Latent variable modeling: A practical guide

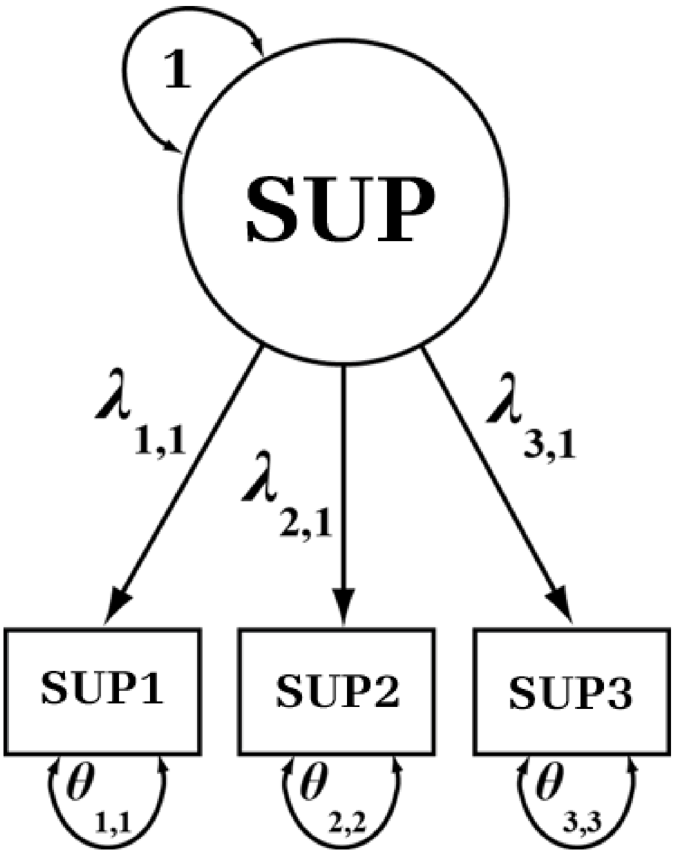
Applied Examples

Waylon Howard

Webinar, March 04, 2025

1.1-CFA-Fixed-Factor.inp

```
DATA: FILE = mydata.dat;
VARIABLE:
NAMES = SUP1 SUP2 SUP3;
MODEL:
SUP by SUP1*
          SUP2
          SUP3;
SUP@1;
OUTPUT: TECH1;
```

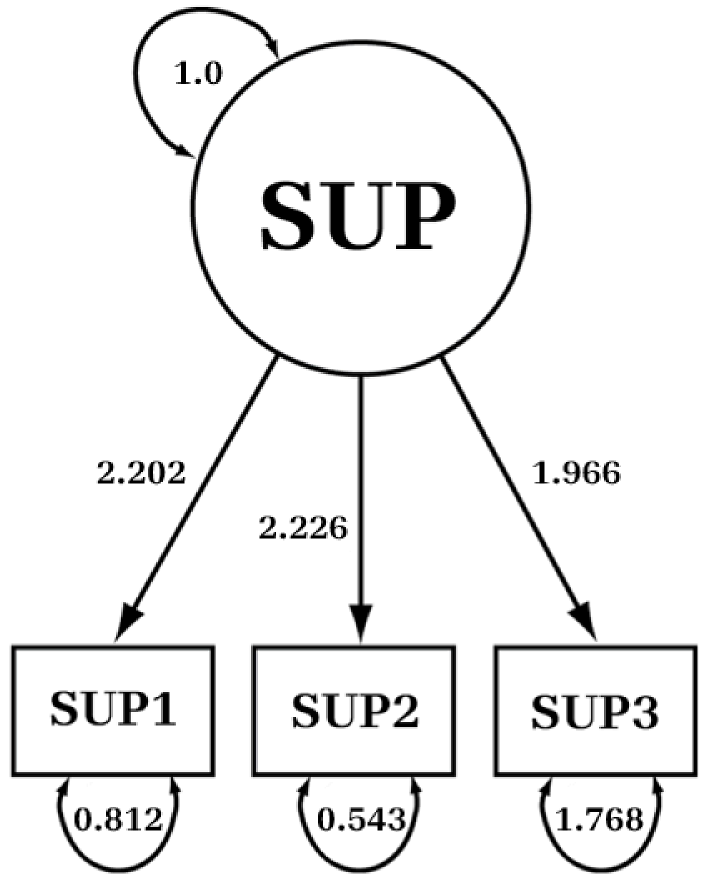


1.1-cfa-fixed-factor.out

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 SUP      BY
    SUP1               2.202      0.155     14.246      0.000
    SUP2               2.226      0.150     14.879      0.000
    SUP3               1.966      0.164     11.990      0.000

 Intercepts
    SUP1               3.287      0.199     16.522      0.000
    SUP2               2.990      0.196     15.239      0.000
    SUP3               3.322      0.198     16.739      0.000

 Variances
    SUP                1.000      0.000    999.000    999.000

 Residual Variances
    SUP1               0.812      0.183      4.429      0.000
    SUP2               0.543      0.171      3.176      0.001
    SUP3               1.768      0.243      7.265      0.000
```

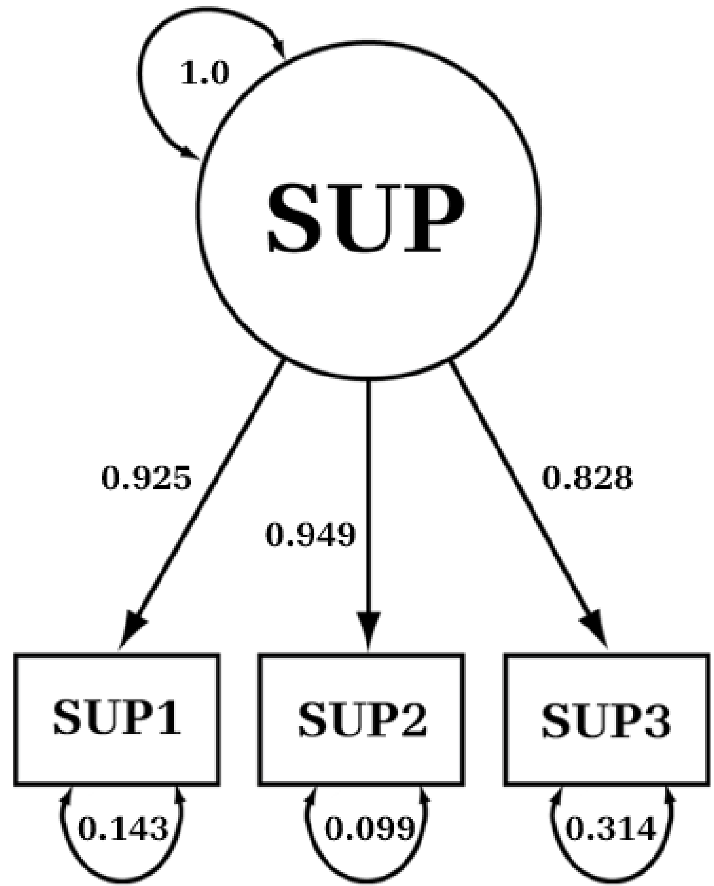
	Estimate	S.E.	Est./S.E.	P-Value
SUP BY				
SUP1	2.202	0.155	14.246	0.000
SUP2	2.226	0.150	14.879	0.000
SUP3	1.966	0.164	11.990	0.000
Intercepts				
SUP1	3.287	0.199	16.522	0.000
SUP2	2.990	0.196	15.239	0.000
SUP3	3.322	0.198	16.739	0.000
Variances				
SUP	1.000	0.000	999.000	999.000
Resid Var				
SUP1	0.812	0.183	4.429	0.000
SUP2	0.543	0.171	3.176	0.001
SUP3	1.768	0.243	7.265	0.000

Unstandardized



STDYX Standardization

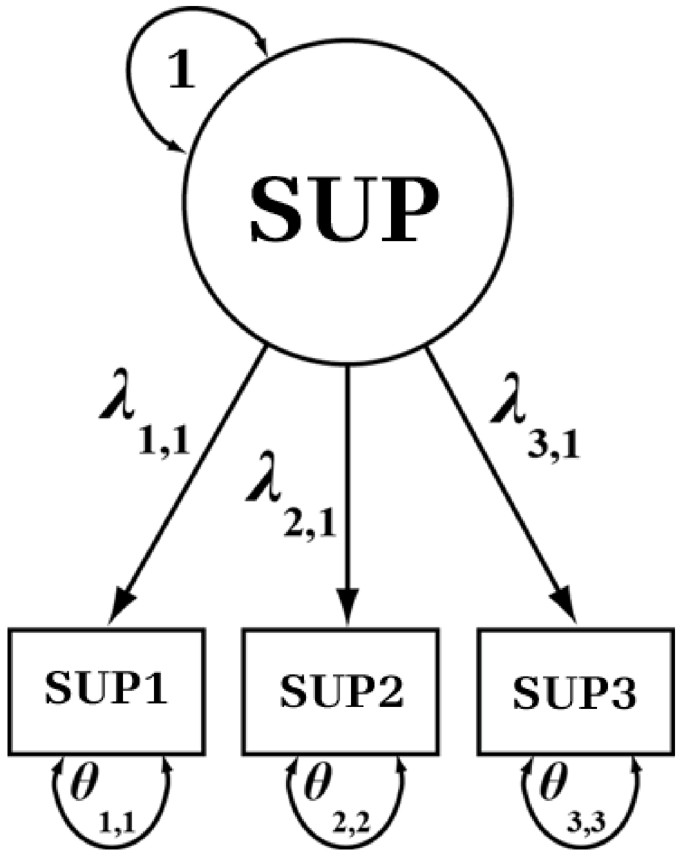
	Estimate	S.E.	Est./S.E.	P-Value
SUP BY				
SUP1	0.925	0.019	48.366	0.000
SUP2	0.949	0.017	54.928	0.000
SUP3	0.828	0.029	28.129	0.000
Intercepts				
SUP1	1.382	0.117	11.818	0.000
SUP2	1.275	0.113	11.318	0.000
SUP3	1.400	0.118	11.897	0.000
Variances				
SUP	1.000	0.000	999.000	999.000
Resid Var				
SUP1	0.143	0.035	4.050	0.000
SUP2	0.099	0.033	3.011	0.003
SUP3	0.314	0.049	6.435	0.000



Perceived Social Support (latent SUP) accounts for **85.6%** $(0.925^2 = 0.856)$ of the variance in the indicator SUP1. Also, 0.856 + 0.143 = 1.00

cfa-examples.R (Lavaan)

```
library(lavaan)
m1 ← '
SUP =~ NA*SUP1 + SUP2 + SUP3
SUP ~~ 1*SUP
'
fit1 ← cfa(m1, data=mydata,
std.lv=T)
summary(fit1, standardized=T,
fit.measures=T, rsquare=T)
```

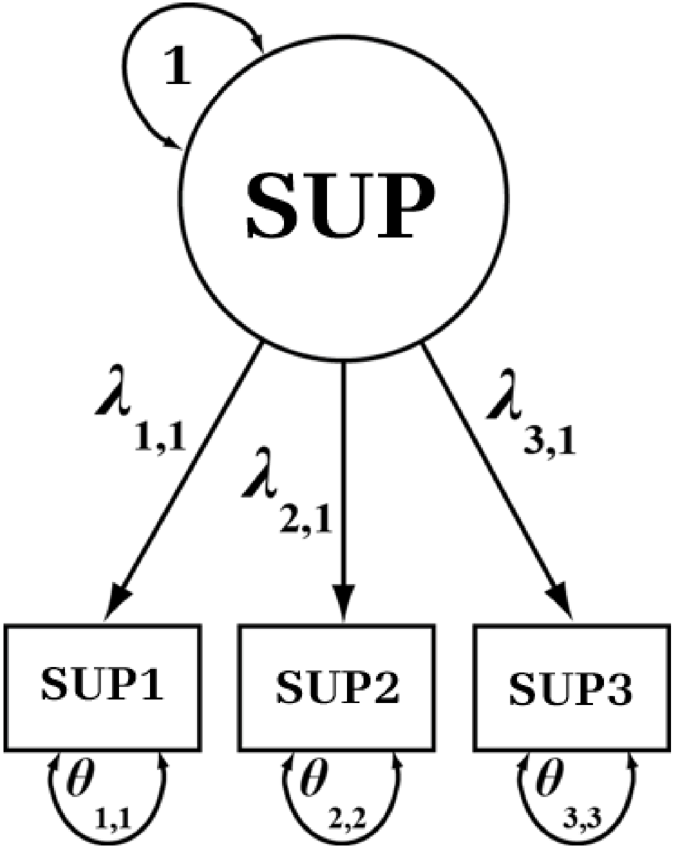


Sample CFA Lavaan Estimates

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  SUP =~
    SUP1              2.211    0.155   14.262    0.000    2.211    0.927
    SUP2              2.230    0.150   14.821    0.000    2.230    0.949
    SUP3              1.978    0.164   12.039    0.000    1.978    0.832

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    SUP               1.000                               1.000    1.000
   .SUP1              0.797    0.180    4.419    0.000    0.797    0.140
   .SUP2              0.554    0.170    3.268    0.001    0.554    0.100
   .SUP3              1.739    0.240    7.240    0.000    1.739    0.308

R-Square:
                   Estimate
    SUP1              0.860
    SUP2              0.900
    SUP3              0.692
```

1.1-CFA-Fixed-Factor.sas (Proc Calis)

```
proc calis data=mydata method=ml;

path SUP → SUP1 SUP2 SUP3 = ly1 - ly3;

pvar SUP = 1,

SUP1 SUP2 SUP3 = te1 - te3;

run;
```



Sample CFA Proc Calis Estimates

The SAS System

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

PATH List

Path			Parameter	Estimate	Standard Error	t Value	Pr > |t|
SUP	===>	SUP1	ly1	2.21910	0.15615	14.2114	<.0001
SUP	===>	SUP2	ly2	2.23813	0.15155	14.7684	<.0001
SUP	===>	SUP3	ly3	1.98466	0.16544	11.9962	<.0001

Variance Parameters

Variance Type	Variable	Parameter	Estimate	Standard Error	t Value	Pr > |t|
Exogenous	SUP		1.00000			
Error	SUP1	theta1	0.80308	0.18240	4.4030	<.0001
	SUP2	theta2	0.55811	0.17139	3.2563	0.0011
	SUP3	theta3	1.75168	0.24281	7.2143	<.0001

Squared Multiple Correlations

Variable	Error Variance	Total Variance	R-Square
SUP1	0.80308	5.72750	0.8598
SUP2	0.55811	5.56733	0.8998
SUP3	1.75168	5.69054	0.6922

The SAS System

The CALIS Procedure
Covariance Structure Analysis: Maximum Likelihood Estimation

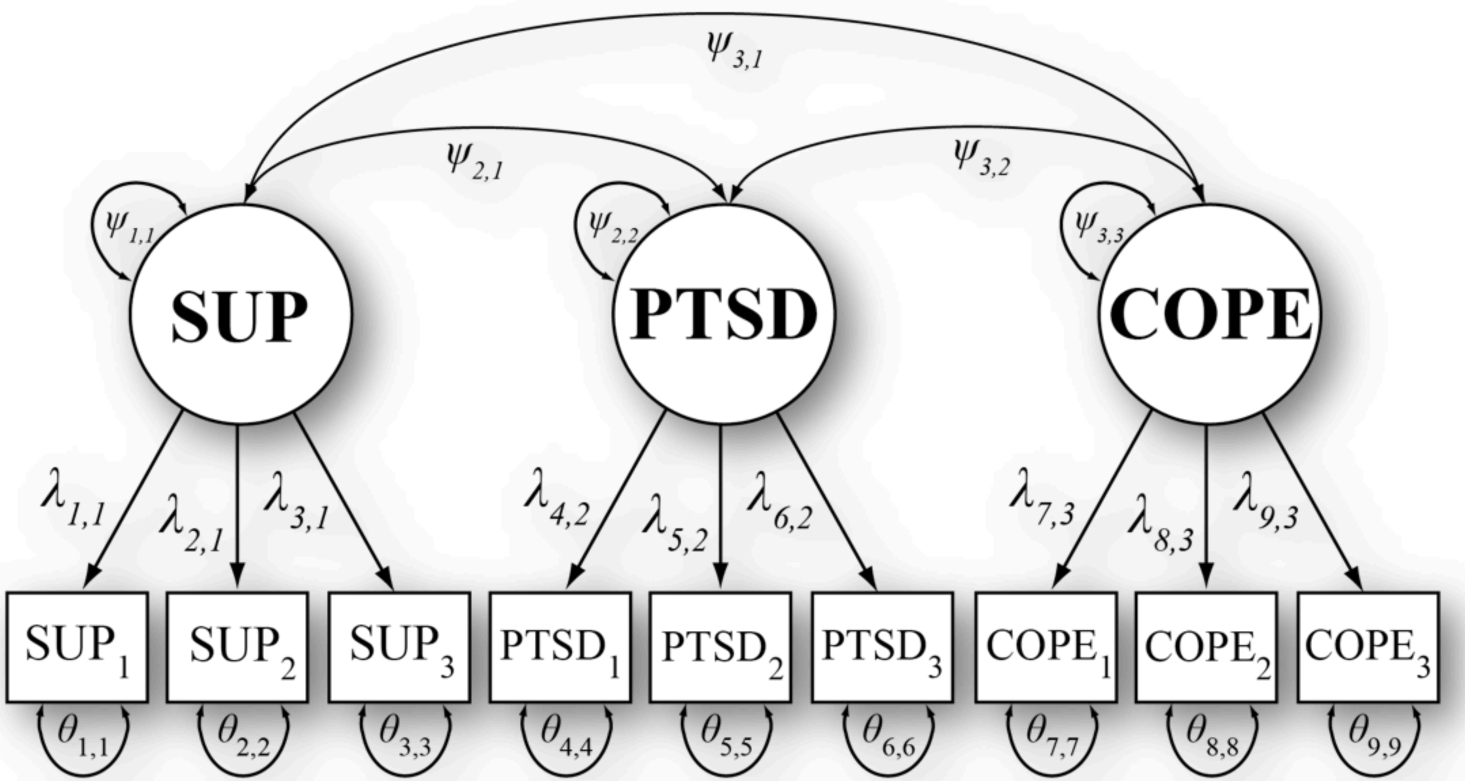
Standardized Results for PATH List

Path			Parameter	Estimate	Standard Error	t Value	Pr > |t|
SUP	===>	SUP1	ly1	0.92725	0.01879	49.3364	<.0001
SUP	===>	SUP2	ly2	0.94855	0.01718	55.2264	<.0001
SUP	===>	SUP3	ly3	0.83197	0.02905	28.6353	<.0001

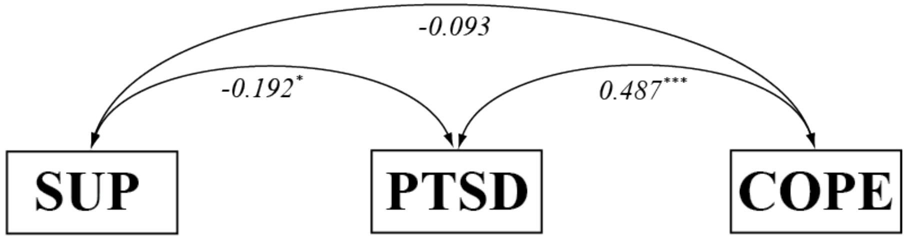
Standardized Results for Variance Parameters

Variance Type	Variable	Parameter	Estimate	Standard Error	t Value	Pr > |t|
Exogenous	SUP		1.00000			
Error	SUP1	theta1	0.14022	0.03485	4.0229	<.0001
	SUP2	theta2	0.10025	0.03258	3.0766	0.0021
	SUP3	theta3	0.30782	0.04834	6.3673	<.0001

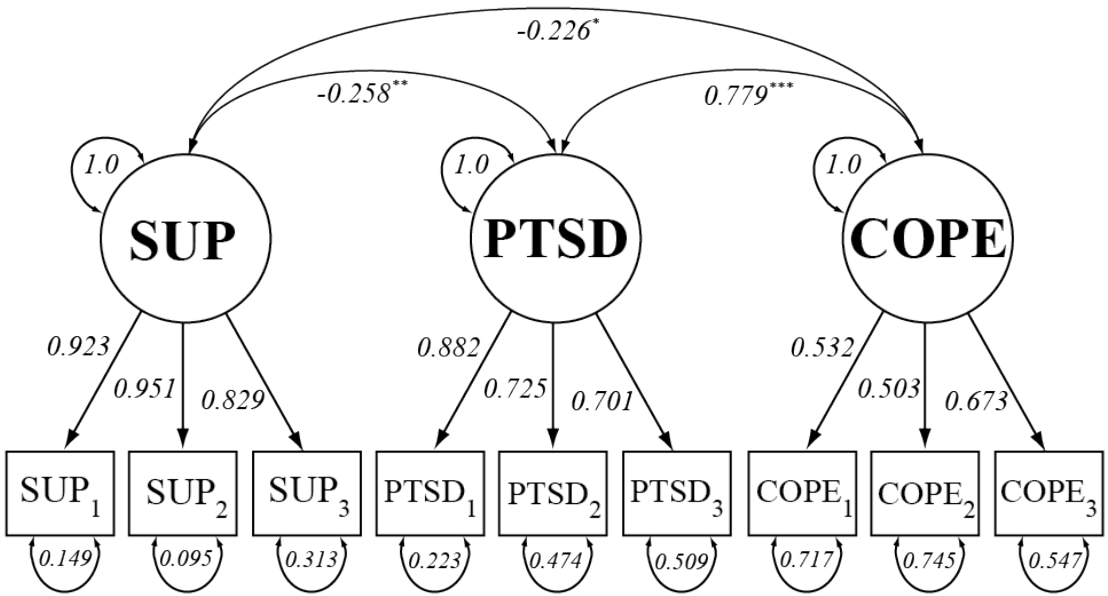
Observed vs. Latent Correlations.

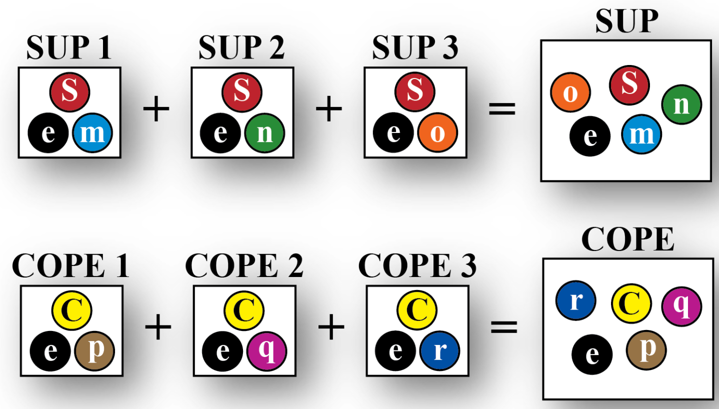


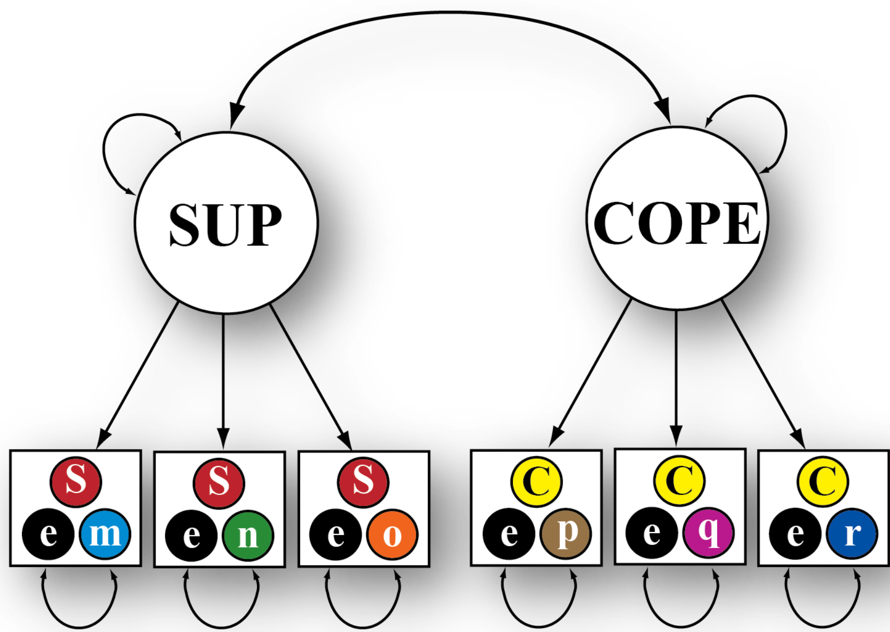
2.1-manifest-correlations.inp

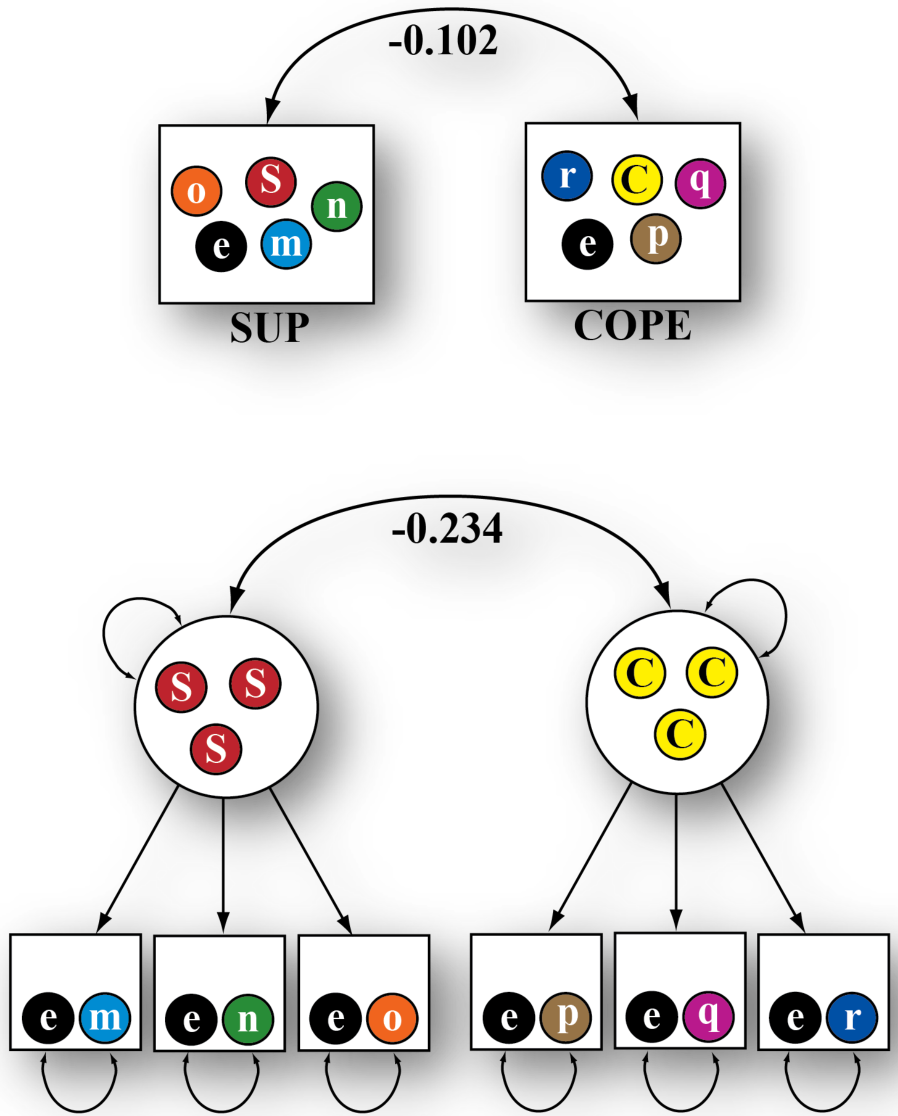


2.2-latent-correlations.inp

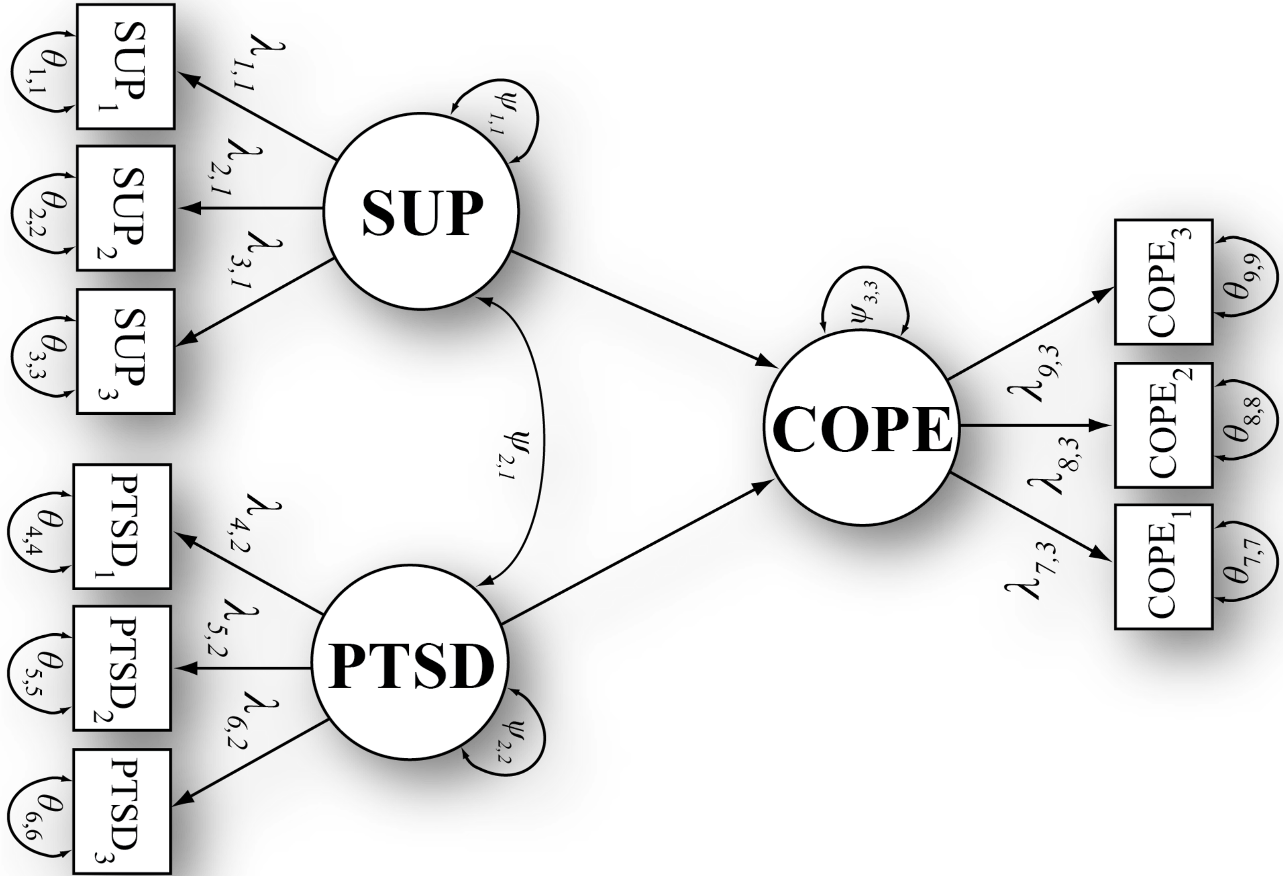



SUP 1
SUP 2
SUP 3
SUP
COPE 1
COPE 2
COPE 3
COPE

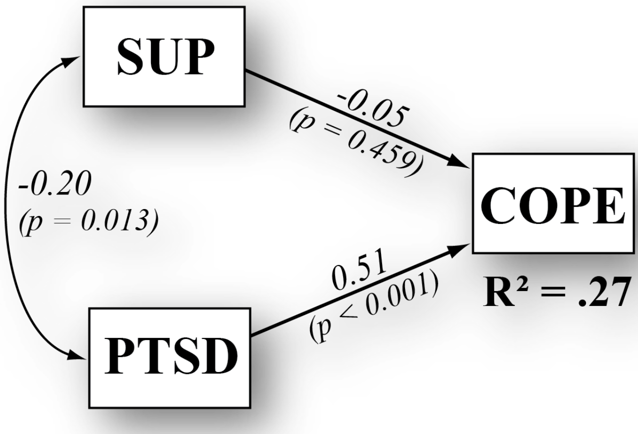

SUP
COPE

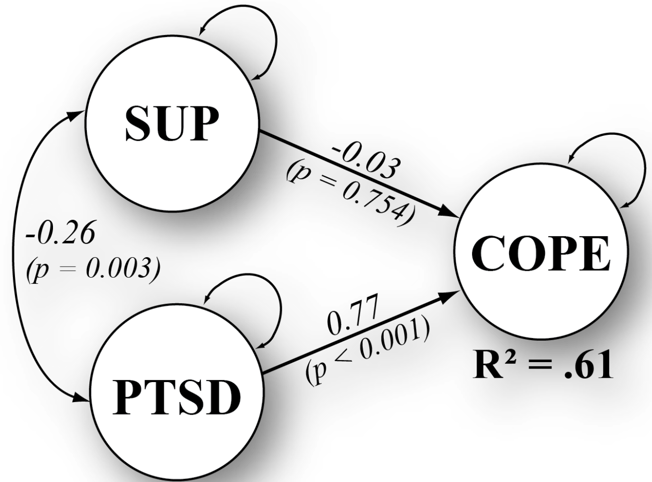

-0.102
o
S
n
e
m
SUP
r
C
q
e
p
COPE
-0.234
S
S
S
C
C
C
e
m
e
n
e
o
e
p
e
q
e
r

Structural Equation Modeling



Latent vs. Manifest/Observed

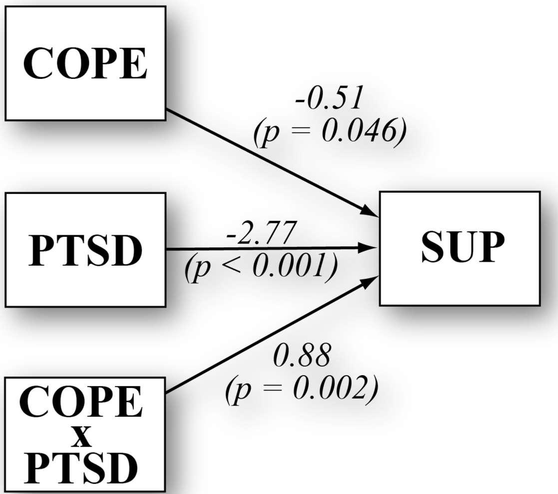




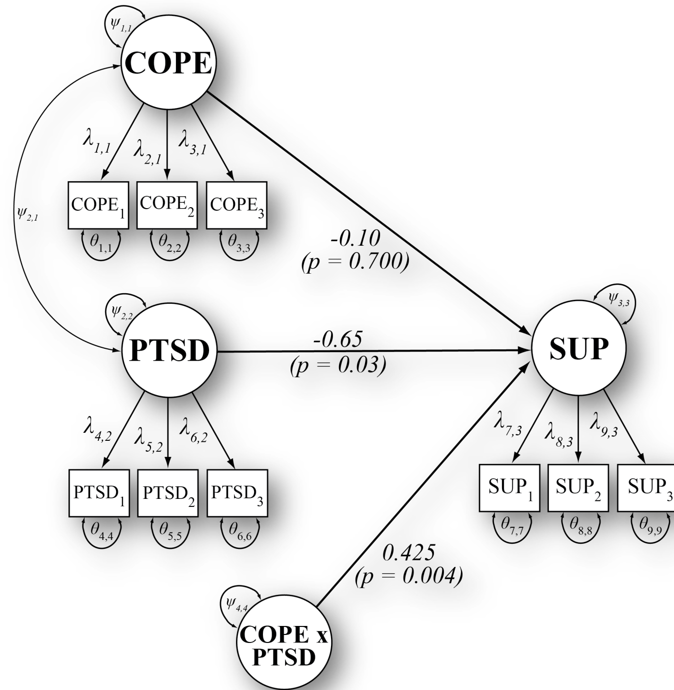
Measurement error in the observed variables can reduce the accuracy of the regression model.

Model Fit: $\chi^2(24, n = 144) = 36.14$; RMSEA = .059 (.000; .097) ; CFI = .980; TLI = .970

Manifest ($R^2 = 0.11$)

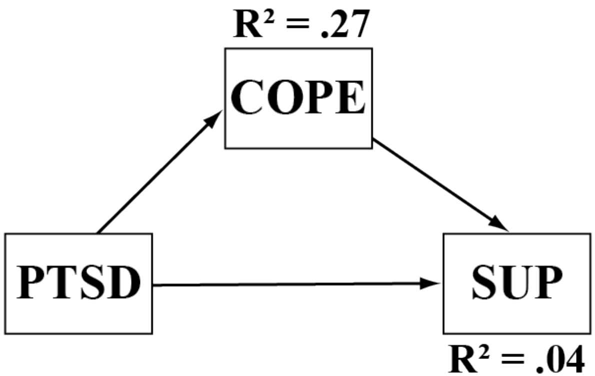


Latent ($R^2 = 0.46$)

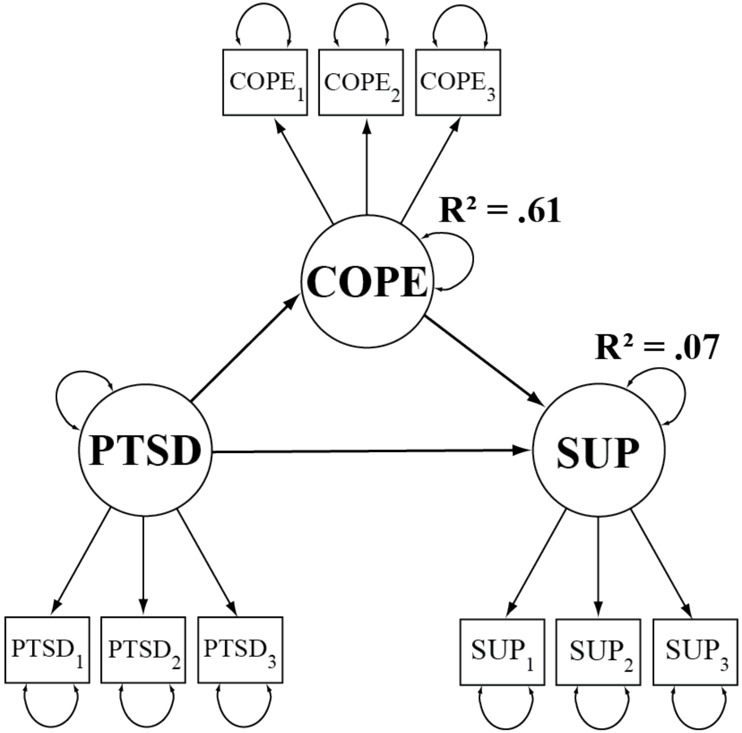


PTSD severity is a relatively modest / key determinant in whether coping strategies help or hinder social support.

Manifest ab = -0.10 [-0.37, 0.18]



Latent ab = -0.05 [-0.64, 0.54]



While this might seem less desirable, SEM accounts for measurement error, which can lead to more realistic estimates.

Power Analysis: Model Parameters

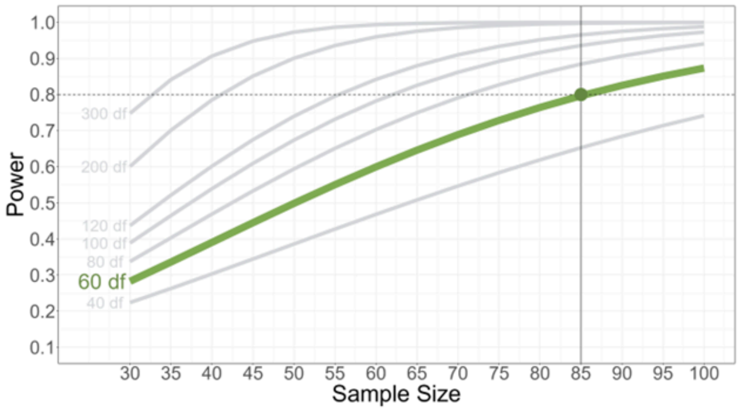
```
TITLE: Monte Carlo Simulation for Power;

MONTECARLO:
  NAMES ARE sup1 sup2 sup3 cope1 cope2 cope3;
  NOBSERVATIONS = 150;
  NREPS = 2000;
  SEED = 461981;
  !save = mypower.dat;

MODEL POPULATION:
    SUP BY sup1*0.924 sup2*0.953 sup3*0.832;
    COPE BY cope1*0.532 cope2*0.505 cope3*0.675;
    !SUP WITH COPE*-0.234;
    SUP WITH COPE*-0.300;
    sup1*0.146 sup2*0.093 sup3*0.308;
    cope1*0.717 cope2*0.745 cope3*0.545;
    SUP@1; COPE@1;

MODEL:
    SUP BY sup1*0.924 sup2*0.953 sup3*0.832;
    COPE BY cope1*0.532 cope2*0.505 cope3*0.675;
    !SUP WITH COPE*-0.234;
    SUP WITH COPE*-0.300;
    sup1*0.146 sup2*0.093 sup3*0.308;
    cope1*0.717 cope2*0.745 cope3*0.545;
    SUP@1; COPE@1;

ANALYSIS: ESTIMATOR = ML;
OUTPUT: TECH9;
```

```
MODEL RESULTS

                              ESTIMATES   95%  % Sig
               Population   Average     Cover Coeff

 SUP      BY
  SUP1               0.924      0.9203   0.955 1.000
  SUP2               0.953      0.9491   0.945 1.000
  SUP3               0.832      0.8293   0.957 1.000

 COPE     BY
  COPE1              0.532      0.5300   0.944 0.995
  COPE2              0.505      0.4991   0.949 0.993
  COPE3              0.675      0.6837   0.952 0.997

 SUP      WITH
  COPE              -0.300     -0.2928   0.940 0.797
```

80% power for a correlation of at least -0.30 given *N* = 150.

Power Analysis: Model Fit

```
w = NULL # We want to store results in here

alpha = .05
d_list = c(40, 60, 80, 100, 120, 200, 300) #degrees of freedom
n_list = c(seq(30, 100, by = 5)) #sample size
rmsea0 <- 0.05 #null hypothesized RMSEA (.05 close fit, .00 exact fit)
rmseaa_list = c(.10)

for (d in d_list){
  for (n in n_list){
    for (rmseaa in rmseaa_list){

      ncp0 <- (n-1)*d*rmsea0^2
      ncpa <- (n-1)*d*rmseaa^2

      #Compute power
      if(rmsea0<rmseaa) {
        cval <- qchisq(alpha,d,ncp=ncp0,lower.tail=F)
        pow <- pchisq(cval,d,ncp=ncpa,lower.tail=F)
      }
      if(rmsea0>rmseaa) {
        cval <- qchisq(1-alpha,d,ncp=ncp0,lower.tail=F)
        pow <- 1-pchisq(cval,d,ncp=ncpa,lower.tail=F)
      }

      w = rbind(w, c(d, n, rmseaa, pow))

      colnames(w) = c('df', 'n', 'rmseaa', 'Power')
      simdf <- as.data.frame(w)
      simdf$grp <- as.factor(simdf$df)

    }
  }
}
```



80% power for a CFA with at least 60 df given *N* = 85.

End