R Notebook: Predictors of Dropout

Contents

clear environment

```
rm(list = ls(all.names = TRUE))
```

Load library

```
#install.packages("lavaan")
#install.packages("xtable")
#install.packages("kableExtra")
#install.packages("semPlot")

library(tidyverse)
library(lavaan)
library(GGally)
library(xtable)
library("kableExtra")
library(knitr)
```

```
library(semPlot)
library(knitr)
library(gridExtra)
library(reshape)
library(pander)
```

Set Kim's Reliability Function

```
my_reliability <- function (items, itemal = TRUE, NA.Delete = TRUE)
{
    if (!all(apply(items, c(1, 2), is.numeric))) {
        items <- apply(items, c(1, 2), as.numeric)
        warning("Data is not numeric. Data has been coerced to be numeric.")
    }
    if (NA.Delete == FALSE) {
        items[is.na(items)] <- 0
        warning("Missing values or NA values are converted to zeros.")
    }
    items <- na.omit(items)
    s <- apply(items, 2, var)
    N <- ncol(items)
    X <- rowSums(items)
    alpha <- (N/(N - 1)) * (1 - sum(s)/var(X))
    if (itemal) {
        alphad <- array(dim = N)
        pbis <- array(dim = N)
        for (i in 1:N) {
            Xd <- rowSums(items[, -i])
            Mean <- sapply(items, mean)
            SD <- sapply(items, sd)
            Min <- sapply(items, min)
            Max <- sapply(items, max)
            Nobs <- sapply(items, length)
            alphad[i] <- ((N - 1)/(N - 2)) * (1 - sum(s[-i])/var(Xd))
            pbis[i] <- cor(items[, i], Xd)
            out <- list(nItem = N, nPerson = nrow(items), alpha = alpha,
                scaleMean = mean(X), scaleSD = sd(X), Alpha.If.Deleted = alphad,
                R.With.Total = pbis, Mean = Mean, SD = SD, Min = Min, Max = Max, Nobs=Nobs)
        }
    }
    else out <- list(nItem = N, nPerson = nrow(items), alpha = alpha,
        scaleMean = mean(X), scaleSD = sd(X))
    class(out) <- "my_reliability"
    out
}

print <- function(x, ...) {
    if (is.numeric(x)) base::print(round(x, digits=3), ...)
    else base::print(x, ...)
}
```

Load Data

```
setwd("T:/Research folders/CCWTG/Analyses/Papers/Dropout/lavaan")

att  <-   read_csv("R_pois.csv")
drop <-   read_csv("R_log.csv")
```

Subset items

```
sem_items <- drop %>%
  select(risk_1:risk_32)

items <- c("risk_1","risk_2","risk_3","risk_4","risk_5","risk_6","risk_7","risk_8","risk_9","risk_10","
```

Frequency table

```
#Count responses: plyr::count
count_all <- map(sem_items, plyr::count)

#COnvert to datframe
count_all <-  as.data.frame(count_all)

#Get necessary info
count_all <- count_all %>%
  mutate(selection = ifelse(risk_1.x == 1, "yes", NA),
         selection = ifelse(risk_1.x == 0, "no", selection)) %>%
  select(selection, ends_with(".freq"))

count_all
```

```
##   selection risk_1.freq risk_2.freq risk_3.freq risk_4.freq risk_5.freq
## 1        no         405         435         329         420         373
## 2       yes          53          23         129          38          85
## 3      <NA>          36          36          36          36          36
##   risk_6.freq risk_7.freq risk_8.freq risk_9.freq risk_10.freq
## 1         355         327         440         424          358
## 2         101         130          17          34          100
## 3          38          37          37          36           36
##   risk_11.freq risk_12.freq risk_13.freq risk_14.freq risk_15.freq
## 1          387          289          393          422          333
## 2           71          169           65           36          124
## 3           36           36           36           36           37
##   risk_16.freq risk_17.freq risk_18.freq risk_19.freq risk_20.freq
## 1          376          407          284          310          235
## 2           82           51          171          147          222
## 3           36           36           39           37           37
##   risk_21.freq risk_22.freq risk_23.freq risk_24.freq risk_25.freq
## 1          213          254          306          435          348
## 2          245          203          152           23          109
```

```
## 3           36           37           36           36           37
##   risk_26.freq risk_27.freq risk_28.freq risk_29.freq risk_30.freq
## 1          390          357          431          453          408
## 2           68          101           27            5           49
## 3           36           36           36           36           37
##   risk_31.freq risk_32.freq
## 1          256          264
## 2          202          194
## 3           36           36
```

Item corrlations

```
cor <- round(cor(sem_items, use = "pairwise.complete.obs") ,2)

upper<-cor
upper[upper.tri(cor)]<-""
upper<-as.data.frame(upper)


kable(upper, "html", full_width = T)
```

risk_1

risk_2

risk_3

risk_4

risk_5

risk_6

risk_7

risk_8

risk_9

risk_10

risk_11

risk_12

risk_13

risk_14

risk_15

risk_16

risk_17

risk_18

risk_19

risk_20

risk_21

risk_22

risk_23

risk_24

risk_25

risk_26

risk_27

risk_28

risk_29

risk_30

risk_31

risk_32

risk_1

1

risk_2

0.2

1

risk_3

0.02

0.19

1

risk_4

0.06

0.11

0.11

1

risk_5

0.04

0.1

0.19

0.08

1

risk_6

0.26

0.05

0.2

0.14

0.43

1

risk_7

0.2

0.05

0.27

0.13

0.31

0.52

1

risk_8

0

0.22

-0.07

0.02

0

0.01

-0.02

1

risk_9

0.05

0.16

-0.01

0.07

0.01

0.01

0.04

0.52

1

risk_10

-0.11

0.07

0.07

0.05

-0.05

-0.08

0.04

0.2

0.23

1

risk_11

-0.08

0.12

0.04

0.05

0

0

0.04

0.2

0.2

0.52

1

risk_12

0.16

0.07

0.08

0.16

0.28

0.32

0.27

-0.08

-0.04

0.02

0.17

1

risk_13

0.07

0.19

0.11

-0.01

0.02

0.07

0.08

0.35

0.24

0.21

0.22

0.09

1

risk_14

0.02

0.04

0.03

0.03

0.03

0.02

0.05

0.11

0.07

0.18

0.17

0.21

0.14

1

risk_15

0.08

0.15

0.07

0.03

0.04

-0.01

0.1

0.11

0.03

0.19

0.16

0.08

0.1

0.17

1

risk_16

-0.01

0.1

0.14

0.07

0

0.07

0.11

0.15

0.06

0.1

0.16

0.09

0.2

0.16

0.13

1

risk_17

0.09

0.11

0.26

0.17

0.13

0.16

0.22

0

-0.02

0.03

0.1

0.13

0.23

0.03

0.08

0.11

1
risk_18
0.14
0.02
0.12
0.08
0.05
0.15
0.24
0.12
0.05
0.17
0.25
0.14
0.08
0.02
0.03
0.03
0.21
1
risk_19
-0.01
0.08
0.09
0.05
0.04
0
-0.07
0.11
0.11
0.08
0.05
0.02
0.11
-0.03
-0.01

0.02

0.02

-0.01

1

risk_20

-0.02

0.04

0.07

0.08

-0.07

-0.01

-0.15

-0.08

0.01

-0.01

0.04

-0.01

0.03

0.02

0.04

0.12

0.02

-0.06

0.17

1

risk_21

-0.07

-0.03

0.05

0.06

0.04

0.07

0

-0.03

-0.07

-0.01

0.04
-0.1
0.05
-0.02
-0.02
-0.02
-0.03
0.06
-0.03
0.14
1
risk_22
-0.08
0.02
0.09
0.05
-0.03
0
-0.09
-0.01
0.05
-0.01
0.01
0.02
0.01
-0.07
0.01
0.03
0.03
-0.06
0.18
0.28
0.16
1
risk_23
-0.04

0.07

0.12

0.12

0.02

0.06

0.1

-0.04

-0.02

-0.04

0.11

0.03

0.02

0

0

0.02

0.09

0.06

-0.05

0.09

0.29

0.13

1

risk_24

0.1

-0.05

-0.03

0

0.04

0.17

0.03

-0.05

0.01

-0.07

-0.04

-0.05

-0.04

0.12

-0.05

-0.03

0.01

0.07

0.06

0.02

-0.03

-0.06

0.01

1

risk_25

-0.12

-0.01

-0.01

0.13

0.06

-0.01

0.03

-0.03

0.04

0.19

0.16

-0.04

0.12

0.03

0.04

-0.05

0.05

-0.05

-0.08

-0.08

0.2

0.02

0.24

-0.01

1
risk_26
-0.02
0.02
0
0.16
-0.01
-0.02
-0.06
0.02
0
0.05
0.08
-0.04
0.04
0.06
-0.01
0
0.05
-0.02
0.05
0.13
0.16
0.11
0.16
0.13
0.27
1
risk_27
-0.06
0.09
0.05
0.16
0.02
0
0.01

0.03

0.03

0.09

0.11

0.01

0.1

0.04

0.08

0.01

0.13

0.02

0.04

0.03

0.11

0.08

0.17

0

0.39

0.39

1

risk_28

-0.03

0.03

0.03

0.09

-0.05

-0.04

0.07

0.1

0.11

0.11

0.15

-0.02

0.11

0.03

0.2

0.03

0.12

0.06

0.09

0.03

0.12

0.07

0.2

-0.02

0.32

0.18

0.27

1

risk_29

-0.04

-0.02

-0.07

0.12

-0.05

-0.06

-0.02

-0.02

-0.03

0.1

0.01

0.01

-0.04

0.05

0.03

-0.05

0.03

-0.08

0.06

-0.02

0.1

0.08

0.06

-0.02

0.14

0.13

0.15

0.15

1

risk_30

-0.04

0.15

0.05

0.16

-0.02

0

0.02

-0.03

0.09

0.04

0.07

-0.04

-0.02

-0.02

0.11

-0.05

0.08

-0.02

0.08

0.12

0.15

0.19

0.03

0.02

0.14

0.33

0.14

0.24

0.1

1

risk_31

-0.01

0.02

0.06

0

-0.08

-0.08

-0.06

-0.03

0.05

0.08

0.02

-0.04

0.04

-0.05

0.15

0.11

-0.01

-0.01

0.11

0.23

0.15

0.24

0.17

0.02

0.12

0.12

0.08

0.17

0.08

0.15

1

risk_32

-0.02

0.11

0.02

0.03

-0.03

-0.04

-0.15

0.07

0.08

0.06

0.05

-0.04

0.09

-0.05

0.04

0.03

-0.02

-0.05

0.15

0.24

0.07

0.38

0.07

-0.1

0.11

0.13

0.17

0.09

0.08

0.13

0.32

1

```
#kable(upper, "html") %>%
#  kable_styling(full_width = T, position = "left") # <- nicer chart but only works whenknit to html.
```

Confirmatory Factor Analyses

Models

```
eco_risk <- "eco_risk  =~ risk_1 + risk_2 + risk_3 + risk_4 + risk_5 + risk_6 + risk_7"
fam_str  <- "fam_str   =~ risk_8 + risk_9 + risk_10 + risk_11 + risk_12 + risk_13 + risk_14 + risk_15 +
peer_dif <- "peer_dif  =~ risk_19 + risk_20"

aca_cha  <- "aca_cha   =~ risk_21 + risk_22 + risk_23"
prob_be  <- "prob_be   =~ risk_25 + risk_26 + risk_27 + risk_28  + risk_30  #+ risk_29"
men_he   <- "men_he    =~ risk_31 + risk_32"
```

Economic Risk

Item Analysis

```
items <- sem_items %>%
    select(risk_1 , risk_2 , risk_3 , risk_4 , risk_5 , risk_6 , risk_7)

variables <- c("risk_1","risk_2","risk_3","risk_4","risk_5","risk_6","risk_7")


ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)

pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))


temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total, ca$Alpha.
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha: 0.62

```
set.alignment('center', row.names = 'center')
pander(temp)
```

Table 1: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_1	456	0.114	0.3182	0	1
risk_2	456	0.05044	0.2191	0	1
risk_3	456	0.2829	0.4509	0	1
risk_4	456	0.08333	0.2767	0	1
risk_5	456	0.1842	0.3881	0	1
risk_6	456	0.2215	0.4157	0	1
risk_7	456	0.2829	0.4509	0	1

Correlation with total	Cronbach's Alpha if item is deleted
0.2065	0.613
0.1921	0.6141
0.296	0.595

Correlation with total	Cronbach's Alpha if item is deleted
0.182	0.6171
0.3678	0.5641
0.542	0.4933
0.4952	0.5096

```
rm(ca)
rm(temp)
rm(pipealpha)
```

```
cfa <- cfa(eco_risk, data = sem_items)
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  56 iterations
##
##                                                   Used       Total
##   Number of observations                            456         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic                63.147
##   Degrees of freedom                                 14
##   P-value (Chi-square)                            0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic               381.559
##   Degrees of freedom                                 21
##   P-value                                         0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                     0.864
##   Tucker-Lewis Index (TLI)                        0.796
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -1007.856
##   Loglikelihood unrestricted model (H1)       -976.282
##
##   Number of free parameters                          14
##   Akaike (AIC)                                2043.712
##   Bayesian (BIC)                              2101.427
##   Sample-size adjusted Bayesian (BIC)         2056.995
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.088
##   90 Percent Confidence Interval          0.066  0.110
##   P-value RMSEA <= 0.05                          0.002
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.058
```

```
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   eco_risk =~
##     risk_1            1.000
##     risk_2            0.305    0.143    2.127    0.033
##     risk_3            1.609    0.397    4.049    0.000
##     risk_4            0.603    0.199    3.026    0.002
##     risk_5            2.202    0.460    4.783    0.000
##     risk_6            3.715    0.737    5.038    0.000
##     risk_7            3.322    0.660    5.034    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_1            0.093    0.006   14.673    0.000
##    .risk_2            0.047    0.003   15.023    0.000
##    .risk_3            0.182    0.013   14.532    0.000
##    .risk_4            0.074    0.005   14.906    0.000
##    .risk_5            0.112    0.008   13.239    0.000
##    .risk_6            0.063    0.011    5.934    0.000
##    .risk_7            0.115    0.011   10.355    0.000
##     eco_risk          0.008    0.003    2.641    0.008
```

```
rm(cfa)
```

Family structure

Item Analysis

```
items <- sem_items %>%
    select(risk_8 , risk_9 , risk_10 , risk_11 , risk_12 , risk_13 , risk_14 , risk_15 , risk_16 , risk

variables <- c("risk_8","risk_9","risk_10","risk_11","risk_12","risk_13","risk_14","risk_15","risk_16",


ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)

pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))


temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total, ca$Alpha.
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha: 0.59

Item	N	Average	Standard Deviation	Minimum	Maximum

Table 3: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_8	453	0.03091	0.1733	0	1
risk_9	453	0.07064	0.2565	0	1
risk_10	453	0.2163	0.4122	0	1
risk_11	453	0.1523	0.3597	0	1
risk_12	453	0.3709	0.4836	0	1
risk_13	453	0.1369	0.3441	0	1
risk_14	453	0.07506	0.2638	0	1
risk_15	453	0.2671	0.4429	0	1
risk_16	453	0.1744	0.3799	0	1
risk_17	453	0.1126	0.3164	0	1
risk_18	453	0.3753	0.4847	0	1

Correlation with total	Cronbach's Alpha if item is deleted
0.2763	0.5753
0.1936	0.5813
0.3715	0.5391
0.472	0.5189
0.2081	0.5868
0.337	0.5515
0.2403	0.5738
0.1974	0.5862
0.2087	0.5798
0.2291	0.5746
0.2361	0.5786

```
cfa <- cfa(fam_str, data = sem_items)
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  93 iterations
##
##                                                   Used       Total
##   Number of observations                            453         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic               250.687
##   Degrees of freedom                                 44
##   P-value (Chi-square)                            0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic               572.744
##   Degrees of freedom                                 55
##   P-value                                         0.000
##
## User model versus baseline model:
##
```

```
##   Comparative Fit Index (CFI)                    0.601
##   Tucker-Lewis Index (TLI)                       0.501
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -1556.172
##   Loglikelihood unrestricted model (H1)      -1430.828
##
##   Number of free parameters                         22
##   Akaike (AIC)                                3156.344
##   Bayesian (BIC)                              3246.893
##   Sample-size adjusted Bayesian (BIC)         3177.073
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.102
##   90 Percent Confidence Interval          0.090  0.114
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.080
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   fam_str =~
##     risk_8            1.000
##     risk_9            1.369    0.281    4.864    0.000
##     risk_10           3.919    0.629    6.230    0.000
##     risk_11           3.680    0.584    6.302    0.000
##     risk_12           1.402    0.449    3.126    0.002
##     risk_13           2.076    0.398    5.211    0.000
##     risk_14           0.935    0.255    3.674    0.000
##     risk_15           1.587    0.429    3.704    0.000
##     risk_16           1.280    0.362    3.530    0.000
##     risk_17           0.894    0.292    3.058    0.002
##     risk_18           2.260    0.506    4.469    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_8            0.026    0.002   13.941    0.000
##    .risk_9            0.057    0.004   14.128    0.000
##    .risk_10           0.102    0.010   10.594    0.000
##    .risk_11           0.070    0.007    9.386    0.000
##    .risk_12           0.225    0.015   14.805    0.000
##    .risk_13           0.099    0.007   13.816    0.000
##    .risk_14           0.066    0.004   14.676    0.000
##    .risk_15           0.185    0.013   14.667    0.000
##    .risk_16           0.137    0.009   14.714    0.000
```

```
##    .risk_17          0.096    0.007   14.818    0.000
##    .risk_18          0.212    0.015   14.371    0.000
##     fam_str          0.004    0.001    3.513    0.000
```

```
rm(cfa)
```

Peer difficulties

Item Analysis

```
items <- sem_items %>%
    select(risk_19 , risk_20)

variables <- c("risk_19","risk_20")


ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)

pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))


temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total)
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha NA (2-items scale)

Table 5: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_19	456	0.3202	0.4671	0	1
risk_20	456	0.4868	0.5004	0	1

Correlation with total
0.1685
0.1685

```
cfa <- cfa(peer_dif, data = sem_items)
```

```
## Warning in lav_model_vcov(lavmodel = lavmodel, lavsamplestats = lavsamplestats, : lavaan WARNING: co
##   lavaan NOTE: this may be a symptom that the model is not identified.
```

```
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  12 iterations
##
##                                                   Used       Total
##   Number of observations                           456         494
##
##   Estimator                                         ML
##   Minimum Function Test Statistic                   NA
##   Degrees of freedom                                -1
```

```
##   Minimum Function Value               0.0000000000000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                        NA
##   Tucker-Lewis Index (TLI)                           NA
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)               -623.613
##   Loglikelihood unrestricted model (H1)       -623.613
##
##   Number of free parameters                          4
##   Akaike (AIC)                                1255.226
##   Bayesian (BIC)                              1271.716
##   Sample-size adjusted Bayesian (BIC)         1259.022
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                              NA
##   90 Percent Confidence Interval             NA      NA
##   P-value RMSEA <= 0.05                              NA
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                            0.000
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   peer_dif =~
##     risk_19           1.000
##     risk_20           0.256       NA
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_19           0.064       NA
##    .risk_20           0.240       NA
##     peer_dif          0.154       NA
```

```
rm(cfa)
```

Academic Challenges

Item Analysis

```
items <- sem_items %>%
    select(risk_21 , risk_22 , risk_23, risk_24)

variables <- c("risk_21","risk_22","risk_23","risk_24")
```

```
ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)

pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))


temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total, ca$Alpha.
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha: 0.34

Table 7: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_21	457	0.5339	0.4994	0	1
risk_22	457	0.4442	0.4974	0	1
risk_23	457	0.3326	0.4717	0	1
risk_24	457	0.05033	0.2189	0	1

Correlation with total	Cronbach's Alpha if item is deleted
0.2816	0.1237
0.1518	0.3103
0.2737	0.1433
-0.04165	0.416

```
cfa <- cfa(aca_cha, data = sem_items)
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  23 iterations
##
##                                                   Used       Total
##   Number of observations                            457         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic                 0.000
##   Degrees of freedom                                  0
##
## Model test baseline model:
##
##   Minimum Function Test Statistic                56.524
##   Degrees of freedom                                  3
##   P-value                                         0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                     1.000
##   Tucker-Lewis Index (TLI)                        1.000
##
## Loglikelihood and Information Criteria:
##
```

```
##   Loglikelihood user model (H0)                -935.716
##   Loglikelihood unrestricted model (H1)        -935.716
##
##   Number of free parameters                           6
##   Akaike (AIC)                                1883.432
##   Bayesian (BIC)                              1908.181
##   Sample-size adjusted Bayesian (BIC)         1889.138
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.000
##   90 Percent Confidence Interval          0.000  0.000
##   P-value RMSEA <= 0.05                             NA
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.000
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   aca_cha =~
##     risk_21           1.000
##     risk_22           0.423    0.161    2.631    0.009
##     risk_23           0.765    0.308    2.484    0.013
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_21           0.158    0.038    4.125    0.000
##    .risk_22           0.231    0.017   13.875    0.000
##    .risk_23           0.169    0.024    6.949    0.000
##     aca_cha           0.091    0.039    2.338    0.019
```

```
rm(cfa)
```

Problem Behavior

Item Analysis

```
items <- sem_items %>%
    select(risk_25:risk_30)

variables <- c("risk_25","risk_26","risk_27","risk_28", "risk_29", "risk_30")


ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)

pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))
```

```
temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total, ca$Alpha.
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha: 0.63

Table 9: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_25	456	0.239	0.427	0	1
risk_26	456	0.1491	0.3566	0	1
risk_27	456	0.2215	0.4157	0	1
risk_28	456	0.05921	0.2363	0	1
risk_29	456	0.01096	0.1043	0	1
risk_30	456	0.1075	0.31	0	1

Correlation with total	Cronbach's Alpha if item is deleted
0.4278	0.5627
0.4533	0.5475
0.469	0.5394
0.3822	0.588
0.2056	0.6398
0.2918	0.6118

```
cfa <- cfa(prob_be, data = sem_items)
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  31 iterations
##
##                                                   Used       Total
##   Number of observations                            456         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic                41.923
##   Degrees of freedom                                  5
##   P-value (Chi-square)                            0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic               293.209
##   Degrees of freedom                                 10
##   P-value                                         0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                     0.870
##   Tucker-Lewis Index (TLI)                        0.739
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)                -656.581
```

```
##   Loglikelihood unrestricted model (H1)       -635.620
##
##   Number of free parameters                         10
##   Akaike (AIC)                                1333.162
##   Bayesian (BIC)                              1374.387
##   Sample-size adjusted Bayesian (BIC)         1342.650
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.127
##   90 Percent Confidence Interval          0.093  0.164
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.057
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   prob_be =~
##     risk_25           1.000
##     risk_26           0.823    0.110    7.472    0.000
##     risk_27           1.090    0.141    7.748    0.000
##     risk_28           0.442    0.067    6.623    0.000
##     risk_30           0.458    0.082    5.578    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_25           0.123    0.011   11.358    0.000
##    .risk_26           0.087    0.008   11.523    0.000
##    .risk_27           0.102    0.010    9.771    0.000
##    .risk_28           0.044    0.003   13.114    0.000
##    .risk_30           0.084    0.006   13.984    0.000
##     prob_be           0.059    0.011    5.176    0.000
```

```
rm(cfa)
```

Mental Health Challenges

Item Analysis

```
items <- sem_items %>%
    select( risk_31, risk_32)

variables <- c("risk_31","risk_32")


ca <- my_reliability(items, itemal=TRUE, NA.Delete=TRUE)
```

```
pipealpha <- paste(formatC(round(ca$alpha, digits=2), 2, format = "f"))


temp <- cbind.data.frame(variables, ca$Nobs, ca$Mean, ca$SD, ca$Min, ca$Max, ca$R.With.Total)
colnames(temp) <- c("Item", "N", "Average", "Standard Deviation", "Minimum", "Maximum", "Correlation wi
rownames(temp) <- NULL
```

Cronbach's Alpha NA (2-items scale)

Table 11: Table continues below

Item	N	Average	Standard Deviation	Minimum	Maximum
risk_31	458	0.441	0.4971	0	1
risk_32	458	0.4236	0.4947	0	1

Correlation with total
0.3154
0.3154

```
cfa <- cfa(men_he, data = sem_items)
```

```
## Warning in lav_model_vcov(lavmodel = lavmodel, lavsamplestats = lavsamplestats, : lavaan WARNING: cou
##   lavaan NOTE: this may be a symptom that the model is not identified.
```

```
summary(cfa, fit.measures = TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after  11 iterations
##
##                                                   Used       Total
##   Number of observations                            458         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic                    NA
##   Degrees of freedom                                 -1
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                        NA
##   Tucker-Lewis Index (TLI)                           NA
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)               -632.217
##   Loglikelihood unrestricted model (H1)       -632.217
##
##   Number of free parameters                          4
##   Akaike (AIC)                               1272.433
##   Bayesian (BIC)                             1288.941
##   Sample-size adjusted Bayesian (BIC)        1276.246
##
## Root Mean Square Error of Approximation:
##
```

```
##   RMSEA                                           NA
##   90 Percent Confidence Interval          NA      NA
##   P-value RMSEA <= 0.05                           NA
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                         0.000
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   men_he =~
##     risk_31           1.000
##     risk_32           0.395       NA
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_31           0.051       NA
##    .risk_32           0.214       NA
##     men_he            0.196       NA
```

```
rm(cfa)
```

SEM

Models

Current problems with model convergence when idnetifying the overall environmental & individual risk factors scale. I will also take all items loaded onto the environmental & individual risk factors scale. This will be model 2.

```
model1 <- "
       #Environmental risk factors
          eco_risk  =~ risk_1 + risk_2 + risk_3 + risk_4 + risk_5 + risk_6 + risk_7
          fam_str   =~ risk_8 + risk_9 + risk_10 + risk_11 + risk_12 + risk_13 + risk_14 + risk_15 + ris
          peer_dif  =~ risk_19 + risk_20

        #Individual Risk factors
          aca_cha   =~ risk_21 + risk_22 + risk_23                        #+risk_24
          prob_be   =~ risk_25 + risk_26 + risk_27 + risk_28  + risk_30   #+ risk_29
          men_he    =~ risk_31 + risk_32

        #env_risk =~ eco_risk + fam_str + peer_dif
        #ind_risk =~ aca_cha + prob_be + men_he

        #risk =~ env_risk + ind_risk
"

model2 <- "
```

```
       #Environmental risk factors
          env_risk =~ risk_1 + risk_2 + risk_3 + risk_4 + risk_5 + risk_6 + risk_7 +
          risk_8 + risk_9 + risk_10 + risk_11 + risk_12 + risk_13 + risk_14 + risk_15 + risk_16 + risk_
          risk_19 + risk_20


        #Individual Risk factors
         ind_risk   =~ risk_21 + risk_22 + risk_23 +                        #+risk_24
          risk_25 + risk_26 + risk_27 + risk_28  + risk_30 +   #+ risk_29
          risk_31 + risk_32

       # risk =~ env_risk + ind_risk
"
```

Fit Model 1

```
fit <-sem(model1, sem_items)
summary(fit, fit.measures =TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after 272 iterations
##
##                                                   Used       Total
##   Number of observations                            450         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic              1053.146
##   Degrees of freedom                                390
##   P-value (Chi-square)                            0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic              2235.313
##   Degrees of freedom                                435
##   P-value                                         0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                     0.632
##   Tucker-Lewis Index (TLI)                        0.589
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -5241.215
##   Loglikelihood unrestricted model (H1)      -4714.642
##
##   Number of free parameters                         75
##   Akaike (AIC)                               10632.431
##   Bayesian (BIC)                             10940.624
##   Sample-size adjusted Bayesian (BIC)        10702.603
##
## Root Mean Square Error of Approximation:
##
```

```
##   RMSEA                                          0.061
##   90 Percent Confidence Interval          0.057  0.066
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.071
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                   Estimate  Std.Err  z-value  P(>|z|)
##   eco_risk =~
##     risk_1           1.000
##     risk_2           0.384    0.152    2.528    0.011
##     risk_3           1.735    0.436    3.981    0.000
##     risk_4           0.681    0.215    3.174    0.002
##     risk_5           2.258    0.495    4.565    0.000
##     risk_6           3.679    0.761    4.832    0.000
##     risk_7           3.572    0.741    4.819    0.000
##   fam_str =~
##     risk_8           1.000
##     risk_9           1.402    0.288    4.864    0.000
##     risk_10          3.777    0.619    6.102    0.000
##     risk_11          3.581    0.576    6.215    0.000
##     risk_12          1.564    0.462    3.383    0.001
##     risk_13          2.225    0.418    5.324    0.000
##     risk_14          0.979    0.261    3.752    0.000
##     risk_15          1.734    0.443    3.915    0.000
##     risk_16          1.362    0.372    3.657    0.000
##     risk_17          1.090    0.307    3.545    0.000
##     risk_18          2.279    0.513    4.445    0.000
##   peer_dif =~
##     risk_19          1.000
##     risk_20          2.292    0.684    3.353    0.001
##   aca_cha =~
##     risk_21          1.000
##     risk_22          1.440    0.275    5.227    0.000
##     risk_23          0.926    0.208    4.449    0.000
##   prob_be =~
##     risk_25          1.000
##     risk_26          0.772    0.103    7.462    0.000
##     risk_27          1.017    0.129    7.912    0.000
##     risk_28          0.452    0.065    6.956    0.000
##     risk_30          0.488    0.082    5.976    0.000
##   men_he =~
##     risk_31          1.000
##     risk_32          1.074    0.155    6.942    0.000
##
## Covariances:
##                   Estimate  Std.Err  z-value  P(>|z|)
```

```
##   eco_risk ~~
##     fam_str           0.001    0.000    2.399    0.016
##     peer_dif         -0.001    0.001   -1.123    0.262
##     aca_cha           0.001    0.001    0.822    0.411
##     prob_be           0.000    0.001    0.270    0.787
##     men_he           -0.003    0.002   -1.733    0.083
##   fam_str ~~
##     peer_dif          0.001    0.001    1.351    0.177
##     aca_cha           0.001    0.001    0.766    0.444
##     prob_be           0.005    0.001    3.467    0.001
##     men_he            0.002    0.001    1.581    0.114
##   peer_dif ~~
##     aca_cha           0.016    0.006    2.912    0.004
##     prob_be           0.004    0.003    1.390    0.164
##     men_he            0.025    0.008    3.199    0.001
##   aca_cha ~~
##     prob_be           0.024    0.006    4.188    0.000
##     men_he            0.042    0.009    4.772    0.000
##   prob_be ~~
##     men_he            0.030    0.007    4.509    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_1            0.095    0.006   14.591    0.000
##    .risk_2            0.043    0.003   14.869    0.000
##    .risk_3            0.182    0.013   14.354    0.000
##    .risk_4            0.070    0.005   14.744    0.000
##    .risk_5            0.114    0.009   13.192    0.000
##    .risk_6            0.073    0.010    7.246    0.000
##    .risk_7            0.109    0.011    9.721    0.000
##    .risk_8            0.026    0.002   13.930    0.000
##    .risk_9            0.058    0.004   14.058    0.000
##    .risk_10           0.108    0.010   11.238    0.000
##    .risk_11           0.073    0.007   10.055    0.000
##    .risk_12           0.222    0.015   14.698    0.000
##    .risk_13           0.097    0.007   13.590    0.000
##    .risk_14           0.066    0.005   14.600    0.000
##    .risk_15           0.183    0.013   14.548    0.000
##    .risk_16           0.137    0.009   14.628    0.000
##    .risk_17           0.095    0.007   14.658    0.000
##    .risk_18           0.212    0.015   14.324    0.000
##    .risk_19           0.202    0.014   13.929    0.000
##    .risk_20           0.165    0.030    5.462    0.000
##    .risk_21           0.217    0.016   13.591    0.000
##    .risk_22           0.181    0.017   10.674    0.000
##    .risk_23           0.193    0.014   13.650    0.000
##    .risk_25           0.122    0.011   11.514    0.000
##    .risk_26           0.091    0.007   12.218    0.000
##    .risk_27           0.111    0.010   11.015    0.000
##    .risk_28           0.042    0.003   12.961    0.000
##    .risk_30           0.083    0.006   13.792    0.000
##    .risk_31           0.174    0.016   11.102    0.000
##    .risk_32           0.160    0.016    9.897    0.000
##     eco_risk          0.007    0.003    2.516    0.012
```

```
##     fam_str           0.004    0.001    3.467    0.001
##     peer_dif          0.016    0.008    2.099    0.036
##     aca_cha           0.032    0.011    3.034    0.002
##     prob_be           0.060    0.011    5.325    0.000
##     men_he            0.072    0.016    4.593    0.000
```

Fit model 2

```
fit2 <-sem(model2, sem_items)
summary(fit2, fit.measures =TRUE)
```

```
## lavaan (0.5-23.1097) converged normally after 155 iterations
##
##                                                   Used       Total
##   Number of observations                            450         494
##
##   Estimator                                          ML
##   Minimum Function Test Statistic              1372.673
##   Degrees of freedom                                404
##   P-value (Chi-square)                            0.000
##
## Model test baseline model:
##
##   Minimum Function Test Statistic              2235.313
##   Degrees of freedom                                435
##   P-value                                         0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                     0.462
##   Tucker-Lewis Index (TLI)                        0.421
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)               -5400.979
##   Loglikelihood unrestricted model (H1)       -4714.642
##
##   Number of free parameters                          61
##   Akaike (AIC)                                10923.958
##   Bayesian (BIC)                              11174.622
##   Sample-size adjusted Bayesian (BIC)         10981.031
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                           0.073
##   90 Percent Confidence Interval          0.069  0.077
##   P-value RMSEA <= 0.05                           0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                            0.082
##
## Parameter Estimates:
```

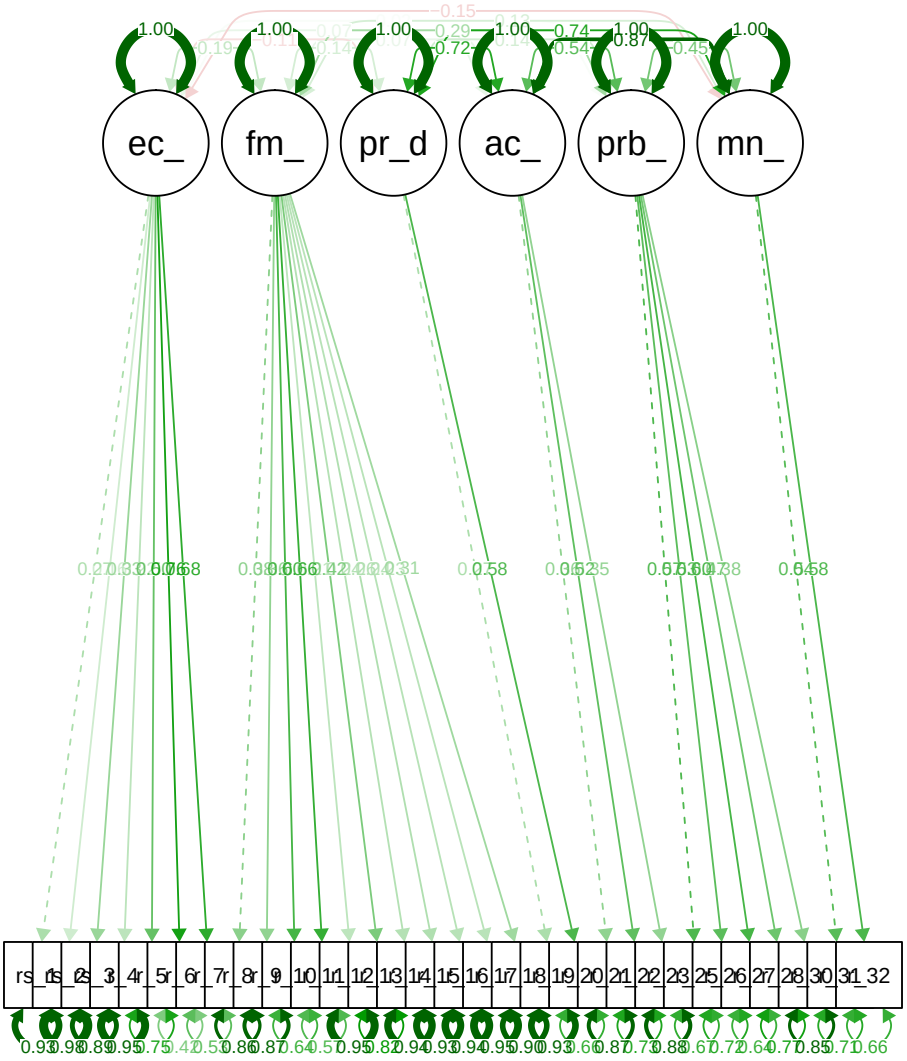
```
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   env_risk =~
##     risk_1            1.000
##     risk_2            0.543    0.165    3.288    0.001
##     risk_3            1.943    0.457    4.249    0.000
##     risk_4            0.810    0.226    3.587    0.000
##     risk_5            2.031    0.449    4.529    0.000
##     risk_6            3.052    0.626    4.875    0.000
##     risk_7            3.392    0.693    4.892    0.000
##     risk_8            0.213    0.115    1.858    0.063
##     risk_9            0.352    0.172    2.042    0.041
##     risk_10           0.586    0.277    2.111    0.035
##     risk_11           0.844    0.272    3.103    0.002
##     risk_12           2.493    0.552    4.517    0.000
##     risk_13           0.980    0.281    3.488    0.000
##     risk_14           0.465    0.185    2.507    0.012
##     risk_15           0.819    0.314    2.610    0.009
##     risk_16           0.869    0.286    3.038    0.002
##     risk_17           1.346    0.318    4.229    0.000
##     risk_18           1.766    0.445    3.972    0.000
##     risk_19           0.233    0.293    0.796    0.426
##     risk_20          -0.218    0.313   -0.697    0.486
##   ind_risk =~
##     risk_21           1.000
##     risk_22           0.855    0.207    4.137    0.000
##     risk_23           1.030    0.217    4.740    0.000
##     risk_25           1.315    0.242    5.440    0.000
##     risk_26           1.073    0.199    5.402    0.000
##     risk_27           1.301    0.238    5.467    0.000
##     risk_28           0.617    0.120    5.154    0.000
##     risk_30           0.717    0.148    4.852    0.000
##     risk_31           1.025    0.223    4.597    0.000
##     risk_32           1.011    0.221    4.576    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   env_risk ~~
##     ind_risk          0.001    0.001    1.217    0.224
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .risk_1            0.094    0.006   14.577    0.000
##    .risk_2            0.042    0.003   14.721    0.000
##    .risk_3            0.175    0.012   14.137    0.000
##    .risk_4            0.069    0.005   14.618    0.000
##    .risk_5            0.120    0.009   13.614    0.000
##    .risk_6            0.102    0.009   11.273    0.000
##    .risk_7            0.115    0.011   10.938    0.000
##    .risk_8            0.030    0.002   14.939    0.000
```

```
##    .risk_9           0.065    0.004   14.924    0.000
##    .risk_10          0.166    0.011   14.918    0.000
##    .risk_11          0.123    0.008   14.769    0.000
##    .risk_12          0.185    0.014   13.645    0.000
##    .risk_13          0.111    0.008   14.656    0.000
##    .risk_14          0.068    0.005   14.874    0.000
##    .risk_15          0.190    0.013   14.860    0.000
##    .risk_16          0.139    0.009   14.783    0.000
##    .risk_17          0.086    0.006   14.161    0.000
##    .risk_18          0.210    0.015   14.405    0.000
##    .risk_19          0.217    0.014   14.990    0.000
##    .risk_20          0.250    0.017   14.992    0.000
##    .risk_21          0.219    0.016   14.117    0.000
##    .risk_22          0.225    0.016   14.373    0.000
##    .risk_23          0.189    0.014   13.912    0.000
##    .risk_25          0.131    0.011   12.425    0.000
##    .risk_26          0.093    0.007   12.576    0.000
##    .risk_27          0.123    0.010   12.308    0.000
##    .risk_28          0.043    0.003   13.284    0.000
##    .risk_30          0.082    0.006   13.780    0.000
##    .risk_31          0.215    0.015   14.053    0.000
##    .risk_32          0.213    0.015   14.072    0.000
##     env_risk         0.008    0.003    2.567    0.010
##     ind_risk         0.030    0.009    3.142    0.002
```

Diagrams

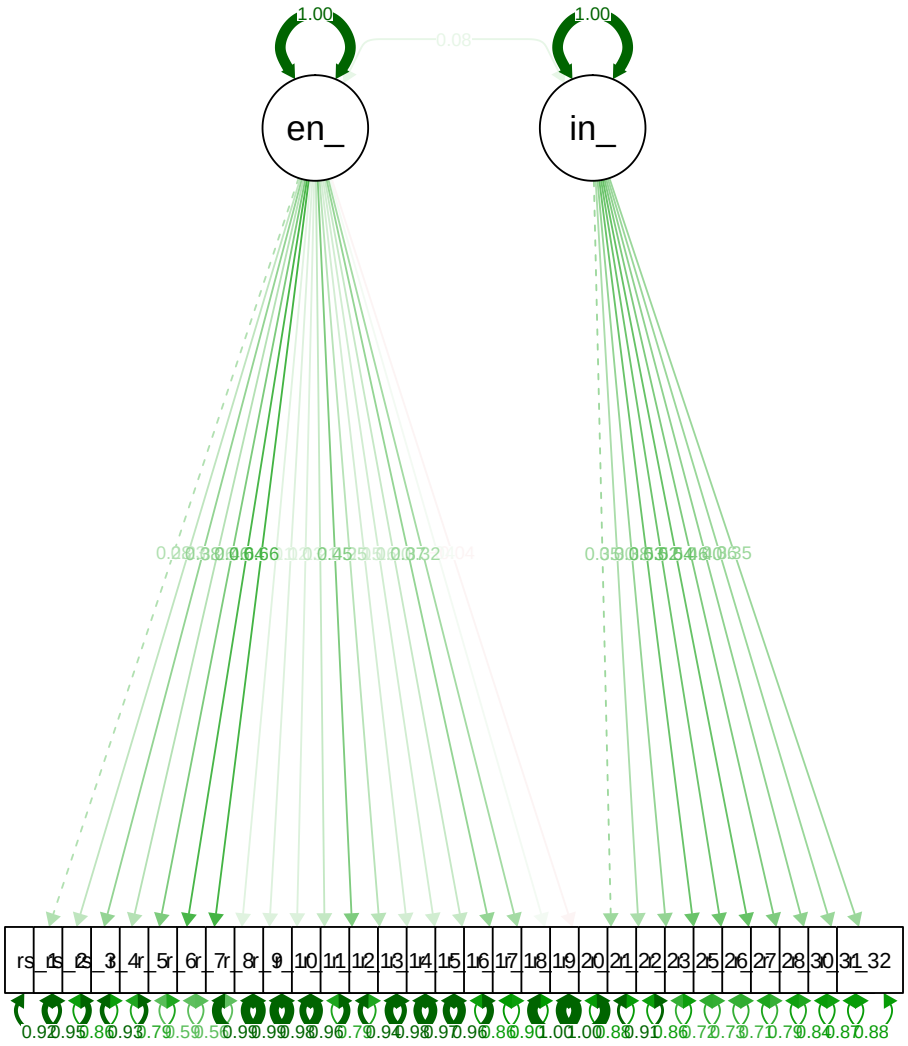
Model 1

```
semPaths(fit, "std", title = FALSE, curvePivot = TRUE)
```



##Model 2

```
semPaths(fit2, "std", title = FALSE, curvePivot = TRUE)
```


1.00
0.08
1.00
en_
in_