderate", "ADOS_Total"])))
```

```
## [1] 29
```

```
length(which(!is.na(datafile[datafile$subgroup == "ASDsevere", "ADOS_Total"])))
```

```
## [1] 25
```

```
length(which(!is.na(datafile[datafile$subgroup == "ASDmoderate", "CARS"])))
```

```
## [1] 26
```

```
length(which(!is.na(datafile[datafile$subgroup == "ASDsevere", "CARS"])))
```

```
## [1] 34
```

```
k <- dim(subj_info_ASD[!is.na(subj_info_ASD$ADOS_Total),])[1]
kk <- dim(subj_info_ASD[!is.na(subj_info_ASD$CARS),])[1]
kkk <- dim(datafile[!is.na(datafile$CARS[datafile$Dx == "TD"]),])[1]

table(subj_info_ASD$Sex[!is.na(subj_info_ASD$ADOS_Total)])
```

```
##
##  1  2
## 47  7
```

```
table(subj_info_ASD$Sex[!is.na(subj_info_ASD$CARS)])
```

```
##
##  1  2
## 49 11
```

```
print(paste0("n = ", k, " children with ASD had ADOS"))
```

```
## [1] "n = 54 children with ASD had ADOS"
```

```
print(paste0("n = ", kk, " children with ASD had CARS"))
```

```
## [1] "n = 60 children with ASD had CARS"
```

```
print(paste0("n = ", kkk, " TD children had CARS"))
```

```
## [1] "n = 8 TD children had CARS"
```

Plot demographic information in ASD subgroups and TD

```
datafile$subgroup <- factor(datafile$subgroup, levels = c("TD","ASDmoderate",

for (beh in colnames(datafile)[6:11]) {

    # boxplot
    p <- ggplot(datafile,aes_string(x = "subgroup", beh))+
        #theme(axis.title = element_blank()) + #delete x,y,title
        geom_boxplot(width = 0.3,fill = c('#F94700','#77933B','#498399')) +
        geom_jitter() +
        #geom_jitter(size = 4,width = 0.2, alpha = 0.4, col = "blue")+
        theme_classic() +# set y/x color as black, set backgroud as white
        #scale_x_discrete(labels=c("nonASD" = "non-ASD")) +
        theme(legend.position  = c(1.5,0.5),
              axis.text.y =element_text(size=16,face = "bold",color = 'black'),
              axis.text.x =  element_blank(),
              axis.title = element_blank(),
              #axis.title.y = element_text(size=18,face="bold"),
              axis.line.x = element_line(size = 1,color = 'black'),
              axis.line.y = element_line(size = 1,color = 'black'))
    #   ylab(ylab)# change y label +

    p
        ggsave(here(paste0("results/beh_diff_",beh,".ASDsubgroups.png")),
               height = 3, width = 4)

}
```

Gruop differences in global volume of GM, WM, CSF, and TIV

```
# differences between ASD and TD
p1 <- t.test(datafile$TIV[datafile$Dx == "TD"], datafile$TIV[datafile$Dx == "ASD"])

p2 <- t.test(datafile$GM[datafile$Dx == "TD"], datafile$GM[datafile$Dx == "ASD"])

p3 <- t.test(datafile$WM[datafile$Dx == "TD"], datafile$WM[datafile$Dx == "ASD"])

p4 <- t.test(datafile$CSF[datafile$Dx == "TD"], datafile$CSF[datafile$Dx == "ASD"])

p.adjust(c(p1$p.value,p2$p.value,p3$p.value,p4$p.value))
```

```
## [1] 0.20408126 0.07143551 0.51325799 0.51325799
```

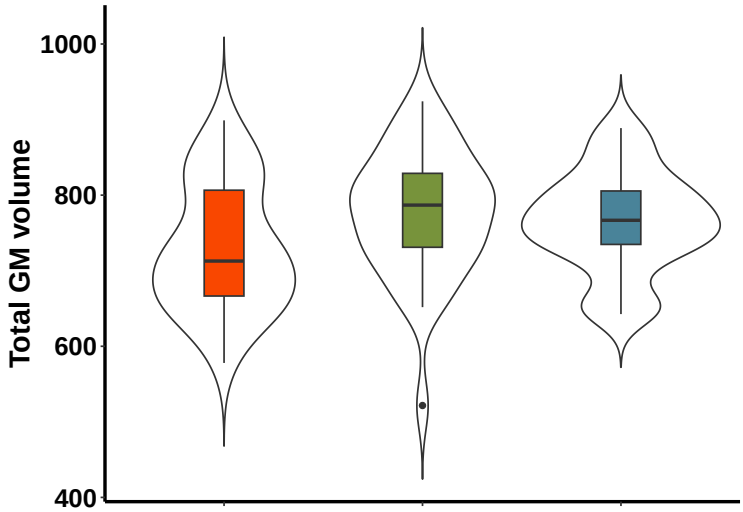
```
# regression analysis controlling for age, sex, and TIV
diff_all <- as.data.frame(matrix(0, 4, 4))

i <- 0
for (Vol in c("GM","WM","CSF","TIV")) {
    i <- i+1

    if (Vol != "TIV") {
        ll <- anova(lm(paste0(Vol, " ~ Dx + Age + Sex + TIV"), datafile))
    } else {
        ll <- anova(lm(paste0(Vol, " ~ Dx + Age + Sex"), datafile))
    }


    diff_all[i, 1] <- Vol
    diff_all[i, 2] <- round(ll$`Pr(>F)`[1], 4)

    #
    dd <- effsize::cohen.d(datafile[datafile$Dx =='ASD', Vol],
                datafile[datafile$Dx =='TD', Vol])

    diff_all[i, 3:4] <- c(round(dd$estimate,2), paste0("[",round(dd$conf.int[1],2),
                                ", ", round(dd$conf.int[2],2),"]"))
}

colnames(diff_all) <- c("measure","p-value","cohenD","CI")

#View(diff_all)

diff_all$`p-value`
```

```
## [1] 0.0000 0.0010 0.3427 0.0113
```

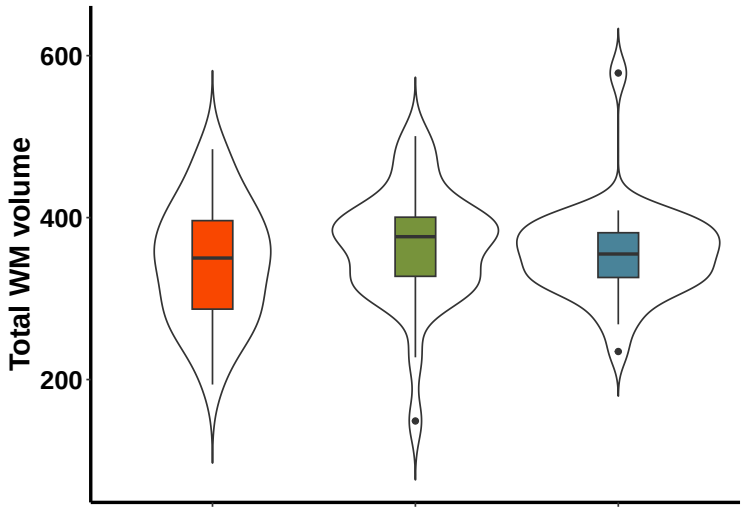
```
p.adjust(diff_all$`p-value`)
```

```
## [1] 0.0000 0.0030 0.3427 0.0226
```

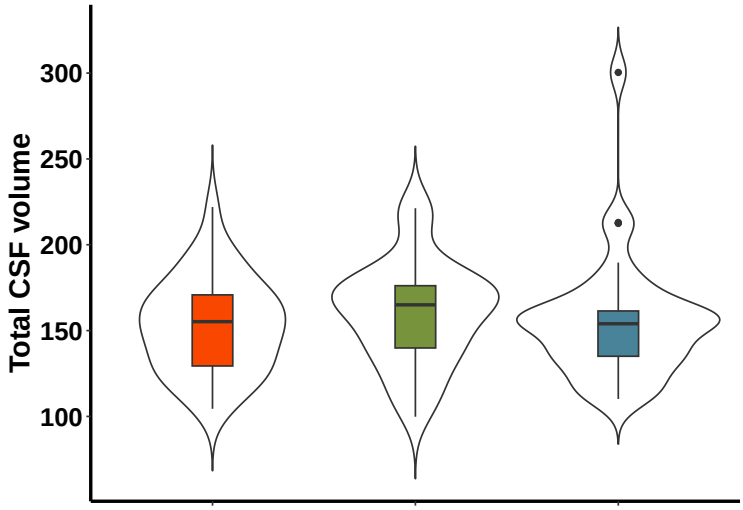
```
writexl::write_xlsx(diff_all,"results/global_volume_diff.xlsx", col_names = T)

# plot for three groups
diff_volume("subgroup","GM","Total GM volume")
```



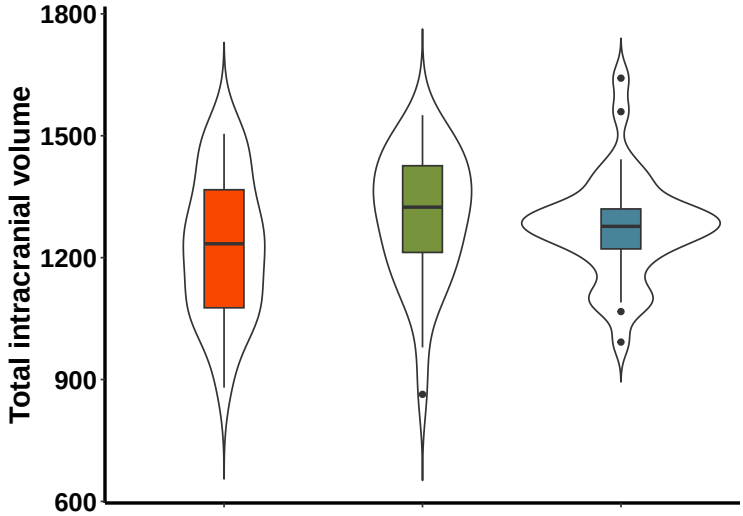

```
diff_volume("subgroup","WM","Total WM volume")
```




```
diff_volume("subgroup","CSF", "Total CSF volume")
```




```
diff_volume("subgroup","TIV", "Total intracranial volume")
```



Normative modeling, revision on Sep.12

```
# install required packages
# load data
datafile <- read.csv("data/datafile_clean.csv", header = T)
td_data <- subset(datafile, subgroup == "TD")
asd_moderate_data <- subset(datafile, subgroup == "ASDmoderate")
asd_severe_data <- subset(datafile, subgroup == "ASDsevere")

norm_tt <- as.data.frame(as.matrix(0, 4, 10))

i <- 0
for (tt in c("GM","WM","CSF","TIV")) {
```

```
    i <- i + 1

# build the normative model for the TD group
    if (tt != "TIV") {
        norm_model <- lm(as.formula(paste0(tt, "~ Age + Sex + TIV")),
                 data = td_data)
    } else {
        norm_model <- lm(as.formula(paste0(tt, "~ Age + Sex")),
                 data = td_data)
    }



# deviations for ASD moderate subgroup
asd_moderate_data$Pred_ASDmoderete <- predict(norm_model, newdata = asd_moderate_data)

# Calculate deviations
asd_moderate_data$Deviation <- asd_moderate_data[,tt] - asd_moderate_data$Pred_ASDmoderete

# deviation_mean <- mean(asd_moderate_data$Deviation)
# deviation_sd <- sd(asd_moderate_data$Deviation)
#
# asd_moderate_data$z_score <- (asd_moderate_data$Deviation - deviation_mean)/deviation_sd

td_data$Predicted_Volume <- predict(norm_model, newdata = td_data)
td_data$Deviation <- td_data[,tt] - td_data$Predicted_Volume

t_test_result1 <- t.test(td_data$Deviation, asd_moderate_data$Deviation)

# deviations for ASD severe subgroup
asd_severe_data$Pred_ASDsevere <- predict(norm_model, newdata = asd_severe_data)

# Calculate deviations
asd_severe_data$Deviation <- asd_severe_data[,tt] - asd_severe_data$Pred_ASDsevere

t_test_result2 <- t.test(td_data$Deviation, asd_severe_data$Deviation)


# ASD moderate vs. ASD severe
t_test_result3 <- t.test(asd_severe_data$Deviation, asd_moderate_data$Deviation)

norm_tt[i, 1] <- tt
norm_tt[i, 2:10] <- c("TD vs. ASDmoderate", round(t_test_result1$statistic,3),
             round(t_test_result1$p.value,3),"TD vs. ASDsevere",
             round(t_test_result2$statistic,3),round(t_test_result2$p.value,3),
             "ASDmoderate vs. ASDsevere",round(t_test_result3$statistic,3),
             round(t_test_result3$p.value,3))

}

#View(norm_tt)

p.adjust(norm_tt$V4,"fdr")
```

```
## [1] 0.384 0.492 0.492 0.016
```

```
p.adjust(norm_tt$V7,"fdr")
```

```
## [1] 0.2080000 0.7410000 0.2813333 0.0160000
```

Differences in global volumes between ASD subgroups and TD

```
diff_all <- as.data.frame(matrix(0, 4, 11))

table(datafile$subgroup)
```

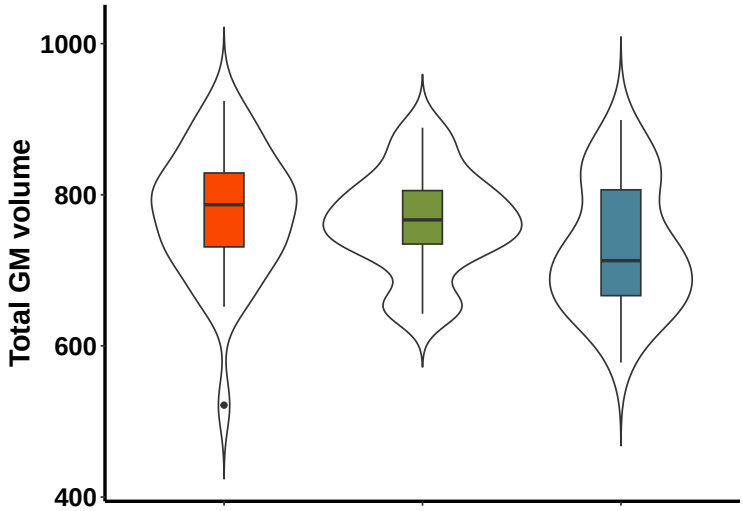
```
##
## ASDmoderate   ASDsevere          TD
##          34          34          37
```

```
i <- 0
for (Vol in c("GM","WM","CSF","TIV")) {
    i <- i+1
    if (Vol != "TIV") {
        ll <- anova(lm(as.formula(paste0(Vol, " ~ subgroup + Age + Sex + TIV")),
                   datafile))
        diff_all[i, 1] <- Vol
        diff_all[i, 2:3] <- c(round(ll$`F value`[1], 3),
                    round(ll$`Pr(>F)`[1], 3))
    } else {
        ll <- anova(lm(as.formula(paste0(Vol, " ~ subgroup + Age + Sex")),
                      datafile))
        diff_all[i, 1] <- Vol
        diff_all[i, 2:3] <- c(round(ll$`F value`[1], 3),
                    round(ll$`Pr(>F)`[1], 3))
    }

    tt <- pairwise.t.test(datafile[,Vol], datafile[, "subgroup"],
                  p.adjust="fdr", pool.sd = T)

    # differences between ASD high vs. TD
    #k1 <- anova(lm(paste0(Vol, " ~ subgroup + Age + Sex"),
    #          datafile[datafile$subgroup != "ASD_lowLang",]))

    dd1 <- effsize::cohen.d(datafile[datafile$subgroup =='TD', Vol],
                    datafile[datafile$subgroup =='ASDmoderate', Vol])

    dd2 <- effsize::cohen.d(datafile[datafile$subgroup =='TD', Vol],
                    datafile[datafile$subgroup =='ASDsevere', Vol])

    dd3 <- effsize::cohen.d(datafile[datafile$subgroup =='ASDmoderate', Vol],
                    datafile[datafile$subgroup =='ASDsevere', Vol])

    diff_all[i, 4:6] <- c(round(tt$p.value[1], 3),round(dd1$estimate,2),
```

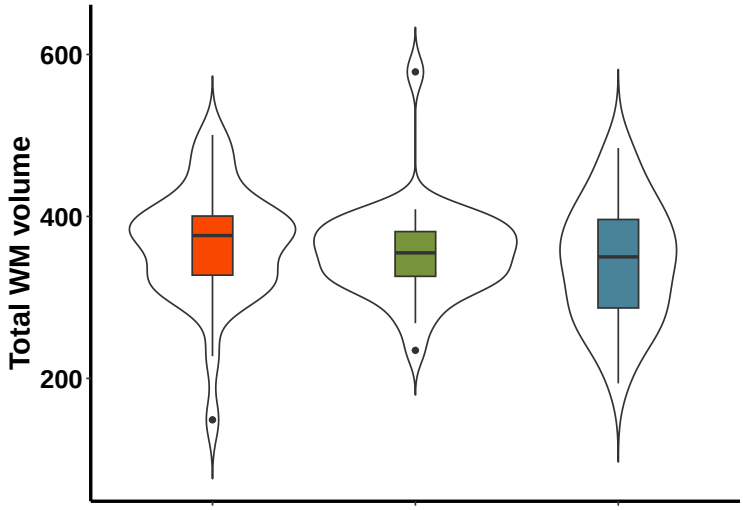
```
                  paste0("[",round(dd1$conf.int[1],2),    ", ",
                         round(dd1$conf.int[2],2),"]"))

    diff_all[i,7:9] <- c(round(tt$p.value[2,1], 3),round(dd2$estimate,2),
                  paste0("[",round(dd2$conf.int[1],2),    ", ",
                         round(dd2$conf.int[2],2),"]"))

    diff_all[i,10:12] <- c(round(tt$p.value[2,2], 3),round(dd3$estimate,2),
                  paste0("[",round(dd3$conf.int[1],2),    ", ",
                         round(dd3$conf.int[2],2),"]"))
}

colnames(diff_all) <- c("Volume", "all.F","all.p","TDvsASDmoderate.p",
            "TDvsASDmoderate.d","TDvsASDmoderate.CI","TDvsASDsevere.p",
            "TDvsASDsevere.d","TDvsASDsevere.CI","ASDmoderateVSsevere.p",
            "ASDmoderateVSsevere.d","ASDmoderateVSsevere.CI")

#View(diff_all)

p.adjust(diff_all$all.p)
```

```
## [1] 0.000 0.003 0.468 0.056
```

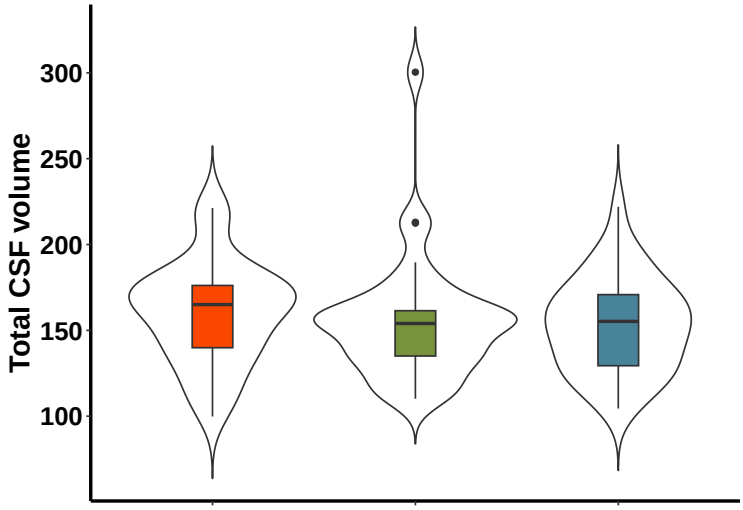
```
# plot for three groups
diff_volume("subgroup","GM","Total GM volume")
```



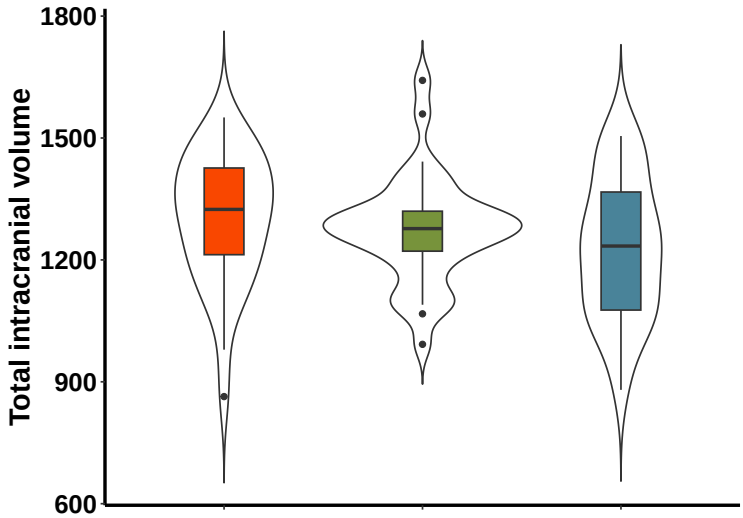

```
diff_volume("subgroup","WM","Total WM volume")
```




```
diff_volume("subgroup","CSF", "Total CSF volume")
```




```
diff_volume("subgroup","TIV", "Total intracranial volume")
```




```
writexl::write_xlsx(diff_all,"results/global_volume_diff_3groups.xlsx", col_names = T)
```

Partial correlations for the significant clusters

```
# load ROI signal
ROIsignal_low <- read.table("data/ROIsignal_low.txt", header = F)
ROIsignal_high <- read.table("data/ROIsignal_high.txt", header = F)
ROIsignal_TD <- read.table("data/ROIsignal_TD.txt", header = F)

colnames(ROIsignal_high) <- c("Lhippocampus", "LSTG","LIPL","Rhippocampus",
                  "RSTG","PCC")
colnames(ROIsignal_low) <- c("Lhippocampus", "LSTG","LIPL","Rhippocampus",
                 "RSTG","PCC")
colnames(ROIsignal_TD) <- c("Lhippocampus", "LSTG","LIPL","Rhippocampus",
                "RSTG","PCC")

dim(ROIsignal_high)
```

```
## [1] 34  6
```

```
dim(ROIsignal_low)
```

```
## [1] 34  6
```

```
dim(ROIsignal_TD)
```

```
## [1] 37  6
```

```
#View(ROIsignal_low)

datafile_ROIsignal <- rbind.data.frame(cbind(datafile[datafile$subgroup=="ASDmoderate", ],
```

```
                    ROIsignal_high),
                    cbind(datafile[datafile$subgroup == "ASDsevere", ],
                    ROIsignal_low),
                    cbind(datafile[datafile$subgroup == "TD", ],
                    ROIsignal_TD))

# Partial correlations
corr_sum <- as.data.frame(matrix(0, 18, 4))

i <- 0
for (gr in c("ASDmoderate", "ASDsevere","TD")) {
for (ROI in c("Lhippocampus", "LSTG", "LIPL","Rhippocampus", "RSTG","PCC")) {
    i <- i+1

pp <- ppcor::pcor.test(datafile_ROIsignal[datafile_ROIsignal$subgroup == gr,
                    "Gesell_Lang"],
                datafile_ROIsignal[datafile_ROIsignal$subgroup == gr,ROI],
                datafile_ROIsignal[datafile_ROIsignal$subgroup == gr,
                        c("Age","Sex","TIV")])

        corr_sum[i, 1:2] <- c(gr,ROI)

        corr_sum[i, 3:4] <- c(round(pp$estimate,2), pp$p.value)

        permutation_results <- numeric(5000)

        # run permutation tests
    for (k in 1:5000) {

        set.seed(123)

        shuffled_data <- datafile_ROIsignal[datafile_ROIsignal$subgroup == gr, ]
        shuffled_data$Gesell_Lang <- sample(shuffled_data[,ROI])

        # Calculate partial correlation for the shuffled data
        pcor_perm <- ppcor::pcor.test(shuffled_data$Gesell_Lang,
                        shuffled_data[, ROI],
                    shuffled_data[, c("Age", "Sex", "TIV")])$estimate

        # Store the result
        permutation_results[k] <- pcor_perm
}
    # Calculate p-value
    p_value <- mean(abs(permutation_results) >= abs(pp$estimate))

    corr_sum[i, 5] <- p_value

}
}

colnames(corr_sum) <- c("group","ROI", "r value", "p value", "adjusted_p_value")

#View(corr_sum)
```

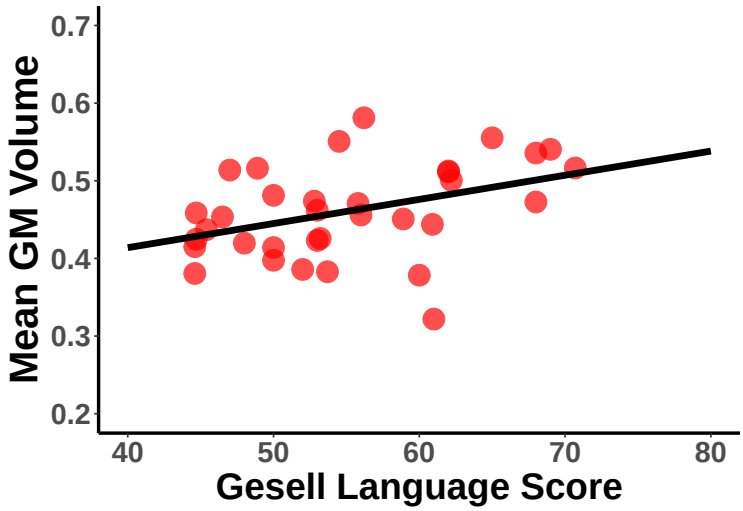
```
writexl::write_xlsx(corr_sum,"results/significant_corr.xlsx")
```

Scatter plots on the significant clusters

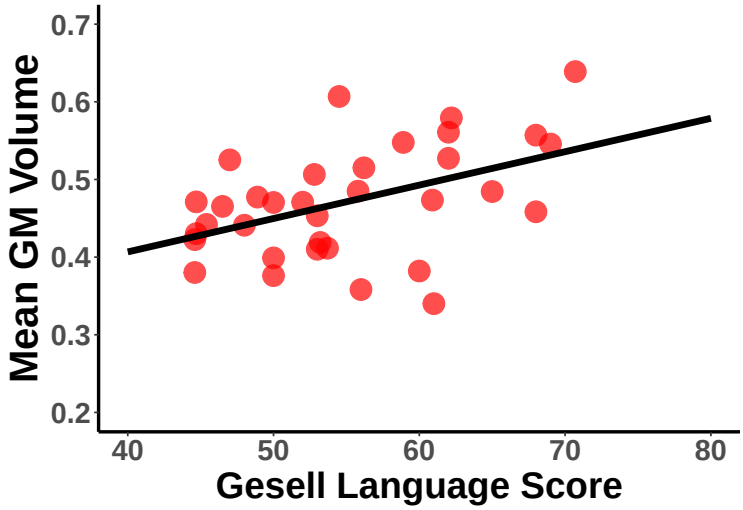
```
# mean and group differences in grey matter volume
gr_diff <- as.data.frame(matrix(0, 6, 10))

i <- 0
for (vol in colnames(ROIsignal_high)) {
    i <- i+1
    gr_diff[i, 1] <- vol
    gr_diff[i, 2] <- mean(datafile_ROIsignal[datafile_ROIsignal$subgroup=="TD",
                          vol])
    gr_diff[i, 3] <- mean(datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDmoderate",
                          vol])
    gr_diff[i, 4] <- mean(datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDsevere",
                          vol])
    t1 <- t.test(datafile_ROIsignal[datafile_ROIsignal$subgroup=="TD", vol],
           datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDmoderate", vol])
    t2 <- t.test(datafile_ROIsignal[datafile_ROIsignal$subgroup=="TD", vol],
           datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDsevere", vol])
    t3 <- t.test(datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDmoderate",
                     vol],
           datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDsevere", vol])

    gr_diff[i, 5:6] <- c(round(t1$statistic,3),round(t1$p.value,3))
    gr_diff[i, 7:8] <- c(round(t2$statistic,3),round(t2$p.value,3))
    gr_diff[i, 9:10] <- c(round(t3$statistic,3),round(t3$p.value,3))

}

#View(gr_diff)

colnames(gr_diff) <- c("region","mean_TD", "mean_ASDhigh", "mean_ASDlow", "TDvsASDhigh-t",
               "TDvsASDhigh-p", "TDvsASDlow-t", "TDvsASDlow-p", "ASDhighvsASDlow-t",
               "ASDhighvsASDlow-p")

#View(gr_diff)


#View(datafile_ROIsignal)

# correlation plots
# plot
pplot <- function(dat, ROI, xlabel,ll,xl) {

    p <- ggplot(dat,aes_string(x="Gesell_Lang",y=ROI, xmin = 40, xmax = 80)) +
        geom_point(col = "red",size = 6, alpha = 0.7) +
        geom_smooth(method = lm, se = F, size = 2,col = 'black',fullrange = T)+
        theme_classic()+
```

```
        xlab(xlabel)+ ylab('Mean GM Volume') +
        guides(colour = "none") +
        coord_cartesian(ylim= ll,xlim=xl) +
        theme(axis.text=element_text(size=18,face = "bold"),
              axis.title.x = element_text(size=23,face="bold"),
              axis.title.y = element_text(size=25,face="bold"),
              axis.line.x = element_line(size = 1),
              axis.line.y = element_line(size = 1))

    print(p)
    ggsave(here(paste0("results/",ROI,'_Lang.png')),width = 5,height = 4)
}


xlabel <- "Gesell Language Score"

for (ROI in c("Lhippocampus", "LSTG","Rhippocampus", "RSTG","LIPL")) {
    pplot(datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDmoderate",],
          ROI, xlabel, c(0.2,0.7), c(40,80))
}
```

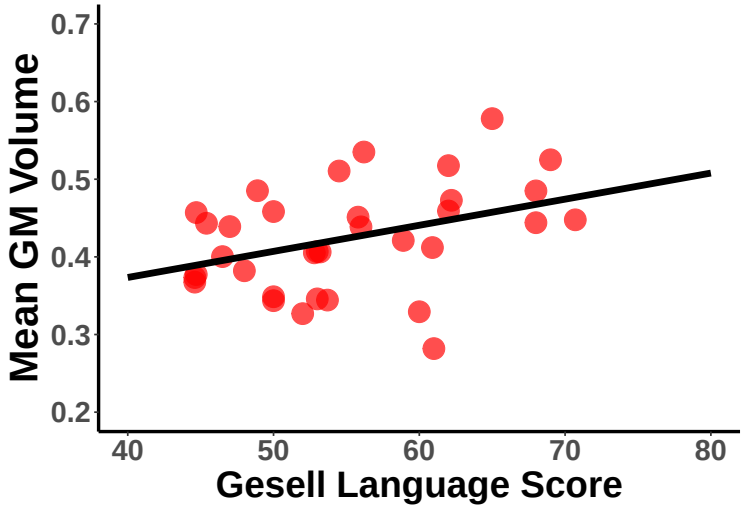
```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```



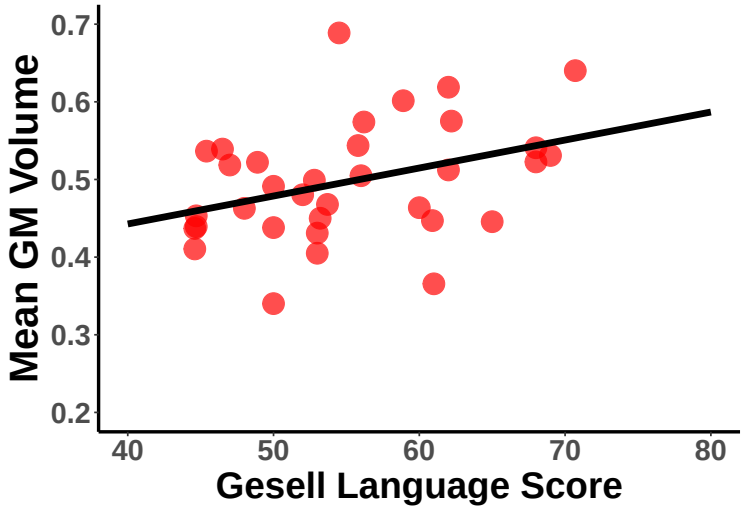

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```



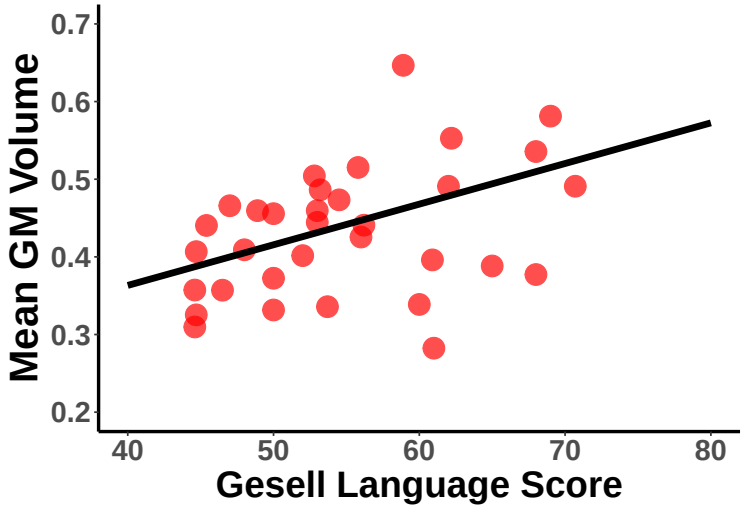

```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```




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

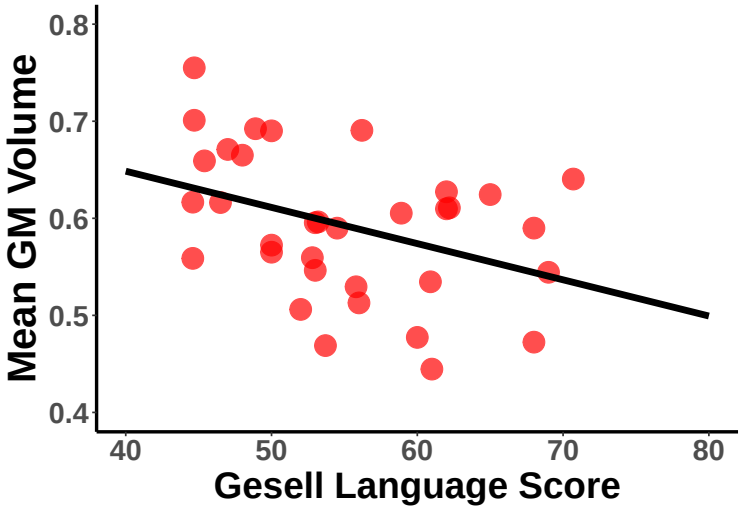



```
## `geom_smooth()` using formula = 'y ~ x'
## `geom_smooth()` using formula = 'y ~ x'
```




```
pplot(datafile_ROIsignal[datafile_ROIsignal$subgroup=="ASDmoderate",],
      "PCC", xlabel, c(0.4, 0.8), c(40,80))
```

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



Correlations between language and other domains 2024-09-23

```
# load data
datafile <- read.csv("data/datafile_clean.csv", header = T)

tt_all <- as.data.frame(as.matrix(0,3*5, 4))

t <- 0
for (gg in c("TD", "ASDmoderate", "ASDsevere"))
{
    tmp <-  subset(datafile, subgroup == gg)

    for (tt  in c("Gesell_Total",
              "Gesell_Adpt","Gesell_GrMot","Gesell_FineMot","Gesell_Social"))
        {
    t <- t + 1
    tt_tmp <- cor.test(tmp$Gesell_Lang, tmp[,tt])
    tt_all[t, 1:2] <- c(gg, tt)
    tt_all[t, 3:4] <- c(round(tt_tmp$estimate,3), round(tt_tmp$p.value,3))
}
}

#View(tt_all)

colnames(tt_all) <- c("Group", "Gesell domain","P value","r value")

writexl::write_xlsx(tt_all, "results/correlation_language.xlsx")
```