

Appendix A: Simulation Studies for Model Selection

Simulate data: The “true model”

Different options for the factor loadings (FL) of the family effect

- Option 1: all observed scores are of equal importance for the family effect (FE) => constrain every loading to 1, sum will be equal to 6 (identical to simulation study with traditional SRM)

```
o1 <- "FE = 1*MF + 1*MO + 1*MY + 1*FO + 1*FY + 1*YO"
```

- Option 2: some FL's are more important than others, but sum of all is equal to 6 (lambda approach)

```
o2 <- "FE = 1.4*MF + 0.8*MO + 0.8*MY + 0.8*FO + 0.8*FY + 1.4*YO"
```

- Option 3: fix 1 FL to 1 and for the other FL's we will use the lambda approach (based on model 3a; lambda approach) => sum is equal to 6

```
o3 <- "FE = 1*MF + 0.9*MO + 1.2*MY + 1.1*FO + 1.2*FY + 0.6*YO"
```

- Option 4: lambda for each person (based on model 3b) => sum is equal to 6

```
o4 <- "FE = 0.7*MF + 1.1*MO + 1.2*MY + 1.2*FO + 1.1*FY + 0.7*YO"
```

```
# Factor loadings for FE:
# option 1: all observed scores are of equal importance for the family effect (FE)
# => constrain every loading to 1, sum will be equal to 6 (identical to simulation
# study with traditional SRM)
o1 <- "FE =~ 1*MF + 1*MO + 1*MY + 1*FO + 1*FY + 1*YO"
# option 2: some FL's are more important than others, but sum of all is equal to
# six (lambda approach)
o2 <- "FE =~ 1.4*MF + 0.8*MO + 0.8*MY + 0.8*FO + 0.8*FY + 1.4*YO"
# option 3: fix 1 FL to 1 and for the other FL's we will use the lambda approach
# (based on model 3a; lambda approach) => sum is equal to 6
o3 <- "FE =~ 1*MF + 0.9*MO + 1.2*MY + 1.1*FO + 1.2*FY + 0.6*YO"
# option 4: lambda for each person (based on model 3b) => sum is equal to 6
o4 <- "FE =~ 0.7*MF + 1.1*MO + 1.2*MY + 1.2*FO + 1.1*FY + 0.7*YO"
FE <- c(o1,o2,o3,o4)

for (i in 1:4){ # Factor loadings for FE
  true_t <- paste0(FE[i],
    ,
    I.M =~ 1*MF + 1*MO + 1*MY
    I.F =~ 1*MF + 1*FO + 1*FY
    I.O =~ 1*MO + 1*FO + 1*YO
    I.Y =~ 1*MY + 1*FY + 1*YO
    D.MF =~ 1*MF
    D.MO =~ 1*MO
    D.MY =~ 1*MY
    D.FO =~ 1*FO
    D.FY =~ 1*FY
    D.OY =~ 1*YO

    # Variances
    FE ~~ 2*FE
    I.M ~~ 1*I.M
```

```

I.F ~~ 1*I.F
I.O ~~ 1*I.O
I.Y ~~ 1*I.Y
D.MF ~~ 0.5*D.MF
D.MO ~~ 0.5*D.MO
D.MY ~~ 0.5*D.MY
D.FO ~~ 0.5*D.FO
D.FY ~~ 0.5*D.FY
D.OY ~~ 0.5*D.OY

# Intercepts
FE ~ 4*1
I.M ~ -0.8*1
I.F ~ -0.8*1
I.O ~ 0.8*1
I.Y ~ 0.8*1
D.MF ~ 1*1
D.MO ~ -0.5*1
D.MY ~ -0.5*1
D.FO ~ -0.5*1
D.FY ~ -0.5*1
D.OY ~ 1*1
')
assign(paste0("PDSRM_th_FE", i), true_t)
}
truemodel <- list(option1 = PDSRM_th_FE1, option2 = PDSRM_th_FE2, option3 = PDSRM_th_FE3,
                  option4 = PDSRM_th_FE4)

```

Fit different types of models with the simulated data

Method 1: PD SRM as an adaptation of Kenny's SRM (model 1)

In this model, the factor loadings are all fixed to 1. This is similar to the traditional SRM.

Note: For each model, the model syntax is printed. The data simulating loop is not printed for every model because this remains almost identical (only the name of the model changes).

```
SRM_pd <- '  
# Latent variables  
FE =~ 1*MF + 1*MO + 1*MY + 1*FO + 1*FY + 1*YO  
I.M =~ 1*MF + 1*MO + 1*MY  
I.F =~ 1*MF + 1*FO + 1*FY  
I.O =~ 1*MO + 1*FO + 1*YO  
I.Y =~ 1*MY + 1*FY + 1*YO  
D.MF =~ 1*MF  
D.MO =~ 1*MO  
D.MY =~ 1*MY  
D.FO =~ 1*FO  
D.FY =~ 1*FY  
D.OY =~ 1*YO  
  
# Intragen. covariances  
#   I.M ~~ I.F  
#   I.O ~~ I.Y  
  
# Variances  
FE ~~ VAR.FE*FE  
I.M ~~ VAR.I.M*I.M  
I.F ~~ VAR.I.F*I.F  
I.O ~~ VAR.I.O*I.O  
I.Y ~~ VAR.I.Y*I.Y  
D.MF ~~ VAR.D.MF*D.MF  
D.MO ~~ VAR.D.MO*D.MO  
D.MY ~~ VAR.D.MY*D.MY  
D.FO ~~ VAR.D.FO*D.FO  
D.FY ~~ VAR.D.FY*D.FY  
D.OY ~~ VAR.D.OY*D.OY  
  
# Intercepts  
FE ~ mean.FE*1  
I.M ~ mean.I.M*1  
I.F ~ mean.I.F*1  
I.O ~ mean.I.O*1  
I.Y ~ mean.I.Y*1  
D.MF ~ mean.D.MF*1  
D.MO ~ mean.D.MO*1  
D.MY ~ mean.D.MY*1  
D.FO ~ mean.D.FO*1  
D.FY ~ mean.D.FY*1  
D.OY ~ mean.D.OY*1  
  
# Constraints
```

```

mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0
mean.D.MF + mean.D.MO + mean.D.MY == 0
mean.D.MF + mean.D.FO + mean.D.FY == 0
mean.D.MY + mean.D.FY + mean.D.OY == 0
mean.D.MO + mean.D.FO + mean.D.OY == 0
'

# Simulations
#####
setwd("/user/home/gent/vsc408/vsc40825/modelselectie/model1/resultaten")
N <- 500 # number of simulations
nobs <- c(50,75,100,125,150,200,500,1000) # different sample sizes
a <- 1
for( k in 1:4){ # which option for FE
  for (j in 1:length(nobs)){
    cfi <- c()
    tli <- c()
    rmsea <- c()
    chi2 <- c()
    srmr <- c()
    neg <- c()
    for(i in 1:N) {
      tryCatch({
        sim.data <- simulateData(truemodel[[k]], seed = a, sample.nobs=nobs[j])
        fit <- lavaan(SRM_pd, data=sim.data)
        cfi[i] <- fitMeasures(fit, "cfi")
        rmsea[i] <- fitMeasures(fit, "rmsea")
        chi2[i] <- fitMeasures(fit, "pvalue")
        tli[i] <- fitMeasures(fit, "tli")
        srmr[i] <- fitMeasures(fit, "srmr")
        name <- paste('nobs', nobs[j], sep='')
        fitvalues <- list(TLI = tli, CFI = cfi, RMSEA = rmsea, CHI2 = chi2, SRMR = srmr)
        neg[i] <- length(which(parameterEstimates(fit)[c(25:35),"est"] < 0))
        a <- a + 1
      }, error=function(e){cat("ERROR i=",i, "j=",j,"k=",k , ":",conditionMessage(e),
                                "\n")})
    }
    if (!is.null(cfi)){
      # only write table when there is at least 1 model that converged
      write.table(fitvalues, paste0("dataT0",k,"_m1_fit_nobs",nobs[j], ".txt"))
      # T = theoretical, O = which option for FE
      write.table(neg, paste0("dataT0",k,"_m1_neg_nobs",nobs[j], ".txt"))
    }
  }
}
}

```

Results model 1

Below, the code for analyzing these results is provided. This is only print this once, for the other models I will directly provide the output.

In this output section, different kinds of tables are shown. The first four tables (option1_fit - option4_fit) contain information about 5 different fit measures: The CFI (with a cut-off of .90 and .95), the TLI (with a

cut-off of .90 and .95), the RMSEA (with a cut-off of .08 and .04), the SRMR (with a cut-off of .10 and .08) and the p-value of the chi-square ($p > .05$: there is no significant difference between the data and the model). All results are given in percentage.

The fifth table provides information about how many models converged (%). The last table shows how many models contained at least one negative variance (%).

Note: We expect that the chi-square test yields more significant results with larger samples, as documented in literature.

```
setwd("H:/home/Doctoraat/Studie 3 - PD SRM/Code/Gebruikt vanaf 10-2-2017/Simulaties model selection/3)
nobs <- c(50,75,100,125,150,200,500,1000)
fit_all <- as.data.frame(matrix(NA, ncol= 9, nrow = 32))
colnames(fit_all) <- c("CFI90", "CFI95", "TLI90", "TLI95", "RMSEA08", "RMSEA04", "CHI2",
                      "SRMR10", "SRMR08")
neg_all <- as.data.frame(matrix(NA, ncol= 4, nrow = 8))
colnames(neg_all) <- c("optie 1", "optie 2", "optie 3", "optie 4")
conv <- as.data.frame(matrix(NA, ncol = 4, nrow = 8))
colnames(conv) <- c("optie 1", "optie 2", "optie 3", "optie 4")
# which option for FE
for( k in 1:4){
  # 8 different sample sizes
  for (j in 1:length(nobs)){
    # make index i, so every unique combination (of option and sample size)
    # is written on a new (ith) row
    l <- k-1
    i <- j + (l*8)
    # fit measures:
    assign(paste0("dataT0",k,"_m1_fit_nobs",nobs[j]),
          read.table(paste0("dataT0",k,"_m1_fit_nobs",nobs[j], ".txt")))
    # T = theoretical data
    # o = which option for FE
    # m = which model
    #how many models have a good fit (%)?
    fit_all[i,1] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$CFI")))) > .90)))/
      length(eval(parse(text=(paste0("dataT0", k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,2] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$CFI")))) > .95)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,3] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$TLI")))) > .90)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,4] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$TLI")))) > .95)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,5] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$RMSEA")))) < .08)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,6] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$RMSEA")))) < .04)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
    fit_all[i,7] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
                                                         "$CHI2")))) > .05)))/
      length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))) *100
```

```

fit_all[i,8] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
"$SRMR")))) < .10))/
length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))*100
fit_all[i,9] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
"$SRMR")))) < .08))/
length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))*100
# negative variances:
assign(paste0("dataT0",k,"_m1_neg_nobs",nobs[j]),
read.table(paste0("dataT0",k,"_m1_neg_nobs",nobs[j], ".txt")))
# percentage of simulations that had at least 1 negative variance
neg_all[j,k] <- length(which(eval(parse(text=(paste0("dataT0",k,"_m1_neg_nobs",nobs[j],
"[,1]")))) != 0))/
length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],"$CFI"))))*100
# percentage of all simulations that converged
conv[j,k] <- length(eval(parse(text=(paste0("dataT0",k,"_m1_fit_nobs",nobs[j],
"$CFI")))))/5
}
}

# results in table format
option1.1_fit <- fit_all[c(1:8),]
rownames(option1.1_fit) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200",
"n=500", "n=1000")
option1.2_fit <- fit_all[c(9:16),]
rownames(option1.2_fit) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200",
"n=500", "n=1000")
option1.3_fit <- fit_all[c(17:24),]
rownames(option1.3_fit) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200",
"n=500", "n=1000")
option1.4_fit <- fit_all[c(25:32),]
rownames(option1.4_fit) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200",
"n=500", "n=1000")
rownames(neg_all) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200", "n=500",
"n=1000")
rownames(conv) <- c("n=50", "n=75", "n=100", "n=125", "n=150", "n=200", "n=500",
"n=1000")

# How did the fit measures react?
option1.1_fit

##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      98.4  89.8  93.8  82.6    75.8    57.2  92.8   76.8   49.4
## n=75      99.6  97.4  99.2  88.4    81.6    62.6  94.2   94.0   73.6
## n=100     100.0  98.6  99.6  94.8    90.4    68.4  93.6   98.6   90.6
## n=125     100.0  99.4 100.0  99.0    95.6    69.8  95.8   99.4   95.6
## n=150     100.0 100.0 100.0  98.8    96.0    70.2  93.0  100.0   97.8
## n=200     100.0 100.0 100.0  99.6    98.4    81.2  95.6  100.0  100.0
## n=500     100.0 100.0 100.0 100.0   100.0    94.0  94.2  100.0  100.0
## n=1000    100.0 100.0 100.0 100.0   100.0    99.4  93.8  100.0  100.0

option1.2_fit

##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      57.6  23.6  33.2  16.0    12.2     6.4  31.8   12.8    3.4
## n=75      61.2  20.8  33.0  10.0     8.0     2.4  18.4   22.6    4.0

```

```
## n=100 62.2 15.2 29.6 5.4 3.0 0.2 6.6 24.8 3.6
## n=125 62.6 10.4 21.8 3.6 2.6 0.6 2.6 27.2 4.4
## n=150 65.6 7.8 22.8 1.4 0.8 0.0 0.6 32.8 3.6
## n=200 67.0 3.8 18.4 0.4 0.2 0.0 0.0 37.8 2.4
## n=500 74.4 0.0 5.6 0.0 0.0 0.0 0.0 48.8 0.6
## n=1000 80.8 0.0 1.0 0.0 0.0 0.0 0.0 59.4 0.0
```

option1.3_fit

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      79.6 49.2 62.4 37.2 29.8 17.6 58.4 29.4 10.2
## n=75      89.4 55.0 67.0 37.2 28.4 11.6 45.8 42.6 18.8
## n=100     92.2 51.4 69.2 31.2 21.4 8.4 30.0 46.2 17.4
## n=125     95.8 56.0 76.2 30.8 22.6 4.6 23.4 55.6 22.6
## n=150     97.0 56.8 80.4 25.4 14.8 3.0 12.6 60.0 22.6
## n=200     98.2 55.8 80.2 25.0 11.8 0.4 4.8 63.8 26.0
## n=500    100.0 59.4 92.0 12.6 2.6 0.0 0.0 81.6 26.6
## n=1000   100.0 55.6 98.2 4.2 0.0 0.0 0.0 91.8 22.4
```

option1.4_fit

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      67.2 32.4 43.2 21.8 18.2 9.0 42.8 17.2 5.4
## n=75      73.4 25.4 40.8 16.4 12.6 3.0 22.2 22.2 6.2
## n=100     70.0 21.0 37.0 9.6 7.0 1.2 10.8 22.8 6.8
## n=125     72.4 20.2 36.8 5.2 4.2 0.8 4.4 25.0 5.6
## n=150     78.0 17.2 40.2 5.6 2.8 0.4 1.4 28.6 5.6
## n=200     81.0 16.0 37.4 3.8 1.6 0.2 0.8 32.8 7.0
## n=500     90.2 4.0 25.0 0.0 0.0 0.0 0.0 30.8 3.0
## n=1000    97.6 0.2 16.6 0.0 0.0 0.0 0.0 30.6 0.4
```

How many models converged (%)?

conv

```
##          optie 1 optie 2 optie 3 optie 4
## n=50          100      100      100      100
## n=75          100      100      100      100
## n=100         100      100      100      100
## n=125         100      100      100      100
## n=150         100      100      100      100
## n=200         100      100      100      100
## n=500         100      100      100      100
## n=1000        100      100      100      100
```

How many models had at least 1 negative variance (%)?

round(neg_all,2)

```
##          optie 1 optie 2 optie 3 optie 4
## n=50          2.0      7.6      4.2      5.6
## n=75          0.6      1.0      0.4      1.2
## n=100         0.0      0.2      0.0      0.4
## n=125         0.0      0.0      0.0      0.0
## n=150         0.0      0.0      0.0      0.0
## n=200         0.0      0.0      0.0      0.0
## n=500         0.0      0.0      0.0      0.0
## n=1000        0.0      0.0      0.0      0.0
```

Conclusion model 1

If in reality all observed scores are of equal importance for the family effect (*option 1* of the simulated data), this model suits very well with these simulated data (as expected).

If, however, some observed scores are more important than others (*option 2, 3 and 4*), this model will terribly underperform.

Note: The RMSEA will underperform with almost all presented models.

Method 2: Set factor loadings free of FE, but fix at least 1

In models 2a, 2b and 2c, the FL's of the family effect are constrained:

(2a) the FL's of the dyads of the same generation are fixed to 1

(2b) the FL's of the parent-child dyads are fixed to 1

(2c) the FL's of the same person are constrained (here: mother)

Here, the assumption is made that the fixed observed variables are of equal importance for the family effect. They also serve as the baseline to which the other observed scores can be compared.

Model 2d uses a similar approach to ULI (i.e., unit loading identification) for the family effect. Here, the FL of one observed score is constrained to 1, while the others are set free. The constrained observed score serves as a baseline to which the other observed scores can be compared: Are they more (or less) important for the family effect?

In search of a general version of the PD SRM, a priori, model 2d is the most suitable of all three models. Models 2a, 2b and 2c can be used for specific research questions.

Model 2a: Constrain dyads from the same generation (6df)

```
SRM_pd2a <- '
# Latent variables
FE =~ 1*MF + MO + MY + FO + FY + 1*YO
I.M =~ 1*MF + 1*MO + 1*MY
I.F =~ 1*MF + 1*FO + 1*FY
I.O =~ 1*MO + 1*FO + 1*YO
I.Y =~ 1*MY + 1*FY + 1*YO
D.MF =~ 1*MF
D.MO =~ 1*MO
D.MY =~ 1*MY
D.FO =~ 1*FO
D.FY =~ 1*FY
D.OY =~ 1*YO

# Variances
FE ~~ VAR.FE*FE
I.M ~~ VAR.I.M*I.M
I.F ~~ VAR.I.F*I.F
I.O ~~ VAR.I.O*I.O
I.Y ~~ VAR.I.Y*I.Y
D.MF ~~ VAR.D.MF*D.MF
D.MO ~~ VAR.D.MO*D.MO
D.MY ~~ VAR.D.MY*D.MY
D.FO ~~ VAR.D.FO*D.FO
D.FY ~~ VAR.D.FY*D.FY
D.OY ~~ VAR.D.OY*D.OY

# Intercepts
FE ~ mean.FE*1
I.M ~ mean.I.M*1
I.F ~ mean.I.F*1
I.O ~ mean.I.O*1
```

```

I.Y ~ mean.I.Y*1
D.MF ~ mean.D.MF*1
D.MO ~ mean.D.MO*1
D.MY ~ mean.D.MY*1
D.FO ~ mean.D.FO*1
D.FY ~ mean.D.FY*1
D.OY ~ mean.D.OY*1

# Constraints
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0
mean.D.MF + mean.D.MO + mean.D.MY == 0
mean.D.MF + mean.D.FO + mean.D.FY == 0
mean.D.MY + mean.D.FY + mean.D.OY == 0
mean.D.MO + mean.D.FO + mean.D.OY == 0

```

Results model 2a

```

# How did the fit measures react?
round(option2a.1_fit,2)

```

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	94.29	92.38	76.19	69.52	57.14	94.29	97.14	91.43
## n=75	100	97.80	96.60	86.40	79.40	65.20	94.00	99.80	97.40
## n=100	100	99.60	98.40	89.40	84.80	62.60	92.60	99.60	99.20
## n=125	100	99.80	99.40	96.00	90.40	70.80	95.60	100.00	100.00
## n=150	100	99.80	99.80	97.40	93.40	69.40	95.00	100.00	100.00
## n=200	100	100.00	99.80	99.00	96.80	74.00	94.80	100.00	100.00
## n=500	100	100.00	100.00	100.00	100.00	93.00	96.00	100.00	100.00
## n=1000	100	100.00	100.00	100.00	100.00	99.00	95.20	100.00	100.00

option2a.2_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99	94.4	92.2	77.8	71.4	56.8	94.6	97.2	89.8
## n=75	100	97.8	96.2	85.6	80.4	64.0	94.4	99.4	98.2
## n=100	100	99.4	98.4	90.8	86.0	66.0	94.8	99.8	99.4
## n=125	100	99.8	99.4	94.2	89.8	65.0	95.2	100.0	99.4
## n=150	100	100.0	100.0	96.4	93.0	73.0	95.4	100.0	99.6
## n=200	100	100.0	100.0	98.2	96.4	77.8	95.0	100.0	100.0
## n=500	100	100.0	100.0	100.0	99.8	90.6	94.4	100.0	100.0
## n=1000	100	100.0	100.0	100.0	100.0	98.4	94.0	100.0	100.0

option2a.3_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	99.8	99.6	98.6	98.6	97.0	94.2	99.8	99.2
## n=75	100	97.0	94.4	85.2	82.0	69.4	89.8	96.4	89.8
## n=100	100	97.0	94.2	86.2	81.4	68.6	89.2	97.0	92.0
## n=125	100	97.6	96.2	82.4	75.6	44.0	83.2	99.0	93.0
## n=150	100	99.8	99.2	84.0	72.4	42.6	77.8	99.6	93.8
## n=200	100	99.8	99.4	91.2	80.0	41.2	74.4	99.8	97.8
## n=500	100	100.0	100.0	98.8	91.2	28.0	40.2	100.0	100.0
## n=1000	100	100.0	100.0	100.0	97.4	13.6	6.0	100.0	100.0

```

option2a.4_fit

##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      100  99.8  99.4  99.0    96.2    15.0  8.4   99.6   99.2
## n=75      100  99.6  99.4  97.6    94.0    23.0 22.8  100.0   99.4
## n=100     100 100.0  99.8  97.4    93.8    30.4 34.2  100.0   99.8
## n=125     100  99.6  99.4  95.2    90.6    68.8 95.6   99.8   99.0
## n=150     100 100.0  99.8  95.4    93.0    69.2 94.6  100.0   99.8
## n=200     100 100.0 100.0  99.0    95.6    76.2 94.6  100.0  100.0
## n=500     100 100.0 100.0 100.0   100.0    88.6 93.0  100.0  100.0
## n=1000    100 100.0 100.0 100.0   100.0    98.0 92.8  100.0  100.0

# How many models converged (%)?
conv

##          optie 1 optie 2 optie 3 optie 4
## n=50          21    100    100    100
## n=75          100    100    100    100
## n=100         100    100    100    100
## n=125         100    100    100    100
## n=150         100    100    100    100
## n=200         100    100    100    100
## n=500         100    100    100    100
## n=1000        100    100    100    100

# How many models had at least 1 negative variance (%)?
round(neg_all,2)

##          optie 1 optie 2 optie 3 optie 4
## n=50        13.33    17.6     0.6     0.6
## n=75         3.40     4.6     6.8     1.2
## n=100        0.00     4.0     5.6     0.8
## n=125        0.80     2.0     0.8     1.0
## n=150        0.00     0.6     1.2     0.0
## n=200        0.00     0.2     0.4     0.0
## n=500        0.00     0.0     0.0     0.0
## n=1000       0.00     0.0     0.0     0.0

```

- Option 1: Only 21% of all models converged with $n = 50$. With the other three options, all models converged.

Starting from $n = 75$: good performance (exception: RMSEA performs only well with larger samples)

- Option 2 and 3: Also perform well with $n = 50$ (*Note:* with option 3 the chi-square test becomes more significant as sample size increases) .
- Option 4: Performs well (expectation chi-square if $n \leq 100$)
- With $n = 50$: data simulated under option 1 and 2 result in fitted models with the most negative variances.

RMSEA only performs adequately starting from $n = 200$

Model 2b: Constrain the parent-child dyads (8df)

```
SRM_pd2b <- '  
# Latent variables  
FE =~ MF + 1*MO + 1*MY + 1*FO + 1*FY + YO  
I.M =~ 1*MF + 1*MO + 1*MY  
I.F =~ 1*MF + 1*FO + 1*FY  
I.O =~ 1*MO + 1*FO + 1*YO  
I.Y =~ 1*MY + 1*FY + 1*YO  
D.MF =~ 1*MF  
D.MO =~ 1*MO  
D.MY =~ 1*MY  
D.FO =~ 1*FO  
D.FY =~ 1*FY  
D.OY =~ 1*YO  
  
# Variances  
FE ~~ VAR.FE*FE  
I.M ~~ VAR.I.M*I.M  
I.F ~~ VAR.I.F*I.F  
I.O ~~ VAR.I.O*I.O  
I.Y ~~ VAR.I.Y*I.Y  
D.MF ~~ VAR.D.MF*D.MF  
D.MO ~~ VAR.D.MO*D.MO  
D.MY ~~ VAR.D.MY*D.MY  
D.FO ~~ VAR.D.FO*D.FO  
D.FY ~~ VAR.D.FY*D.FY  
D.OY ~~ VAR.D.OY*D.OY  
  
# Intercepts  
FE ~ mean.FE*1  
I.M ~ mean.I.M*1  
I.F ~ mean.I.F*1  
I.O ~ mean.I.O*1  
I.Y ~ mean.I.Y*1  
D.MF ~ mean.D.MF*1  
D.MO ~ mean.D.MO*1  
D.MY ~ mean.D.MY*1  
D.FO ~ mean.D.FO*1  
D.FY ~ mean.D.FY*1  
D.OY ~ mean.D.OY*1  
  
# Constraints  
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0  
mean.D.MF + mean.D.MO + mean.D.MY == 0  
mean.D.MF + mean.D.FO + mean.D.FY == 0  
mean.D.MY + mean.D.FY + mean.D.OY == 0  
mean.D.MO + mean.D.FO + mean.D.OY == 0  
'
```

Results model 2b

How did the fit measures react?

`round(option2b.1_fit,2)`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	97.33	93.33	93.33	82.67	78.67	58.67	94.67	89.33	72.0
## n=75	99.80	97.80	98.00	89.80	82.20	63.40	95.00	98.20	88.6
## n=100	100.00	99.40	99.60	92.40	85.80	64.80	94.40	99.00	95.0
## n=125	100.00	99.80	99.80	98.00	93.60	71.60	96.00	100.00	99.0
## n=150	100.00	99.80	100.00	98.20	96.20	71.00	95.80	100.00	99.6
## n=200	100.00	100.00	100.00	99.00	97.60	76.40	94.40	100.00	100.0
## n=500	100.00	100.00	100.00	100.00	100.00	93.20	95.00	100.00	100.0
## n=1000	100.00	100.00	100.00	100.00	100.00	98.80	93.40	100.00	100.0

`option2b.2_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.8	99.0	99.0	98.8	98.4	95.8	93.2	98.6	97.0
## n=75	100.0	99.6	99.8	97.6	95.4	89.0	94.0	98.8	95.4
## n=100	100.0	99.2	99.2	91.8	87.0	68.2	93.2	98.4	92.0
## n=125	100.0	99.6	99.8	94.2	92.0	70.2	93.2	99.2	97.6
## n=150	100.0	100.0	100.0	97.4	94.4	74.0	94.4	100.0	98.8
## n=200	100.0	100.0	100.0	99.2	97.8	76.2	94.8	100.0	100.0
## n=500	100.0	100.0	100.0	100.0	100.0	90.8	93.2	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	99.6	95.2	100.0	100.0

`option2b.3_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100.0	100.0	100.0	99.6	99.4	98.8	95.6	99.6	99.6
## n=75	99.8	94.6	96.0	79.0	71.2	49.0	87.0	94.8	77.8
## n=100	100.0	96.8	97.6	83.4	76.2	47.4	84.6	95.6	87.0
## n=125	100.0	98.6	99.0	87.0	78.6	48.0	82.4	99.2	93.2
## n=150	100.0	99.4	99.4	87.8	79.8	42.8	79.0	99.8	95.4
## n=200	100.0	99.6	99.8	93.4	83.8	44.0	71.6	99.0	98.0
## n=500	100.0	100.0	100.0	99.4	95.8	33.2	37.8	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	99.6	23.6	7.6	100.0	100.0

`option2b.4_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.4	89.6	91.0	77.0	72.8	51.6	84.4	92.2	73.2
## n=75	99.8	95.8	97.6	82.6	75.8	56.0	89.6	97.8	89.8
## n=100	100.0	98.6	98.8	92.6	86.4	61.6	93.6	99.4	97.0
## n=125	100.0	99.4	99.6	94.0	88.6	59.2	92.0	100.0	98.0
## n=150	100.0	100.0	100.0	96.8	91.4	63.6	90.8	100.0	100.0
## n=200	100.0	100.0	100.0	98.4	95.4	63.8	88.6	100.0	99.8
## n=500	100.0	100.0	100.0	100.0	100.0	73.0	79.0	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	83.0	61.8	100.0	100.0

How many models converged (%)?

`conv`

##	optie 1	optie 2	optie 3	optie 4
## n=50	15	100	100	100
## n=75	100	100	100	100

```
## n=100      100      100      100      100
## n=125      100      100      100      100
## n=150      100      100      100      100
## n=200      100      100      100      100
## n=500      100      100      100      100
## n=1000     100      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50      6.67      3.2      0.2      6.6
## n=75      0.40      5.6      1.0      0.8
## n=100     0.20     10.6      0.4      0.0
## n=125     0.20      8.2      0.0      0.0
## n=150     0.00      4.6      0.0      0.0
## n=200     0.00      2.2      0.0      0.0
## n=500     0.00      0.0      0.0      0.0
## n=1000    0.00      0.0      0.0      0.0
```

Similar results as with the previous model

- Option 1: Only 15% of all models converged with $n = 50$. With the other three options, all models converged.
if $n \geq 75$: fit indices all perform very well
- Option 2: also good results with $n=50$
- Option 3: CFI, TLI and SRMR perform good.

Exception: RMSEA only performs well in larger samples.

Model 2c: Constrain everything of 1 person (here: mother)

```
SRM_pd2c <- '  
# Latent variables  
FE =~ 1*MF + 1*MO + 1*MY + FO + FY + YO  
I.M =~ 1*MF + 1*MO + 1*MY  
I.F =~ 1*MF + 1*FO + 1*FY  
I.O =~ 1*MO + 1*FO + 1*YO  
I.Y =~ 1*MY + 1*FY + 1*YO  
D.MF =~ 1*MF  
D.MO =~ 1*MO  
D.MY =~ 1*MY  
D.FO =~ 1*FO  
D.FY =~ 1*FY  
D.OY =~ 1*YO  
  
# Variances  
FE ~~ VAR.FE*FE  
I.M ~~ VAR.I.M*I.M  
I.F ~~ VAR.I.F*I.F  
I.O ~~ VAR.I.O*I.O  
I.Y ~~ VAR.I.Y*I.Y  
D.MF ~~ VAR.D.MF*D.MF  
D.MO ~~ VAR.D.MO*D.MO  
D.MY ~~ VAR.D.MY*D.MY  
D.FO ~~ VAR.D.FO*D.FO  
D.FY ~~ VAR.D.FY*D.FY  
D.OY ~~ VAR.D.OY*D.OY  
  
# Intercepts  
FE ~ mean.FE*1  
I.M ~ mean.I.M*1  
I.F ~ mean.I.F*1  
I.O ~ mean.I.O*1  
I.Y ~ mean.I.Y*1  
D.MF ~ mean.D.MF*1  
D.MO ~ mean.D.MO*1  
D.MY ~ mean.D.MY*1  
D.FO ~ mean.D.FO*1  
D.FY ~ mean.D.FY*1  
D.OY ~ mean.D.OY*1  
  
# Constraints  
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0  
mean.D.MF + mean.D.MO + mean.D.MY == 0  
mean.D.MF + mean.D.FO + mean.D.FY == 0  
mean.D.MY + mean.D.FY + mean.D.OY == 0  
mean.D.MO + mean.D.FO + mean.D.OY == 0  
'
```

Results model 2c

How did the fit measures react?

`round(option2c.1_fit,2)`

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      98.82  91.76  91.76  77.65    71.76    57.65  94.12  95.29  82.35
## n=75     100.00  98.42  98.42  90.00    85.26    70.00  96.32 100.00  96.84
## n=100     100.00  99.20  98.80  91.20    86.20    63.20  92.60  99.80  97.40
## n=125     100.00  99.40  99.20  94.80    89.40    69.00  94.60 100.00  99.60
## n=150     100.00 100.00 100.00  97.60    94.20    71.00  94.40 100.00  99.80
## n=200     100.00 100.00 100.00  99.40    97.20    75.80  94.60 100.00 100.00
## n=500     100.00 100.00 100.00 100.00   100.00    92.00  95.00 100.00 100.00
## n=1000    100.00 100.00 100.00 100.00   100.00   100.00  95.00 100.00 100.00
```

`option2c.2_fit`

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      98.6   97.4   97.2   95.2     95.0     93.2  93.2   95.8   94.6
## n=75      98.4   79.6   77.0   51.2     44.4     28.8  68.2   79.6   52.0
## n=100     98.6   77.8   73.8   45.0     34.0     13.0  52.4   82.2   55.0
## n=125     99.2   81.2   77.4   40.0     29.6     10.4  38.8   90.0   57.4
## n=150     100.0   84.0   80.4   40.8     30.6       7.0  32.2   95.2   62.0
## n=200     100.0   89.0   85.8   37.8     24.0       4.6  15.4   96.6   71.4
## n=500     100.0   96.8   93.0   22.0       9.2       0.2   0.2  100.0   84.0
## n=1000    100.0   99.6   98.8   13.6       1.6       0.0   0.0  100.0   95.2
```

`option2c.3_fit`

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      99.4   96.8   95.8   28.0     18.0     12.6  24.2   95.6   88.2
## n=75     100.0   95.4   94.2   78.8     70.4     50.8  88.2   96.0   84.4
## n=100     100.0   98.0   97.8   83.0     73.4     50.2  86.8   98.2   90.4
## n=125     100.0   98.4   97.8   86.0     78.6     47.8  83.6   99.4   96.2
## n=150     100.0   99.0   98.4   88.0     78.8     44.4  79.8   99.2   95.0
## n=200     100.0  100.0  100.0   93.2     84.2     44.6  77.4   99.6   98.6
## n=500     100.0  100.0  100.0   99.2     95.4     33.8  41.8  100.0  100.0
## n=1000    100.0  100.0  100.0  100.0     98.8     25.6  10.0  100.0  100.0
```

`option2c.4_fit`

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      99.6   98.4   98.2   96.2     94.6     26.4  17.4   96.6   93.8
## n=75      99.0   87.2   85.2   63.2     55.0     34.2  76.8   87.0   69.2
## n=100     99.8   91.4   90.2   64.0     52.8     27.8  69.2   93.2   71.8
## n=125     99.8   94.4   92.6   66.4     52.8     27.0  62.2   94.4   76.8
## n=150     99.8   94.8   93.6   65.8     51.8     20.6  53.8   95.8   80.8
## n=200     100.0   96.6   95.8   65.8     49.8     16.6  41.4   97.6   87.2
## n=500     100.0  100.0   99.8   71.6     45.8       2.4   3.0  100.0   97.6
## n=1000    100.0  100.0  100.0   83.4     47.4       0.0   0.0  100.0   99.8
```

How many models converged (%)?

`conv`

```
##          optie 1  optie 2  optie 3  optie 4
## n=50           17       100       100       100
## n=75           38       100       100       100
```



```
## n=100      100      100      100      100
## n=125      100      100      100      100
## n=150      100      100      100      100
## n=200      100      100      100      100
## n=500      100      100      100      100
## n=1000     100      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50      10.59      3.4      4.2      4.0
## n=75       3.68     17.6      6.0     11.4
## n=100      1.40     12.2      1.2      4.6
## n=125      0.80      6.0      0.8      1.8
## n=150      0.60      5.8      0.6      1.8
## n=200      0.20      3.0      0.2      1.2
## n=500      0.00      0.0      0.0      0.0
## n=1000     0.00      0.0      0.0      0.0
```

- Option 1: Only 17% and 38% of all models converged with sample sizes of 50 and 75, respectively. For the other sample sizes, the results are all good.
- Option 2: Except for the smallest sample, model 2c underperforms. This is the case with almost all fit indices (exception: CFI90)
- Option 3: In general, good fit with $n \geq 125$ (exceptions: RMSEA and Chi-square)
- Option 4: In general, good fit with $n \geq 200$ (exceptions: RMSEA and Chi-square)
- Option 2-4: Chi-square test (and RMSEA) becomes worse with increasing sample sizes. This is what we would expect for the chi-square test. Also, TLI performs badly with option 2.
- Negative variances seem to be present with all options in smaller samples

Conclusion: of all models, this model seems to be the least suitable.

Model 2d: Set factor loadings free of FE, but fix 1 (ULI; 5df)

```
SRM_pd2d <- '  
# Latent variables  
FE =~ 1*MF + MO + MY + FO + FY + YO  
I.M =~ 1*MF + 1*MO + 1*MY  
I.F =~ 1*MF + 1*FO + 1*FY  
I.O =~ 1*MO + 1*FO + 1*YO  
I.Y =~ 1*MY + 1*FY + 1*YO  
D.MF =~ 1*MF  
D.MO =~ 1*MO  
D.MY =~ 1*MY  
D.FO =~ 1*FO  
D.FY =~ 1*FY  
D.OY =~ 1*YO  
  
# Variances  
FE ~~ VAR.FE*FE  
I.M ~~ VAR.I.M*I.M  
I.F ~~ VAR.I.F*I.F  
I.O ~~ VAR.I.O*I.O  
I.Y ~~ VAR.I.Y*I.Y  
D.MF ~~ VAR.D.MF*D.MF  
D.MO ~~ VAR.D.MO*D.MO  
D.MY ~~ VAR.D.MY*D.MY  
D.FO ~~ VAR.D.FO*D.FO  
D.FY ~~ VAR.D.FY*D.FY  
D.OY ~~ VAR.D.OY*D.OY  
  
# Intercepts  
FE ~ mean.FE*1  
I.M ~ mean.I.M*1  
I.F ~ mean.I.F*1  
I.O ~ mean.I.O*1  
I.Y ~ mean.I.Y*1  
D.MF ~ mean.D.MF*1  
D.MO ~ mean.D.MO*1  
D.MY ~ mean.D.MY*1  
D.FO ~ mean.D.FO*1  
D.FY ~ mean.D.FY*1  
D.OY ~ mean.D.OY*1  
  
# Constraints  
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0  
mean.D.MF + mean.D.MO + mean.D.MY == 0  
mean.D.MF + mean.D.FO + mean.D.FY == 0  
mean.D.MY + mean.D.FY + mean.D.OY == 0  
mean.D.MO + mean.D.FO + mean.D.OY == 0  
'
```

Results model 2d

How did the fit measures react?

`round(option2d.1_fit,2)`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100	85.71	85.71	57.14	28.57	100.00	85.71	71.43
## n=75	100	100	92.00	80.00	76.00	64.00	92.00	100.00	100.00
## n=100	100	100	99.35	92.86	88.31	65.58	94.81	100.00	100.00
## n=125	100	100	100.00	93.42	85.19	63.37	93.83	100.00	100.00
## n=150	100	100	99.72	95.87	91.74	72.73	95.04	100.00	100.00
## n=200	100	100	100.00	98.00	96.00	78.60	95.60	100.00	100.00
## n=500	100	100	100.00	100.00	100.00	89.00	94.00	100.00	100.00
## n=1000	100	100	100.00	100.00	100.00	98.00	96.80	100.00	100.00

`option2d.2_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	99.8	99.0	97.6	97.4	94.6	96.4	100.0	99.2
## n=75	100	99.8	99.0	98.0	97.8	95.0	96.2	99.8	99.4
## n=100	100	99.8	97.8	90.4	85.0	64.8	95.2	99.6	99.4
## n=125	100	100.0	98.8	91.0	86.0	64.6	92.8	100.0	99.6
## n=150	100	100.0	99.6	95.8	91.8	73.0	95.0	100.0	100.0
## n=200	100	100.0	100.0	98.0	96.2	80.6	95.8	100.0	100.0
## n=500	100	100.0	100.0	100.0	100.0	89.0	94.0	100.0	100.0
## n=1000	100	100.0	100.0	100.0	100.0	97.2	94.8	100.0	100.0

`option2d.3_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100.0	99.8	99.6	99.2	95.8	95.2	100.0	100.0
## n=75	100	99.8	99.4	97.6	96.6	90.6	95.2	99.8	99.8
## n=100	100	99.8	99.4	97.6	96.6	90.6	95.2	99.8	99.8
## n=125	100	99.8	99.8	98.0	97.6	91.2	94.8	100.0	100.0
## n=150	100	100.0	99.8	96.4	93.0	71.0	95.2	100.0	100.0
## n=200	100	100.0	100.0	99.0	96.2	77.2	95.8	100.0	100.0
## n=500	100	100.0	100.0	100.0	99.8	89.4	94.2	100.0	100.0
## n=1000	100	100.0	100.0	100.0	100.0	98.0	95.6	100.0	100.0

`option2d.4_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100.0	100.0	99.8	99.8	97.8	95.8	100	100.0
## n=75	100	100.0	99.8	98.6	98.0	95.0	95.8	100	99.8
## n=100	100	100.0	98.6	89.2	84.6	68.4	94.8	100	99.8
## n=125	100	99.6	98.8	91.0	85.4	67.0	93.4	100	99.8
## n=150	100	99.6	99.2	93.2	88.0	68.8	95.4	100	99.8
## n=200	100	100.0	100.0	98.0	95.2	75.2	94.4	100	100.0
## n=500	100	100.0	100.0	100.0	100.0	88.2	94.8	100	100.0
## n=1000	100	100.0	100.0	100.0	100.0	98.2	96.4	100	100.0

How many models converged (%)?

`conv`

##	optie 1	optie 2	optie 3	optie 4
## n=50	1.4	100	100	100
## n=75	5.0	100	100	100

```
## n=100      30.8      100      100      100
## n=125      48.6      100      100      100
## n=150      72.6      100      100      100
## n=200     100.0      100      100      100
## n=500     100.0      100      100      100
## n=1000    100.0      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50      57.14      5.2      1.4      1.4
## n=75      16.00      5.2      4.2      0.6
## n=100      5.19     16.6      4.2      4.8
## n=125      3.70      9.0      2.6      2.2
## n=150      0.55      3.6      3.2      0.8
## n=200      0.20      2.0      0.4      0.2
## n=500      0.00      0.0      0.0      0.0
## n=1000     0.00      0.0      0.0      0.0
```

Results:

- If all FL's of the family effect in reality are of equal importance (i.e. *Option1*), this model underperforms with samples smaller than 200. Only 1.4% of all models converged with $n = 50$ and 72.6% of all models with $n = 150$ did. Also, a lot of negative variances seem to be present in small samples. For the models that did converged, results were good for samples if $n \geq 75$ or 100.
- The other three kinds of simulated data (*Option2*, 3 and 4) perform very well: all models converged and the fit indices as well as the chi-square difference tests showed an excellent model fit. *Note* : Also the RMSEA performs well.

Method 3: lambda approach

In this section, I use the lambda approach for the factor loadings of the family effect. This is an alternative identification strategy where the mean of the factor loadings equals 1. Model 3c is the most general model with the most straightforward interpretation of the components.

Model 3a: Constrain one FL as a Reference (Sum other FL's = 5).

```
SRM_pd3a <- '  
  # Latent variables  
  FE =~ 1*MF + lambdaMO*MO + lambdaMY*MY + lambdaFO*FO + lambdaFY*FY + lambdaYO*YO  
  I.M =~ 1*MF + 1*MO + 1*MY  
  I.F =~ 1*MF + 1*FO + 1*FY  
  I.O =~ 1*MO + 1*FO + 1*YO  
  I.Y =~ 1*MY + 1*FY + 1*YO  
  D.MF =~ 1*MF  
  D.MO =~ 1*MO  
  D.MY =~ 1*MY  
  D.FO =~ 1*FO  
  D.FY =~ 1*FY  
  D.OY =~ 1*YO  
  
  # Variances  
  FE ~~ VAR.FE*FE  
  I.M ~~ VAR.I.M*I.M  
  I.F ~~ VAR.I.F*I.F  
  I.O ~~ VAR.I.O*I.O  
  I.Y ~~ VAR.I.Y*I.Y  
  D.MF ~~ VAR.D.MF*D.MF  
  D.MO ~~ VAR.D.MO*D.MO  
  D.MY ~~ VAR.D.MY*D.MY  
  D.FO ~~ VAR.D.FO*D.FO  
  D.FY ~~ VAR.D.FY*D.FY  
  D.OY ~~ VAR.D.OY*D.OY  
  
  # Intercepts  
  FE ~ mean.FE*1  
  I.M ~ mean.I.M*1  
  I.F ~ mean.I.F*1  
  I.O ~ mean.I.O*1  
  I.Y ~ mean.I.Y*1  
  D.MF ~ mean.D.MF*1  
  D.MO ~ mean.D.MO*1  
  D.MY ~ mean.D.MY*1  
  D.FO ~ mean.D.FO*1  
  D.FY ~ mean.D.FY*1  
  D.OY ~ mean.D.OY*1  
  
  # Constraints  
  mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0  
  mean.D.MF + mean.D.MO + mean.D.MY == 0  
  mean.D.MF + mean.D.FO + mean.D.FY == 0
```

```

mean.D.MY + mean.D.FY + mean.D.OY == 0
mean.D.MO + mean.D.FO + mean.D.OY == 0
# set constraints on factor loadings FE for identifiability
lambdaMO+ lambdaMY+ lambdaFO + lambdaFY + lambdaYO==5

```

Results model 3a

How did the fit measures react?
option3a.1_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.4	94.4	91.2	78.2	72.0	59.6	94.8	97.2	90.2
## n=75	100.0	98.4	96.2	86.0	81.6	64.2	93.2	98.8	97.2
## n=100	100.0	99.8	98.6	93.2	88.8	65.6	95.8	99.8	99.8
## n=125	100.0	100.0	99.8	95.8	90.4	70.6	95.4	99.8	99.8
## n=150	100.0	100.0	100.0	96.6	92.0	69.4	94.8	100.0	100.0
## n=200	100.0	100.0	100.0	99.0	96.8	78.4	95.4	100.0	100.0
## n=500	100.0	100.0	100.0	100.0	100.0	90.6	95.4	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	98.0	93.8	100.0	100.0

option3a.2_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	98.8	86.4	80.2	58.4	51.2	39.8	85.2	86.2	63.0
## n=75	99.2	90.0	83.0	59.8	51.6	32.6	77.2	91.4	77.2
## n=100	99.6	94.2	87.4	64.2	53.8	32.0	71.8	97.0	84.6
## n=125	100.0	93.2	87.4	60.6	50.8	23.6	62.6	97.2	86.6
## n=150	100.0	96.0	91.8	61.2	52.4	22.0	57.8	98.6	89.4
## n=200	100.0	98.0	93.6	60.8	44.8	12.4	40.2	99.4	93.6
## n=500	100.0	100.0	98.6	58.0	35.2	1.0	2.8	100.0	99.6
## n=1000	100.0	100.0	100.0	53.8	22.2	0.0	0.0	100.0	100.0

option3a.3_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.8	95.0	90.8	78.4	70.0	56.8	94.2	97.0	87.0
## n=75	99.8	98.0	96.8	85.4	80.2	65.0	94.8	99.2	98.0
## n=100	100.0	99.2	98.2	90.8	85.8	63.4	94.6	99.8	99.0
## n=125	100.0	100.0	99.6	93.4	89.2	66.6	93.4	100.0	99.6
## n=150	100.0	99.6	99.6	96.4	93.2	73.4	94.8	100.0	100.0
## n=200	100.0	100.0	100.0	99.2	95.2	70.2	92.8	100.0	100.0
## n=500	100.0	100.0	100.0	100.0	100.0	89.2	93.4	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	97.4	94.6	100.0	100.0

option3a.4_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.0	86.4	81.0	63.0	57.8	44.8	86.4	76.8	59.0
## n=75	100.0	92.4	87.4	66.8	58.2	40.0	81.2	87.4	71.0
## n=100	99.6	93.6	88.6	63.8	56.6	31.8	70.6	89.2	73.6
## n=125	100.0	96.2	91.8	67.8	58.2	31.2	67.4	92.4	79.4
## n=150	100.0	97.4	94.6	71.0	59.0	29.8	63.0	95.8	80.2
## n=200	100.0	98.8	96.6	71.8	58.0	21.8	52.2	96.4	86.6
## n=500	100.0	100.0	99.6	77.2	56.4	5.0	9.0	99.8	92.4

```
## n=1000 100.0 100.0 100.0 89.2 59.0 0.6 0.0 100.0 99.0
```

```
# How many models converged (%)?
```

```
conv
```

```
##      optie 1 optie 2 optie 3 optie 4
## n=50      100      100      100      100
## n=75      100      100      100      100
## n=100     100      100      100      100
## n=125     100      100      100      100
## n=150     100      100      100      100
## n=200     100      100      100      100
## n=500     100      100      100      100
## n=1000    100      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
```

```
round(neg_all,2)
```

```
##      optie 1 optie 2 optie 3 optie 4
## n=50      14.8      51.2      21.8      35.2
## n=75       4.0      38.4       9.8      17.2
## n=100       1.4      31.4       2.4       9.8
## n=125       0.8      27.6       0.6       5.4
## n=150       0.0      23.8       0.0       3.2
## n=200       0.0      16.0       0.0       1.6
## n=500       0.0       5.0       0.0       0.0
## n=1000      0.0       1.8       0.0       0.0
```

All models converged. With small samples, a lot of negative variances were found.

Options 1 and 3: CFI and chi-square test perform well with both small ($n = 50$) and larger samples. TLI performs well starting from $n = 75$. RMSEA only performs adequately with samples of 200 families or more.

Option 2 and 4: chi-square test and RMSEA perform worse with larger samples. The former is what we could expect.

Option 3: Good results if $n \geq 75$ for least stringent cut offs.

Drawback: not so easy to interpret the different components.

Model 3b: lambda for every person

```
SRM_pd3b <- '
# Latent variables
FE =~ lambdaMF*MF + lambdaMO*MO + lambdaMY*MY + lambdaFO*FO + lambdaFY*FY + lambdaYO*YO
I.M =~ 1*MF + 1*MO + 1*MY
I.F =~ 1*MF + 1*FO + 1*FY
I.O =~ 1*MO + 1*FO + 1*YO
I.Y =~ 1*MY + 1*FY + 1*YO
D.MF =~ 1*MF
D.MO =~ 1*MO
D.MY =~ 1*MY
D.FO =~ 1*FO
D.FY =~ 1*FY
D.OY =~ 1*YO

# Variances
FE ~~ VAR.FE*FE
I.M ~~ VAR.I.M*I.M
I.F ~~ VAR.I.F*I.F
I.O ~~ VAR.I.O*I.O
I.Y ~~ VAR.I.Y*I.Y
D.MF ~~ VAR.D.MF*D.MF
D.MO ~~ VAR.D.MO*D.MO
D.MY ~~ VAR.D.MY*D.MY
D.FO ~~ VAR.D.FO*D.FO
D.FY ~~ VAR.D.FY*D.FY
D.OY ~~ VAR.D.OY*D.OY

# Intercepts
FE ~ mean.FE*1
I.M ~ mean.I.M*1
I.F ~ mean.I.F*1
I.O ~ mean.I.O*1
I.Y ~ mean.I.Y*1
D.MF ~ mean.D.MF*1
D.MO ~ mean.D.MO*1
D.MY ~ mean.D.MY*1
D.FO ~ mean.D.FO*1
D.FY ~ mean.D.FY*1
D.OY ~ mean.D.OY*1

# Constraints
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0
mean.D.MF + mean.D.MO + mean.D.MY == 0
mean.D.MF + mean.D.FO + mean.D.FY == 0
mean.D.MY + mean.D.FY + mean.D.OY == 0
mean.D.MO + mean.D.FO + mean.D.OY == 0

# set constraints on factor loadings FE for identifiability
lambdaMF + lambdaMO + lambdaMY == 3
lambdaMF + lambdaFO + lambdaFY == 3
lambdaMO + lambdaFO + lambdaYO == 3
```



```
lambdaMY + lambdaFY + lambdaY0 == 3
,
```

Results model 3b

How did the fit measures react?

option3b.1_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	98.4	91.6	92.4	80.2	74.2	56.6	93.6	87.0	65.0
## n=75	100.0	96.2	97.0	86.2	79.0	62.6	92.4	95.6	81.8
## n=100	100.0	99.2	99.2	93.6	87.4	66.8	94.2	99.8	93.2
## n=125	100.0	99.6	99.6	97.6	92.8	68.8	95.2	99.6	98.0
## n=150	100.0	100.0	100.0	97.0	93.6	68.0	93.2	100.0	99.0
## n=200	100.0	100.0	100.0	99.2	99.0	79.8	95.4	100.0	100.0
## n=500	100.0	100.0	100.0	100.0	100.0	92.2	93.8	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	99.2	95.4	100.0	100.0

option3b.2_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.6	93.2	95.2	78.4	74.8	60.0	95.8	85.4	60.0
## n=75	100.0	97.4	98.0	89.8	86.0	65.0	94.6	95.6	82.2
## n=100	100.0	99.0	99.2	91.8	87.0	65.0	94.2	98.6	92.8
## n=125	100.0	99.8	100.0	95.8	92.2	67.8	95.2	99.4	96.6
## n=150	100.0	100.0	100.0	97.4	95.2	73.6	94.6	100.0	98.4
## n=200	100.0	100.0	100.0	99.0	98.0	76.6	94.8	100.0	99.4
## n=500	100.0	100.0	100.0	100.0	99.8	91.8	93.6	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	98.8	94.6	100.0	100.0

option3b.3_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	96.6	81.4	83.0	62.0	54.4	41.2	82.8	65.4	40.2
## n=75	98.6	88.0	89.0	69.0	59.6	39.8	84.4	78.0	55.2
## n=100	100.0	93.2	94.8	76.4	65.6	38.0	78.4	88.0	63.2
## n=125	99.8	96.2	97.2	78.4	67.8	35.8	73.4	91.2	74.2
## n=150	100.0	97.2	97.8	83.6	73.2	34.2	72.6	96.8	81.2
## n=200	100.0	99.2	99.8	84.4	72.4	31.6	59.2	97.4	84.0
## n=500	100.0	100.0	100.0	97.0	83.8	14.4	18.0	100.0	98.4
## n=1000	100.0	100.0	100.0	98.6	89.6	3.4	1.2	100.0	99.8

option3b.4_fit

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.0	91.0	92.0	79.4	74.6	58.2	94.4	86.4	62.6
## n=75	99.8	96.6	97.2	86.4	81.4	64.4	94.0	94.6	83.2
## n=100	100.0	98.8	99.4	94.4	89.0	68.2	94.6	99.2	94.4
## n=125	100.0	100.0	100.0	95.8	89.8	66.2	92.4	99.8	97.2
## n=150	100.0	99.8	99.8	98.6	97.0	76.8	96.4	100.0	98.2
## n=200	100.0	100.0	100.0	99.0	97.6	75.8	94.2	100.0	99.8
## n=500	100.0	100.0	100.0	100.0	100.0	93.2	95.6	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	99.2	95.8	100.0	100.0

```
# How many models converged (%)?
```

```
conv
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50         100     100     100     100
## n=75         100     100     100     100
## n=100        100     100     100     100
## n=125        100     100     100     100
## n=150        100     100     100     100
## n=200        100     100     100     100
## n=500        100     100     100     100
## n=1000       100     100     100     100
```

```
# How many models had at least 1 negative variance (%)?
```

```
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50         4.0    15.2     9.6     9.4
## n=75         1.4     4.8     4.4     2.2
## n=100        0.0     3.4     0.6     1.0
## n=125        0.0     1.8     0.0     0.2
## n=150        0.0     0.4     0.0     0.2
## n=200        0.0     0.0     0.0     0.0
## n=500        0.0     0.0     0.0     0.0
## n=1000       0.0     0.0     0.0     0.0
```

For all options: All models converged, only with small samples negative variances were found.

Options 1, 2 and 4: CFI and chi-square test perform well with both small ($n = 50$) and larger samples. TLI performs well starting from $n = 75$. RMSEA only performs well with samples of 200 families or more. In general, for $n = 75$: good results with least stringent cut-offs for fit indices. For the stringent cut-offs: good results for $n \geq 100$ or 125.

Option 3: chi-square test and RMSEA perform worse with larger samples.

Model 3c: Average all FL of FC = 1

This model yields the most straightforward interpretation of all models of this section.

```
SRM_pd3c <- '  
# Latent variables  
FE =~ lambdaMF*MF + lambdaMO*MO + lambdaMY*MY + lambdaFO*FO + lambdaFY*FY + lambdaYO*YO  
I.M =~ 1*MF + 1*MO + 1*MY  
I.F =~ 1*MF + 1*FO + 1*FY  
I.O =~ 1*MO + 1*FO + 1*YO  
I.Y =~ 1*MY + 1*FY + 1*YO  
D.MF =~ 1*MF  
D.MO =~ 1*MO  
D.MY =~ 1*MY  
D.FO =~ 1*FO  
D.FY =~ 1*FY  
D.OY =~ 1*YO  
  
# Variances  
FE ~~ VAR.FE*FE  
I.M ~~ VAR.I.M*I.M  
I.F ~~ VAR.I.F*I.F  
I.O ~~ VAR.I.O*I.O  
I.Y ~~ VAR.I.Y*I.Y  
D.MF ~~ VAR.D.MF*D.MF  
D.MO ~~ VAR.D.MO*D.MO  
D.MY ~~ VAR.D.MY*D.MY  
D.FO ~~ VAR.D.FO*D.FO  
D.FY ~~ VAR.D.FY*D.FY  
D.OY ~~ VAR.D.OY*D.OY  
  
# Intercepts  
FE ~ mean.FE*1  
I.M ~ mean.I.M*1  
I.F ~ mean.I.F*1  
I.O ~ mean.I.O*1  
I.Y ~ mean.I.Y*1  
D.MF ~ mean.D.MF*1  
D.MO ~ mean.D.MO*1  
D.MY ~ mean.D.MY*1  
D.FO ~ mean.D.FO*1  
D.FY ~ mean.D.FY*1  
D.OY ~ mean.D.OY*1  
  
# Constraints  
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0  
mean.D.MF + mean.D.MO + mean.D.MY == 0  
mean.D.MF + mean.D.FO + mean.D.FY == 0  
mean.D.MY + mean.D.FY + mean.D.OY == 0  
mean.D.MO + mean.D.FO + mean.D.OY == 0  
  
# set constraints on factor loadings FE for identifiability  
lambdaMF + lambdaMO + lambdaMY + lambdaFO + lambdaFY + lambdaYO == 6  
'
```

Results model 3c

How did the fit measures react?

`round(option3c.1_fit,2)`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100	90.00	90.00	60.00	20.00	100.00	90	90
## n=75	100	100	90.91	81.82	77.27	63.64	90.91	100	100
## n=100	100	100	99.35	92.86	88.31	65.58	94.81	100	100
## n=125	100	100	100.00	93.42	85.19	63.37	93.83	100	100
## n=150	100	100	99.72	95.87	91.74	72.73	95.04	100	100
## n=200	100	100	100.00	98.00	96.00	78.60	95.60	100	100
## n=500	100	100	100.00	100.00	100.00	89.00	94.00	100	100
## n=1000	100	100	100.00	100.00	100.00	98.00	96.80	100	100

`option3c.2_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100.0	99.2	98.2	98.0	95.8	96.6	100.0	99.6
## n=75	100	100.0	99.6	99.2	98.8	96.0	96.6	99.8	99.6
## n=100	100	99.8	97.8	90.4	85.0	64.8	95.2	99.8	99.6
## n=125	100	100.0	98.8	91.0	86.0	64.6	92.8	100.0	99.8
## n=150	100	100.0	99.6	95.8	91.8	73.0	95.0	100.0	100.0
## n=200	100	100.0	100.0	98.0	96.2	80.6	95.8	100.0	100.0
## n=500	100	100.0	100.0	100.0	100.0	89.0	94.0	100.0	100.0
## n=1000	100	100.0	100.0	100.0	100.0	97.2	94.8	100.0	100.0

`option3c.3_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	100	100.0	99.8	99.2	98.8	95.0	95.4	100.0	99.6
## n=75	100	99.8	99.6	97.8	96.8	91.0	95.4	99.8	99.8
## n=100	100	100.0	99.0	92.2	87.6	70.2	95.8	100.0	99.8
## n=125	100	100.0	99.2	94.2	90.0	69.0	95.4	100.0	99.8
## n=150	100	100.0	99.8	96.6	92.6	68.8	95.4	100.0	100.0
## n=200	100	100.0	100.0	98.4	95.8	72.2	95.0	100.0	100.0
## n=500	100	100.0	100.0	100.0	99.6	89.6	95.4	100.0	100.0
## n=1000	100	100.0	100.0	100.0	100.0	98.2	95.2	100.0	100.0

`option3c.4_fit`

##	CFI90	CFI95	TLI90	TLI95	RMSEA08	RMSEA04	CHI2	SRMR10	SRMR08
## n=50	99.8	99.4	98.6	98.0	97.2	94.4	94.8	99.8	98.8
## n=75	100.0	99.2	93.6	82.8	79.0	64.0	93.8	99.6	98.2
## n=100	100.0	99.4	97.8	90.0	86.4	67.0	94.4	100.0	99.8
## n=125	100.0	100.0	99.8	92.2	86.8	68.6	94.8	100.0	100.0
## n=150	100.0	100.0	100.0	96.2	91.8	69.6	95.2	100.0	100.0
## n=200	100.0	100.0	100.0	98.6	96.6	76.0	95.2	100.0	100.0
## n=500	100.0	100.0	100.0	100.0	100.0	88.8	95.0	100.0	100.0
## n=1000	100.0	100.0	100.0	100.0	100.0	98.4	96.4	100.0	100.0

How many models converged (%)?

`conv`

##	optie 1	optie 2	optie 3	optie 4
## n=50	2.0	100	100	100
## n=75	4.4	100	100	100

```
## n=100      30.8      100      100      100
## n=125      48.6      100      100      100
## n=150      72.6      100      100      100
## n=200     100.0      100      100      100
## n=500     100.0      100      100      100
## n=1000    100.0      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50      40.00      2.8      2.2      5.2
## n=75      18.18      2.4      3.8     12.0
## n=100      5.19     16.4      8.8      6.0
## n=125      3.70      8.8      4.6      1.4
## n=150      0.55      3.6      2.0      1.0
## n=200      0.20      2.0      0.6      0.8
## n=500      0.00      0.0      0.0      0.0
## n=1000     0.00      0.0      0.0      0.0
```

- Option 1: A lot of models did not converge (n = 50: only 2% did; n = 150: only 72.6% did). Also, a lot of negative variances were present in smaller samples. Starting from n = 100: good results.
- Option 2, 3 and 4: even with small samples we achieve good results.

UPDATE: The analyses were redone because the models that did converge fitted very well with the data. I hereby increased the number of iterations from 250 to 20000. The results are shown below. Result: By increasing the number of iterations, all models converged for samples n = 150. But, the convergence problem persisted with smaller samples. The models that did converge, show an excellent fit, though. Even with small samples.

```
# How many models converged (%)?
conv
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50         6.2      100      100      100
## n=75        13.0      100      100      100
## n=100       32.0      100      100      100
## n=125       32.0      100      100      100
## n=150      100.0      100      100      100
## n=200      100.0      100      100      100
## n=500      100.0      100      100      100
## n=1000     100.0      100      100      100
```

```
# How many models had at least 1 negative variance (%)?
round(neg_all,2)
```

```
##          optie 1 optie 2 optie 3 optie 4
## n=50      25.81     10.8      1.6      1.6
## n=75      20.00     11.0      1.4     15.0
## n=100      3.75      4.0      4.0      5.8
## n=125      4.38      9.0      4.2      1.4
## n=150      1.40      7.4      1.6      0.2
## n=200      0.40      1.8      0.8      0.0
## n=500      0.00      0.0      0.0      0.0
## n=1000     0.00      0.0      0.0      0.0
```

```
# How did the fit measures react?
round(option3c.1_fit,2)
```

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      100 100.00 100.00  93.55   87.10   67.74 100.00    100  96.77
## n=75      100  98.46  96.92  83.08   78.46   66.15  98.46    100  98.46
## n=100     100 100.00 100.00  85.62   81.88   67.50  93.12    100 100.00
## n=125     100 100.00 100.00  86.25   80.00   68.75  93.12    100 100.00
## n=150     100 100.00  99.80  95.40   91.60   70.00  95.20    100 100.00
## n=200     100 100.00 100.00  98.40   96.80   76.60  96.20    100 100.00
## n=500     100 100.00 100.00 100.00   99.60   87.20  94.60    100 100.00
## n=1000    100 100.00 100.00 100.00  100.00   98.40  96.20    100 100.00
```

```
option3c.2_fit
```

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      99.8  99.0  98.4  95.8    94.6    91.6  96.0    99.8  99.0
## n=75      99.8  99.0  98.2  96.0    94.6    91.4  96.0    99.8  99.0
## n=100     100.0 100.0  99.6  97.4    96.2    90.2  96.4    100.0  99.6
## n=125     100.0  99.8  98.8  90.2    86.0    66.4  93.0    99.8  99.6
## n=150     100.0  99.8  99.6  92.0    87.8    66.6  94.0    100.0  99.8
## n=200     100.0 100.0  99.8  97.6    96.0    76.2  95.0    100.0  99.8
## n=500     100.0 100.0 100.0 100.0   100.0    88.6  94.6    100.0 100.0
## n=1000    100.0 100.0 100.0 100.0   100.0    97.0  93.6    100.0 100.0
```

```
option3c.3_fit
```

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      100    100  99.8  99.4    99.0    95.2  93.8    100    100
## n=75      100    100  99.6  99.2    99.0    95.4  93.6    100    100
## n=100     100    100  99.4  95.4    93.8    85.0  94.4    100    100
## n=125     100    100  99.6  92.4    87.0    65.2  94.6    100    100
## n=150     100    100  99.2  95.8    90.2    71.2  94.2    100    100
## n=200     100    100  99.8  99.0    97.4    77.2  97.0    100    100
## n=500     100    100 100.0 100.0   100.0    88.6  94.4    100    100
## n=1000    100    100 100.0 100.0   100.0    96.8  93.6    100    100
```

```
option3c.4_fit
```

```
##          CFI90  CFI95  TLI90  TLI95  RMSEA08  RMSEA04  CHI2  SRMR10  SRMR08
## n=50      100 100.0  99.8  98.6    98.2    94.6  94.2    99.8  99.8
## n=75      100  98.8  94.8  83.8    79.8    63.6  94.2    99.8  99.6
## n=100     100  99.8  97.8  90.6    85.6    67.4  96.6    100.0  99.6
## n=125     100  99.8  98.4  91.8    86.6    70.4  93.6    100.0 100.0
## n=150     100 100.0 100.0  94.4    91.4    66.6  93.4    100.0 100.0
## n=200     100 100.0 100.0  97.6    96.0    73.8  95.8    100.0 100.0
## n=500     100 100.0 100.0 100.0   100.0    89.0  94.2    100.0 100.0
## n=1000    100 100.0 100.0 100.0   100.0    96.2  94.0    100.0 100.0
```

In order to overcome this convergence problem: constrains are set to the variances. As variances can never be negative, the code below forces them to be positive. These constraints are in line with other software like, for example, EQS.

```
SRM_pd3c_nonegvar <- '
# Latent variables
FE =~ lambdaMF*MF + lambdaMO*MO + lambdaMY*MY + lambdaFO*FO + lambdaFY*FY + lambdaYO*YO
I.M =~ 1*MF + 1*MO + 1*MY
```

```

I.F =~ 1*MF + 1*FO + 1*FY
I.O =~ 1*MO + 1*FO + 1*YO
I.Y =~ 1*MY + 1*FY + 1*YO
D.MF =~ 1*MF
D.MO =~ 1*MO
D.MY =~ 1*MY
D.FO =~ 1*FO
D.FY =~ 1*FY
D.OY =~ 1*YO

# Variances
FE ~~ VAR.FE*FE
I.M ~~ VAR.I.M*I.M
I.F ~~ VAR.I.F*I.F
I.O ~~ VAR.I.O*I.O
I.Y ~~ VAR.I.Y*I.Y
D.MF ~~ VAR.D.MF*D.MF
D.MO ~~ VAR.D.MO*D.MO
D.MY ~~ VAR.D.MY*D.MY
D.FO ~~ VAR.D.FO*D.FO
D.FY ~~ VAR.D.FY*D.FY
D.OY ~~ VAR.D.OY*D.OY

# Intercepts
FE ~ mean.FE*1
I.M ~ mean.I.M*1
I.F ~ mean.I.F*1
I.O ~ mean.I.O*1
I.Y ~ mean.I.Y*1
D.MF ~ mean.D.MF*1
D.MO ~ mean.D.MO*1
D.MY ~ mean.D.MY*1
D.FO ~ mean.D.FO*1
D.FY ~ mean.D.FY*1
D.OY ~ mean.D.OY*1

# no negative variances allowed.
VAR.FE > 0
VAR.I.M > 0
VAR.I.F > 0
VAR.I.O > 0
VAR.I.Y > 0
VAR.D.MF > 0
VAR.D.MO > 0
VAR.D.MY > 0
VAR.D.FO > 0
VAR.D.FY > 0
VAR.D.OY > 0

# Constraints
mean.I.M + mean.I.F + mean.I.O + mean.I.Y == 0
mean.D.MF + mean.D.MO + mean.D.MY == 0
mean.D.MF + mean.D.FO + mean.D.FY == 0

```

```
mean.D.MY + mean.D.FY + mean.D.OY == 0
mean.D.MO + mean.D.FO + mean.D.OY == 0
```

```
# set constraints on factor loadings FE for identifiability
lambdaMF + lambdaMO + lambdaMY + lambdaFO + lambdaFY + lambdaYO == 6
'
```

```
# How did the fit measures react?
round(option3c.1_fit,2)
```

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      99.8  96.8  92.6  78.4    71.8    61.4  96.2   99.8   97.6
## n=75     100.0  99.0  95.0  87.4    80.2    60.8  94.6  100.0   99.8
## n=100     100.0 100.0  98.8  91.2    85.6    65.4  94.0  100.0  100.0
## n=125     100.0  99.8  99.2  94.8    90.2    65.8  95.2  100.0  100.0
## n=150     100.0 100.0  99.8  96.2    92.0    73.0  95.8  100.0  100.0
## n=200     100.0 100.0  99.8  98.4    95.4    73.4  95.2  100.0  100.0
## n=500     100.0 100.0 100.0 100.0   100.0    91.6  95.8  100.0  100.0
## n=1000    100.0 100.0 100.0 100.0   100.0    98.0  95.8  100.0  100.0
```

option3c.2_fit

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      99.6  96.4  90.8  79.6    76.6    65.8  95.4   99.6   97.4
## n=75     100.0  99.0  96.2  84.6    78.2    61.2  95.4   99.8   99.8
## n=100     100.0 100.0  98.6  90.8    86.2    64.0  95.4  100.0  100.0
## n=125     100.0  99.8  99.8  94.2    90.0    68.2  95.8  100.0  100.0
## n=150     100.0 100.0 100.0  96.6    93.8    72.2  95.6  100.0  100.0
## n=200     100.0 100.0 100.0  97.4    92.4    74.4  91.6  100.0  100.0
## n=500     100.0 100.0 100.0 100.0   100.0    87.2  95.0  100.0  100.0
## n=1000    100.0 100.0 100.0 100.0   100.0    97.0  93.6  100.0  100.0
```

option3c.3_fit

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      99.8  95.4  87.2  75.0    69.2    55.4  93.4   99.4   98.4
## n=75     100.0  99.4  94.4  82.2    77.2    61.8  93.4  100.0   98.8
## n=100     100.0  99.8  99.0  91.0    85.4    65.4  95.2  100.0  100.0
## n=125     100.0 100.0  99.0  94.4    89.2    69.0  95.6  100.0  100.0
## n=150     100.0  99.8  99.8  96.8    91.0    71.6  95.8  100.0  100.0
## n=200     100.0 100.0 100.0  98.4    95.4    75.4  94.8  100.0  100.0
## n=500     100.0 100.0 100.0 100.0   100.0    88.6  93.4  100.0  100.0
## n=1000    100.0 100.0 100.0 100.0   100.0    96.6  93.8  100.0  100.0
```

option3c.4_fit

```
##          CFI90 CFI95 TLI90 TLI95 RMSEA08 RMSEA04 CHI2 SRMR10 SRMR08
## n=50      100  97.2  91.0  80.0    74.6    63.6  95.4   99.8   97.4
## n=75      100  99.2  96.8  84.8    78.4    65.4  96.6  100.0   99.4
## n=100      100  99.6  97.4  89.4    84.2    67.0  93.8  100.0  100.0
## n=125      100  99.8  99.0  93.6    87.2    66.4  93.4  100.0  100.0
## n=150      100 100.0 100.0  96.0    91.6    70.2  95.4  100.0  100.0
## n=200      100 100.0 100.0  98.8    97.4    75.6  96.8  100.0  100.0
## n=500      100 100.0 100.0 100.0   100.0    86.8  94.2  100.0  100.0
## n=1000     100 100.0 100.0 100.0   100.0    97.8  96.0  100.0  100.0
```


How many models converged (%)?

```
conv
##      optie 1 optie 2 optie 3 optie 4
## n=50      100    100    100    100
## n=75      100    100    100    100
## n=100     100    100    100    100
## n=125     100    100    100    100
## n=150     100    100    100    100
## n=200     100    100    100    100
## n=500     100    100    100    100
## n=1000    100    100    100    100
```

This models performs very well. By adding these constraints, the convergence problem got completely solved since now all models converged. Also, the fit indices perform very well with all underlying models. General conclusions and guidelines can be given can be made:

Index	Cut_off	Minimal_sample_size
CFI	0.90	50
	0.95	50
TLI	0.90	75
	0.95	150
RMSEA	0.08	200
	0.04	500 < n < 1000
CHI2(p-value)	0.05	50
SRMR	0.10	50
	0.08	50

General conclusion: The different models

The model that uses the lambda approach for the family component (with all variances constrained to be positive) is highly recommended. First of all, its interpretation of the different components is the most straightforward out of all variations of Kenny's specification. Also, the fit measures perform very well and their performance does not depend on the true underlying model.

Detailed results:

Model 1 only fitted well when in reality all observed scores are of equal importance for the family effect (option 1), but not with the other options.

Method 2:

- Models 2a and 2b perform well, especially if $n \geq 75$
- Model 2c performs badly with the simulated datasets of option 2 when $75 \leq n \leq 500$ (exception: CFI)
- Model 2d shows excellent fit indices for all sample sizes. However, with $n \leq 125$, models do not always converge.

Method 3:

- Models 3a and 3b: all models converged using the default 250 iterations. Model 3a resulted in a lot of negative variances in small samples, though. Also, data simulated under option 2 resulted in a lot of negative variances (if $n \leq 200$). For model 3b, this problem only occurred with the smallest samples. For both models, when $n = 75$, we found good results with the least stringent cut-offs for the fit indices. For the stringent cut-offs: good results were found if $n \geq 100$ or 125.
- Model 3c is the preferred model. The fit indices perform well with all underlying true models and all models converge.