

# R Markdown file: Internalizing and externalizing behaviors in school-aged children are related to state anxiety during magnetic resonance imaging

## Contents

1. Data preparation	1
2. Demographic characteristics	2
3. Aim 1: The relationship between internalizing and externalizing behaviors and state anxiety (linear mixed model)	6
4. Aim 2: The relationship between internalizing and externalizing behaviors and MRI participation (logistic regression)	8
5. Aim 3: The relationship between internalizing and externalizing behaviors, state anxiety, and image quality (linear regression)	10
6. Multiple testing correction (FDR-Benjamini Hochberg)	11
7. Print tables including FDR correction	12
7. Correlogram analyzed variables (Figure S3)	14

This Markdown file describes the analyses performed for Eijlers et al. *Internalizing and externalizing behaviors in school-aged children are related to state anxiety during magnetic resonance imaging*.

## 1. Data preparation

Load required libraries

```
library(kableExtra) #used to make tables
library(nlme) #used for mixed models
library(foreign) #used to read in datasets in .sav format
library(GGally) #used to create correlogram
```

```
## Warning: replacing previous import 'vctrs::data_frame' by 'tibble::data_frame'
## when loading 'dplyr'
```

```
library(tidyr) #general use preparation data
library(dplyr) #general use preparation data
```

Load datasets

```
#Set working directory
setwd("V:/medewerkers/074008 Blok, E/Collaborations/Eijlers/3.Data/")

#Read in data
```

```
df_aim1 <- read.spss("df_aim1.sav", to.data.frame = T)
df_aim2 <- read.spss("df_aim2.sav", to.data.frame = T)
df_aim3 <- read.spss("df_aim3.sav", to.data.frame = T)
```

## 2. Demographic characteristics

Create Table 1

```
#Create dataframe that contains only those variables that we use for Table 1.
df_demographics <- merge(df_aim2[c("IDC", "mri_done", "agechild_cbcl_years",
                                   "agechild_evaluation", "Gender", "education_cat",
                                   "income_cat", "sum_int_5", "sum_ext_5")],
                        df_aim1[,c("IDC", "Index1", "child", "parent", "researcher")],
                        by = "IDC")

#Rename the time variable Index1 to understandable names
df_demographics$Index1 <- as.factor(ifelse(df_demographics$Index1 == 1, "before_mock",
                                           ifelse(df_demographics$Index1 == 2, "during_mock",
                                                  "during_mri")))

#Pivot to wide format, as we do not want multiple rows for 1 participant here
df_demographics_wide <- pivot_wider(df_demographics, names_from = c("Index1"),
                                   values_from = c("child", "parent", "researcher"))

#Transform VAS scores to numeric values
#As the levels are specified in the correct order, we can just rely on as.numeric here
df_demographics_wide[,10:18] <- lapply(df_demographics_wide[,10:18], as.numeric)
#However, as.numeric converts to the range of 1-6, while we want to use 0-5,
#so we subtract 1 from all these values
df_demographics_wide[,10:18] <- df_demographics_wide[,10:18] - 1

#Create two separate dataframes,
#df_mri includes all participants who underwent MRI scanning,
#df_nomri includes all participants that did not undergo MRI scanning
df_mri <- subset(df_demographics_wide, mri_done == 1)
df_nomri <- subset(df_demographics_wide, mri_done == 0)

#Specify demographic variables
demvars <- colnames(df_demographics_wide[3:18])

#Create empty matrix to fill
dem_values <- matrix(ncol = 4, nrow = 26)

#Keep track off which row we are at in the loop
rowcount_dfs <- 1

#Loop over all dataframes
dem_dfs <- c("df_mri", "df_nomri")

#Specify distributions of continuous variables
distributions <- matrix(ncol = 1, nrow = 13,
                       dimnames = list(c("agechild_cbcl_years", "agechild_evaluation",
                                           "sum_int_5", "sum_ext_5",
```

```

        "child_before_mock", "child_during_mock",
        "child_during_mri",
        "parent_before_mock", "parent_during_mock",
        "parent_during_mri",
        "researcher_before_mock", "researcher_during_mock",
        "researcher_during_mri",
        c("distribution")))
distributions[,1] <- c("normal", "normal", "skewed", "skewed", "skewed",
                     "skewed", "skewed", "skewed", "skewed", "skewed",
                     "skewed", "skewed", "skewed")

#Create table with demographics
for(x in dem_dfs){
  rowcount <- 0
  for(y in demvars){
    rowcount <- rowcount + 1
    #Calculate means & sds for continuous variables
    if(is.numeric(get(x)[[y]])){
      #Store the total n for continuous measures
      if(is.na(mean(get(x)[[y]]))){
        dem_values[rowcount, rowcount_dfs] <-
          nrow(get(x)) - summary(get(x)[[y]])[[7]]
      } else {
        dem_values[rowcount, rowcount_dfs] <- nrow(get(x))
      }
      if(distributions[y,] == "normal"){
        dem_values[rowcount, rowcount_dfs+1] <-
          paste0(round(mean(get(x)[[y]], na.rm = T),2),
                 " (", round(sd(get(x)[[y]], na.rm = T),2), ")")
      }
      if(distributions[y,] == "skewed"){
        dem_values[rowcount, rowcount_dfs+1] <-
          paste0(round(summary(get(x)[[y]])[[3]],2),
                 " (", round(summary(get(x)[[y]])[[2]],2),
                 "-", round(summary(get(x)[[y]])[[5]],2), ")")
      }
    }
  }
  if(!is.numeric(get(x)[[y]])){
    #Calculate n and % for categorical variables
    for(z in 1:length(summary(get(x)[,y]))){
      rowcount <- rowcount + 1
      dem_values[rowcount, rowcount_dfs] <- summary(get(x)[[y]])[[z]]
      dem_values[rowcount, rowcount_dfs+1] <-
        paste0(round(prop.table(summary(get(x)[[y]]))[[z]]*100,2), "%")
    }
  }
}
rowcount_dfs <- rowcount_dfs + 2
}

#Make colnames and rownames comprehensible again
rownames(dem_values) <- c("Age CBCL (M, SD)", "Age MRI (M, SD)", "Child biological sex (%)",

```

```

levels(df_demographics_wide$Gender), "Maternal education (%)",
levels(df_demographics_wide$education_cat), "NA",
"Household income (%)",
levels(df_demographics_wide$income_cat), "NA",
"Internalizing symptoms (Median, IQR)",
"Externalizing symptoms (Median, IQR)",
"VAS child before mock (Median, IQR)",
"VAS child during mock (Median, IQR)",
"VAS child during MRI (Median, IQR)",
"VAS parent before mock (Median, IQR)",
"VAS parent during mock (Median, IQR)",
"VAS parent during MRI (Median, IQR)",
"VAS researcher before mock (Median, IQR)",
"VAS researcher during mock (Median, IQR)",
"VAS researcher during MRI (Median, IQR)"
colnames(dem_values) <- c("MRI performed n", "MRI performed (M,SD)/%/(Median,IQR)",
"MRI not performed n", "MRI not performed (M,SD)/%/(Median,IQR)")

#Print table
dem_values %>%
  kbl(caption="Sample characteristics",
      col.names = c("MRI performed", "", "MRI not performed", ""),
      align="r") %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 1: Sample characteristics

	MRI performed		MRI not performed	
Age CBCL (M, SD)	981	6.06 (0.45)	159	5.98 (0.41)
Age MRI (M, SD)	1070	7.91 (1)	175	7.62 (0.92)
Child biological sex (%)				
boy	572	53.46%	90	51.43%
girl	498	46.54%	85	48.57%
Maternal education (%)				
High	479	44.77%	64	36.57%
Low	71	6.64%	19	10.86%
Middle	478	44.67%	87	49.71%
NA	42	3.93%	5	2.86%
Household income (%)				
High	615	57.48%	80	45.71%
Low	144	13.46%	19	10.86%
Middle	169	15.79%	43	24.57%
NA	142	13.27%	33	18.86%
Internalizing symptoms (Median, IQR)	972	6 (2-11)	157	9 (4-14)
Externalizing symptoms (Median, IQR)	976	8 (3-15)	157	11 (5-18)
VAS child before mock (Median, IQR)	1059	1 (0-2)	169	2 (0-3)
VAS child during mock (Median, IQR)	1069	0 (0-1)	144	3 (1-4)
VAS child during MRI (Median, IQR)	1064	0 (0-1)	66	4 (2-5)
VAS parent before mock (Median, IQR)	1059	1 (0-2)	167	2 (1-3)
VAS parent during mock (Median, IQR)	1066	1 (0-2)	143	3 (2-4)
VAS parent during MRI (Median, IQR)	1064	1 (0-1)	66	4 (3-5)
VAS researcher before mock (Median, IQR)	1062	1 (1-2)	167	2 (1-3)
VAS researcher during mock (Median, IQR)	1068	1 (0-2)	144	3 (2-4)
VAS researcher during MRI (Median, IQR)	1062	1 (0-1)	66	4.5 (3-5)

Comparison between demographic characteristics for those that participated and those that did not participate in MRI scanning

```
#Run t-tests/chi-square tests to compare subgroups
```

```
#T-tests
```

```
t.test(df_demographics_wide$agechild_cbcl_years~df_demographics_wide$mri_done)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: df_demographics_wide$agechild_cbcl_years by df_demographics_wide$mri_done
```

```
## t = -2.3591, df = 225.85, p-value = 0.01917
```

```
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
```

```
## -0.15340599 -0.01376898
```

```
## sample estimates:
```

```
## mean in group 0 mean in group 1
```

```
## 5.981016 6.064603
```

```
t.test(df_demographics_wide$agechild_evaluation~df_demographics_wide$mri_done)
```

```
##
```

```
## Welch Two Sample t-test
```

```
##
```

```
## data: df_demographics_wide$agechild_evaluation by df_demographics_wide$mri_done
## t = -3.758, df = 245.79, p-value = 0.000214
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.4355472 -0.1359898
## sample estimates:
## mean in group 0 mean in group 1
## 7.621175 7.906944

#Chi-square tests
chisq.test(df_demographics_wide$Gender, df_demographics_wide$mri_done)

##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data: df_demographics_wide$Gender and df_demographics_wide$mri_done
## X-squared = 0.17394, df = 1, p-value = 0.6766

chisq.test(df_demographics_wide$education_cat, df_demographics_wide$mri_done)

##
## Pearson's Chi-squared test
##
## data: df_demographics_wide$education_cat and df_demographics_wide$mri_done
## X-squared = 6.7942, df = 2, p-value = 0.03347

chisq.test(df_demographics_wide$income_cat, df_demographics_wide$mri_done)

##
## Pearson's Chi-squared test
##
## data: df_demographics_wide$income_cat and df_demographics_wide$mri_done
## X-squared = 11.296, df = 2, p-value = 0.003524
```

### 3. Aim 1: The relationship between internalizing and externalizing behaviors and state anxiety (linear mixed model)

```
#Transform VAS scores to numeric values
#As the levels are specified in the correct order, we can just rely on as.numeric here
df_aim1[,c("child", "parent", "researcher")] <-
  lapply(df_aim1[,c("child", "parent", "researcher")], as.numeric)
#However, as.numeric converts to the range of 1-6, while we want to use 0-5,
#so we subtract 1 from all these values
df_aim1[,c("child", "parent", "researcher")] <-
  df_aim1[,c("child", "parent", "researcher")] - 1

#Transform Index1 variable which indicates time point into categorical variable
df_aim1$Index1 <- as.factor(ifelse(df_aim1$Index1 == 1, "before_mock",
  ifelse(df_aim1$Index1 == 2, "during_mock", "during_mri")))

#Run mixed model, collapsing all VAS timepoints
#Model 1
#Internalizing
#Child
coefs_int_child_m1 <- summary(lme(child~sum_int_5+agechild_evaluation+Gender,
```

```

random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Parent
coefs_int_parent_m1 <- summary(lme(parent~sum_int_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Researcher
coefs_int_researcher_m1 <- summary(lme(researcher~sum_int_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Externalizing
#Child
coefs_ext_child_m1 <- summary(lme(child~sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Parent
coefs_ext_parent_m1 <- summary(lme(parent~sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Researcher
coefs_ext_researcher_m1 <- summary(lme(researcher~sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2,c(1,2,3,5)]

#Model 2
#Child
coefs_child_m2 <- summary(lme(child~sum_int_5+sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2:3,c(1,2,3,5)]

#Parent
coefs_parent_m2 <- summary(lme(parent~sum_int_5+sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2:3,c(1,2,3,5)]

#Researcher
coefs_researcher_m2 <- summary(lme(researcher~sum_int_5+sum_ext_5+agechild_evaluation+Gender,
random = (~1|IDC), data = df_aim1,
na.action = na.omit))[[20]][2:3,c(1,2,3,5)]

#Bind together in dataframe
results_aim1_temp <- rbind(coefs_int_child_m1, coefs_child_m2[1,], coefs_ext_child_m1,
coefs_child_m2[2,], coefs_int_parent_m1, coefs_parent_m2[1,],
coefs_ext_parent_m1, coefs_parent_m2[2,], coefs_int_researcher_m1,
coefs_researcher_m2[1,], coefs_ext_researcher_m1,
coefs_researcher_m2[2,]
)

#Rearrange into better looking table
#Create empty dataframe
results_aim1 <- matrix(ncol = 6, nrow = 12)
#Denotes reporter
results_aim1[,1] <- c("Child", rep("",3), "Parent", rep("",3), "Researcher", rep("",3))
results_aim1[,2] <- rep(c("Internalizing","", "Externalizing", ""),3) #Denotes CBCL scale

```

```

results_aim1[,3] <- rep(c(1, 2),6) #Denotes which model
results_aim1[,4] <- round(results_aim1_temp[,1],3) #B's
results_aim1[,5] <- round(results_aim1_temp[,2],3) #SE's
results_aim1[,6] <- round(results_aim1_temp[,4],4) #p-values

#Print table
results_aim1 %>%
  kbl(caption="Association between MRI related state anxiety (measured by VAS),
  internalizing and externalizing behavior and VAS scores.
  Model 1: corrected for age and sex
  Model 2: internalizing and externalizing behavior in one model,
  corrected for age and sex.",
  col.names = c("Reporter VAS", "Behavior", "Model", "B", "S.E.", "p-value"),
  align="r") %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 2: Association between MRI related state anxiety (measured by VAS), internalizing and externalizing behavior and VAS scores. Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior in one model, corrected for age and sex.

Reporter VAS	Behavior	Model	B	S.E.	p-value
Child	Internalizing	1	0.01	0.004	0.0061
		2	0.009	0.005	0.0973
	Externalizing	1	0.008	0.003	0.0262
		2	0.002	0.005	0.752
Parent	Internalizing	1	0.02	0.004	0
		2	0.02	0.005	2e-04
	Externalizing	1	0.014	0.003	1e-04
		2	0	0.005	0.9998
Researcher	Internalizing	1	0.017	0.003	0
		2	0.018	0.005	6e-04
	Externalizing	1	0.011	0.003	7e-04
		2	-0.001	0.005	0.836

#### 4. Aim 2: The relationship between internalizing and externalizing behaviors and MRI participation (logistic regression)

```

#Model 1
#Internalizing
aim2_int_model1 <- glm(mri_done_inversed~sum_int_5+agechild_evaluation+Gender,
  data = df_aim2, family = binomial)

#Externalizing
aim2_ext_model1 <- glm(mri_done_inversed~sum_ext_5+agechild_evaluation+Gender,
  data = df_aim2, family = binomial)

#Model 2 (includes internalizing and externalizing, thus only presented once here)
aim2_model2 <- glm(mri_done_inversed~sum_int_5+sum_ext_5+agechild_evaluation+Gender,
  data = df_aim2, family = binomial)

#Extract results
coefs_int_m1 <- summary(aim2_int_model1)[13][[1]][2,c(1,2,4)]
coefs_ext_m1 <- summary(aim2_ext_model1)[13][[1]][2,c(1,2,4)]

```



```

coefs_m2 <- summary(aim2_model2)[13][[1]][2:3,c(1,2,4)]

#Extract Odds Ratio's
OR_CI_int_m1 <- exp(cbind(coef(aim2_int_model1), confint(aim2_int_model1)))[2,]

## Waiting for profiling to be done...

OR_CI_ext_m1 <- exp(cbind(coef(aim2_ext_model1), confint(aim2_ext_model1)))[2,]

## Waiting for profiling to be done...

OR_CI_m2 <- exp(cbind(coef(aim2_model2), confint(aim2_model2)))[2:3,]

## Waiting for profiling to be done...

#Bind together in dataframe
results_aim2_temp <- cbind(rbind(coefs_int_m1, coefs_m2[1,], coefs_ext_m1, coefs_m2[2,]),
                           rbind(OR_CI_int_m1, OR_CI_m2[1,],OR_CI_ext_m1, OR_CI_m2[2,]))

#Rearrange into better looking table
#Create empty dataframe
results_aim2 <- matrix(ncol = 5, nrow = 4)
results_aim2[,1] <- c("Internalizing", "", "Externalizing", "") #Denotes which CBCL scale
results_aim2[,2] <- c(1,2,1,2) #Denotes which model
results_aim2[,3] <- round(results_aim2_temp[,4],2) #Odds ratio's
results_aim2[,4] <- paste0(round(results_aim2_temp[,5],2), "-",
                           round(results_aim2_temp[,6],2)) #95% CI
results_aim2[,5] <- round(results_aim2_temp[,3],3) #p-values

#Print table
results_aim2 %>%
  kbl(caption="Association between internalizing and externalizing behavior,
and MRI participation (reference group: participated in MRI scanning).
Model 1: corrected for age and sex
Model 2: internalizing and externalizing behavior in one model,
corrected for age and sex.",
      col.names = c("Behavior", "Model", "Odds Ratio MRI participation",
                    "95% Confidence Interval", "p-value"),
      align="r") %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 3: Association between internalizing and externalizing behavior, and MRI participation (reference group: participated in MRI scanning). Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior in one model, corrected for age and sex.

Behavior	Model	Odds Ratio MRI participation	95% Confidence Interval	p-value
Internalizing	1	1.03	1.01-1.06	0.001
	2	1.03	1-1.06	0.088
Externalizing	1	1.03	1.01-1.05	0.004
	2	1.01	0.98-1.04	0.49

## 5. Aim 3: The relationship between internalizing and externalizing behaviors, state anxiety, and image quality (linear regression)

```
#Transform VAS scores to numeric values
df_aim3$Faces_scared_during_cont <-
  ifelse(df_aim3$Faces_scared_during == "face 0", 0,
        ifelse(df_aim3$Faces_scared_during == "face 1", 1,
              ifelse(df_aim3$Faces_scared_during == "face 2", 2,
                    ifelse(df_aim3$Faces_scared_during == "face 3", 3,
                          ifelse(df_aim3$Faces_scared_during == "face 4", 4, 5))))))

#Pivot from long to wide format to analyze VAS reporters separately
df_aim3_wide <-
  pivot_wider(df_aim3, names_from = c("Index1"),
              values_from = c("Faces_scared_during_cont", "Faces_scared_during"))

#Image quality & internalizing problems
coefs_int_model1 <- summary(lm(qal~sum_int_5+agechild_evaluation+Gender,
                              data = df_aim3_wide))[[4]][2, c(1,2,4)]

#Image quality & externalizing problems
coefs_ext_model1 <- summary(lm(qal~sum_ext_5+agechild_evaluation+Gender,
                              data = df_aim3_wide))[[4]][2, c(1,2,4)]

#Image quality & VAS
coefs_VAS_child_model1 <-
  summary(lm(qal~Faces_scared_during_cont_child+agechild_evaluation+Gender,
            data = df_aim3_wide))[[4]][2, c(1,2,4)]
coefs_VAS_parent_model1 <-
  summary(lm(qal~Faces_scared_during_cont_parent+agechild_evaluation+Gender,
            data = df_aim3_wide))[[4]][2, c(1,2,4)]
coefs_VAS_researcher_model1 <-
  summary(lm(qal~Faces_scared_during_cont_researcher+agechild_evaluation+Gender,
            data = df_aim3_wide))[[4]][2, c(1,2,4)]

#All in one model
coefs_model2 <-
  summary(lm(qal~sum_int_5+sum_ext_5+Faces_scared_during_cont_child+
            Faces_scared_during_cont_parent+Faces_scared_during_cont_researcher+
            agechild_evaluation+Gender, data = df_aim3_wide))[[4]][2:6, c(1,2,4)]

#Bind together in dataframe
results_aim3_temp <- rbind(coefs_int_model1, coefs_model2[1,], coefs_ext_model1,
                          coefs_model2[2,], coefs_VAS_child_model1, coefs_model2[3,],
                          coefs_VAS_parent_model1, coefs_model2[4,],
                          coefs_VAS_researcher_model1, coefs_model2[5,])

#Rearrange into better looking table
#Create empty dataframe
results_aim3 <- matrix(ncol = 5, nrow = 10)
#Denotes which CBCL/VAS scale
```

```

results_aim3[,1] <- c("Internalizing", "", "Externalizing", "",
                     "VAS child", "", "VAS parent", "", "VAS researcher", "")
results_aim3[,2] <- rep(1:2,5) #Denotes which model
results_aim3[,3] <- round(results_aim3_temp[,1],2) #B's
results_aim3[,4] <- round(results_aim3_temp[,2],2) #SE's
results_aim3[,5] <- round(results_aim3_temp[,3],3) #p-values

#Print table
results_aim3 %>%
  kbl(caption="Association between MRI quality, internalizing and
externalizing behavior and VAS scores.
Model 1: corrected for age and sex
Model 2: internalizing and externalizing behavior and VAS scores
in one model, corrected for age and sex.",
      col.names = c("Behavior", "Model", "B", "S.E.", "p-value"),
      align="r"
    ) %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 4: Association between MRI quality, internalizing and externalizing behavior and VAS scores. Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior and VAS scores in one model, corrected for age and sex.

Behavior	Model	B	S.E.	p-value
Internalizing	1	-0.19	0.88	0.831
	2	1.88	1.29	0.144
Externalizing	1	-1.1	0.82	0.177
	2	-2.25	1.18	0.056
VAS child	1	-18.4	6.62	0.006
	2	-4.02	8.74	0.646
VAS parent	1	-11.92	6.41	0.063
	2	12.24	8.64	0.157
VAS researcher	1	-31.83	7.39	0
	2	-36.83	10.57	0.001

## 6. Multiple testing correction (FDR-Benjamini Hochberg)

```

#Bind all pvalues together
pvals_all <- c(results_aim1[,6], results_aim2[,5], results_aim3[,5])

#Create empty FDR dataframe
#Nrow corresponds to the number of tests,
#ncol corresponds to columns that we want in our dataframe
 #(space for p-values, FDR thresholds and whether the p-value remains significant)
fdr <- matrix(nrow = length(pvals_all), ncol = 3)

#Calculate threshold for each test
for(x in 1:nrow(fdr)){
  thresh_temp <- x/nrow(fdr)*0.05
  if(x == 1){
    thresholds <- thresh_temp
  }
}

```

```

} else {
  thresholds <- c(thresholds, thresh_temp)
}
}

#Fill FDR dataframe
fdr[,1] <- sort(as.numeric(pvals_all)) #Sort all p-values from low to high
fdr[,2] <- thresholds
fdr[,3] <- ifelse(as.numeric(fdr[,1]) < as.numeric(fdr[,2]), "sig", "nonsig")

#Print which p-values remain significant
fdr

```

```

##      [,1]      [,2]      [,3]
## [1,] "0"      "0.00192307692307692" "sig"
## [2,] "0"      "0.00384615384615385" "sig"
## [3,] "0"      "0.00576923076923077" "sig"
## [4,] "1e-04"  "0.00769230769230769" "sig"
## [5,] "2e-04"  "0.00961538461538462" "sig"
## [6,] "6e-04"  "0.0115384615384615" "sig"
## [7,] "7e-04"  "0.0134615384615385" "sig"
## [8,] "0.001"  "0.0153846153846154" "sig"
## [9,] "0.001"  "0.0173076923076923" "sig"
## [10,] "0.004" "0.0192307692307692" "sig"
## [11,] "0.006" "0.0211538461538462" "sig"
## [12,] "0.0061" "0.0230769230769231" "sig"
## [13,] "0.0262" "0.025" "nonsig"
## [14,] "0.056" "0.0269230769230769" "nonsig"
## [15,] "0.063" "0.0288461538461538" "nonsig"
## [16,] "0.088" "0.0307692307692308" "nonsig"
## [17,] "0.0973" "0.0326923076923077" "nonsig"
## [18,] "0.144" "0.0346153846153846" "nonsig"
## [19,] "0.157" "0.0365384615384615" "nonsig"
## [20,] "0.177" "0.0384615384615385" "nonsig"
## [21,] "0.49"  "0.0403846153846154" "nonsig"
## [22,] "0.646" "0.0423076923076923" "nonsig"
## [23,] "0.752" "0.0442307692307692" "nonsig"
## [24,] "0.831" "0.0461538461538462" "nonsig"
## [25,] "0.836" "0.0480769230769231" "nonsig"
## [26,] "0.9998" "0.05" "nonsig"

```

## 7. Print tables including FDR correction

```

#Add correction to tables
results_aim1_fin <- cbind(results_aim1, ifelse(as.numeric(results_aim1[,6]) <=
  0.0230769230769231, "*", ""))
results_aim2_fin <- cbind(results_aim2, ifelse(as.numeric(results_aim2[,5]) <=
  0.0230769230769231, "*", ""))
results_aim3_fin <- cbind(results_aim3, ifelse(as.numeric(results_aim3[,5]) <=
  0.0230769230769231, "*", ""))

#Print final tables
#Aim 1

```

```

results_aim1_fin %>%
  kbl(caption="Association between MRI related state anxiety (measured by VAS),
    internalizing and externalizing behavior and VAS scores.
    Model 1: corrected for age and sex
    Model 2: internalizing and externalizing behavior in one model,
    corrected for age and sex.
    *Indicates statistical significance after multiple testing correction (q-value 0.05)",
    col.names = c("Reporter VAS", "Behavior", "Model", "B", "S.E.", "p-value", ""),
    align="r") %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 5: Association between MRI related state anxiety (measured by VAS), internalizing and externalizing behavior and VAS scores. Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior in one model, corrected for age and sex. \*Indicates statistical significance after multiple testing correction (q-value 0.05)

Reporter VAS	Behavior	Model	B	S.E.	p-value	
Child	Internalizing	1	0.01	0.004	0.0061	*
		2	0.009	0.005	0.0973	
	Externalizing	1	0.008	0.003	0.0262	
		2	0.002	0.005	0.752	
Parent	Internalizing	1	0.02	0.004	0	*
		2	0.02	0.005	2e-04	*
	Externalizing	1	0.014	0.003	1e-04	*
		2	0	0.005	0.9998	
Researcher	Internalizing	1	0.017	0.003	0	*
		2	0.018	0.005	6e-04	*
	Externalizing	1	0.011	0.003	7e-04	*
		2	-0.001	0.005	0.836	

```

#Aim 2
results_aim2_fin %>%
  kbl(caption="Association between internalizing and externalizing behavior,
    and MRI participation (reference group: participated in MRI scanning).
    Model 1: corrected for age and sex
    Model 2: internalizing and externalizing behavior in one model,
    corrected for age and sex.
    *Indicates statistical significance after multiple testing correction (q-value 0.05)",
    col.names = c("Behavior", "Model", "Odds Ratio MRI participation",
      "95% Confidence Interval", "p-value", ""),
    align="r") %>%
  kable_classic(full_width = F, html_font = "helvetica")

```

Table 6: Association between internalizing and externalizing behavior, and MRI participation (reference group: participated in MRI scanning). Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior in one model, corrected for age and sex. \*Indicates statistical significance after multiple testing correction (q-value 0.05)

Behavior	Model	Odds Ratio MRI participation	95% Confidence Interval	p-value	
Internalizing	1	1.03	1.01-1.06	0.001	*
	2	1.03	1-1.06	0.088	
Externalizing	1	1.03	1.01-1.05	0.004	*
	2	1.01	0.98-1.04	0.49	

```
#Aim3
results_aim3_fin %>%
  kbl(caption="Association between MRI quality, internalizing and
externalizing behavior and VAS scores.
  Model 1: corrected for age and sex
  Model 2: internalizing and externalizing behavior and VAS scores
in one model, corrected for age and sex.
  *Indicates statistical significance after multiple testing correction (q-value 0.05)",
  col.names = c("Behavior", "Model", "B", "S.E.", "p-value", ""),
  align="r"
) %>%
  kable_classic(full_width = F, html_font = "helvetica")
```

Table 7: Association between MRI quality, internalizing and externalizing behavior and VAS scores. Model 1: corrected for age and sex Model 2: internalizing and externalizing behavior and VAS scores in one model, corrected for age and sex. \*Indicates statistical significance after multiple testing correction (q-value 0.05)

Behavior	Model	B	S.E.	p-value	
Internalizing	1	-0.19	0.88	0.831	
	2	1.88	1.29	0.144	
Externalizing	1	-1.1	0.82	0.177	
	2	-2.25	1.18	0.056	
VAS child	1	-18.4	6.62	0.006	*
	2	-4.02	8.74	0.646	
VAS parent	1	-11.92	6.41	0.063	
	2	12.24	8.64	0.157	
VAS researcher	1	-31.83	7.39	0	*
	2	-36.83	10.57	0.001	*

## 7. Correlogram analyzed variables (Figure S3)

```
#Create dataframe with just the variables we need
df_correlogram <-
  select(df_demographics, c("IDC", "Gender", "agechild_evaluation",
    "agechild_cbcl_years", "sum_int_5", "sum_ext_5",
    "Index1", "child", "parent", "researcher"))

#Transform VAS scores to numeric values
#As the levels are specified in the correct order, we can just rely on as.numeric here
df_correlogram[,c("child", "parent", "researcher")] <-
```

```

  lapply(df_correlogram[,c("child", "parent", "researcher")], as.numeric)
#However, as.numeric converts to the range of 1-6, while we want to use 0-5,
#so we subtract 1 from all these values
df_correlogram[,c("child", "parent", "researcher")] <-
  df_correlogram[,c("child", "parent", "researcher")] - 1

#Pivot from long to wide format
df_correlogram_wide <-
  pivot_wider(df_correlogram, names_from = c("Index1"),
    values_from = c("child", "parent", "researcher"))

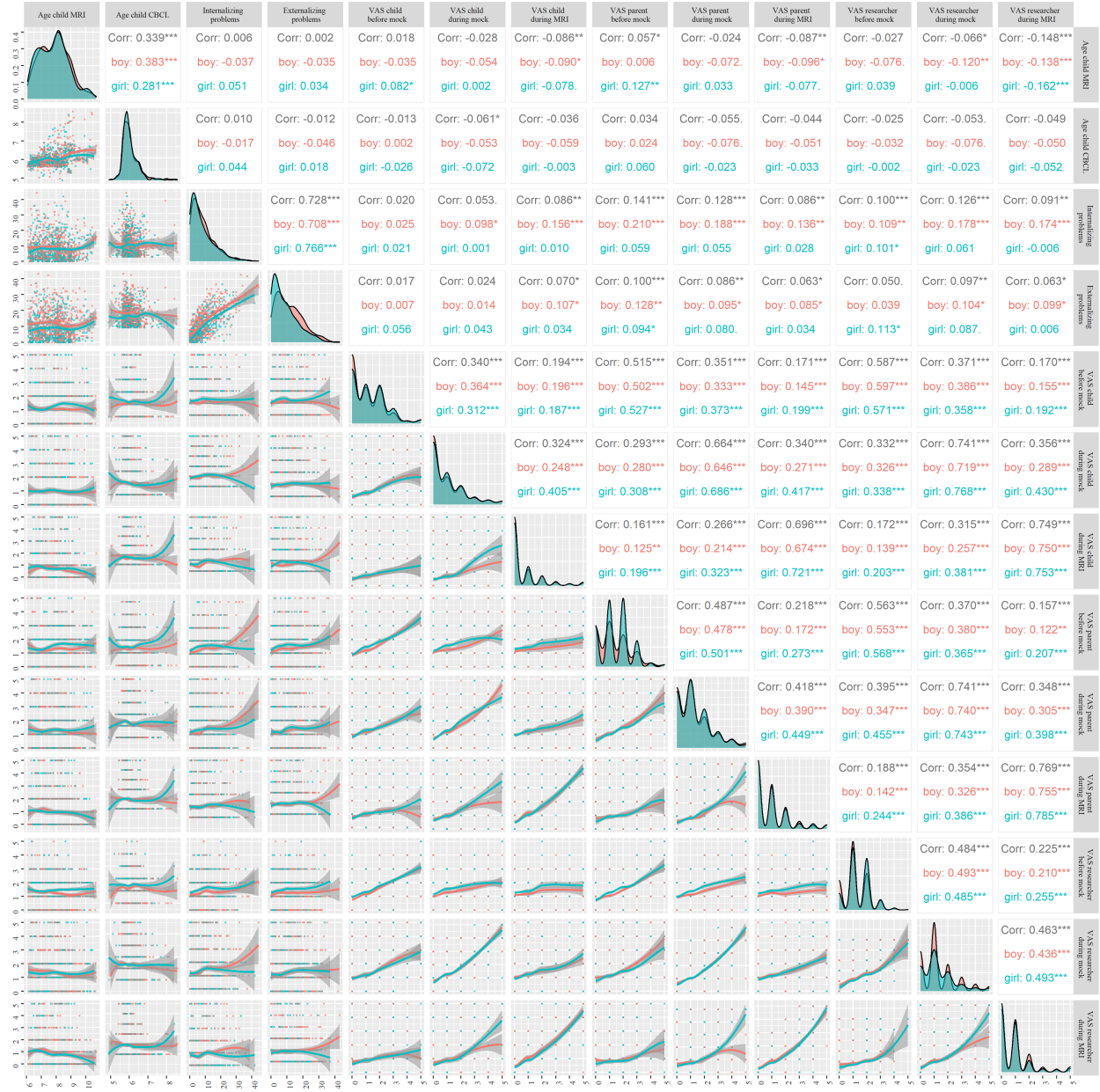
#Make correlogram
#Define colnames in a way that is interpretable
colnames(df_correlogram_wide) <-
  c("IDC", "Biological sex", "Age child MRI", "Age child CBCL",
    "Internalizing \n problems", "Externalizing \n problems",
    "VAS child \n before mock", "VAS child \n during mock",
    "VAS child \n during MRI", "VAS parent \n before mock",
    "VAS parent \n during mock", "VAS parent \n during MRI",
    "VAS researcher \n before mock", "VAS researcher \n during mock",
    "VAS researcher \n during MRI")

#Create correlogram
my_fn <- function(data, mapping, pts=list(), smt=list(), ...){
  ggplot(data = data, mapping = mapping, ...) +
    do.call(geom_point, pts) +
    do.call(geom_smooth, smt)
}

correlogram <-
  ggpairs(df_correlogram_wide[,3:15],
    mapping = aes(color = df_correlogram_wide$`Biological sex`, alpha = 0.3),
    axisLabels = "show",
    lower = list(continuous = wrap(my_fn,
      pts=list(size=.5),
      smt=list(method="loess", se=T, size=1)))) +
  theme(axis.text = element_text(color = "grey20", size = 8, angle = 90,
    hjust = .5, vjust = .5, face = "plain"),
    strip.text = element_text(size = 8),
    text = element_text(family = "serif"))

#Print correlogram
correlogram

```



```
#Set working directory
setwd("V:/medewerkers/074008 Blok, E/Collaborations/Eijlers/4.Results/")
ggsave("FigureS3.png", plot = correlogram, width = 4500, height = 4500, units = "px")
```