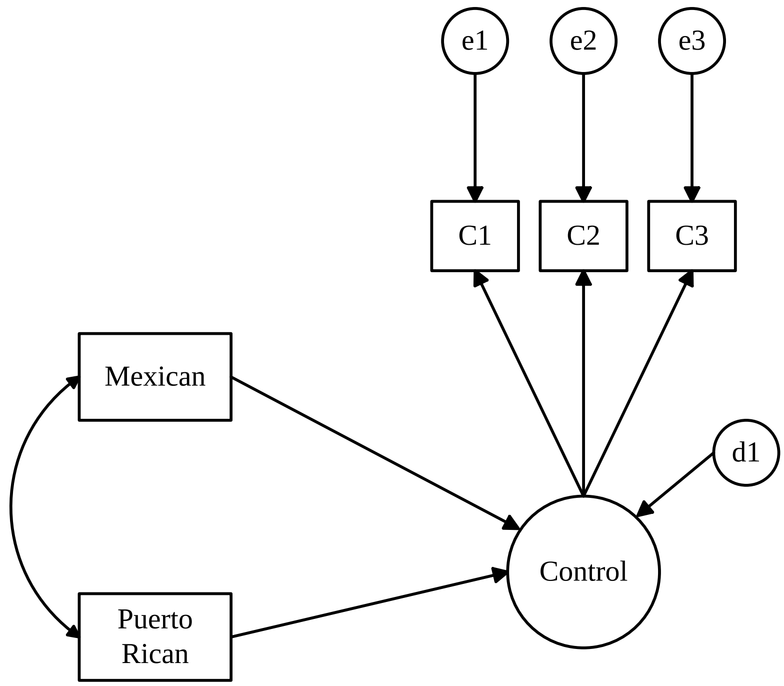
Categorical Exogenous Variables

A researcher collected data on adolescents in high school from three different ethnic groups, Cubans, Mexicans and Puerto Ricans. She want to test if there were mean differences in perceived control over one's life (i.e., locus of control). Three measures of locus of control were obtained, each ranging from 0 to 100. Higher scores indicated more perceived control.


e1
e2
e3
C1
C2
C3
Mexican
d1
Control
Puerto Rican

Categorical Variable Analysis

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	11	5.351	4	.253	1.338
Saturated model	15	.000	0		
Independence model	5	10678.397	10	.000	1067.840

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.017	.999	.996	.266
Saturated model	.000	1.000		
Independence model	115.892	.469	.204	.313

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.999	.999	1.000	1.000	1.000
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.013	.000	.038	.996
Independence model	.731	.719	.743	.000

Standardized Residual Covariances (Group number 1 - Default model)

	PR	MEX	CONTROL3	CONTROL2	CONTROL1
PR	.000				
MEX	.000	.000			
CONTROL3	.232	.069	.000		
CONTROL2	-.346	-.010	.000	.000	
CONTROL1	.147	-.052	.001	-.001	.000

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

	M.I.	Par Change

Variances: (Group number 1 - Default model)

	M.I.	Par Change

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
Lcontrol	<---	MEX	-3.980	1.002	-3.973	***
Lcontrol	<---	PR	-2.236	.921	-2.429	.015
CONTROL1	<---	Lcontrol	1.000			
CONTROL2	<---	Lcontrol	.995	.007	151.301	***
CONTROL3	<---	Lcontrol	.992	.007	146.326	***

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
Lcontrol	<---	MEX	-.092
Lcontrol	<---	PR	-.056
CONTROL1	<---	Lcontrol	.980
CONTROL2	<---	Lcontrol	.980
CONTROL3	<---	Lcontrol	.977

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
MEX	<-->	PR	-.034	.003	-9.897	***

Correlations: (Group number 1 - Default model)

			Estimate
MEX	<-->	PR	-.227

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
Lcontrol	.009
CONTROL3	.955
CONTROL2	.960
CONTROL1	.960

Analysis With Changed Reference Group

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	11	5.351	4	.253	1.338
Saturated model	15	.000	0		
Independence model	5	11726.199	10	.000	1172.620

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.025	.999	.996	.266
Saturated model	.000	1.000		
Independence model	115.892	.437	.156	.291

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	1.000	.999	1.000	1.000	1.000
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.013	.000	.038	.996
Independence model	.766	.754	.778	.000

Standardized Residual Covariances (Group number 1 - Default model)

	PR	cuban	CONTROL3	CONTROL2	CONTROL1
PR	.000				
cuban	.000	.000			
CONTROL3	.232	-.247	.000		
CONTROL2	-.346	.296	.000	.000	
CONTROL1	.147	-.083	.001	-.001	.000

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

	M.I.	Par Change

Variances: (Group number 1 - Default model)

	M.I.	Par Change

Regression Weights: (Group number 1 - Default model)

	M.I.	Par Change

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
Lcontrol	<---	cuban	3.980	1.002	3.973	***
Lcontrol	<---	PR	1.745	1.197	1.458	.145
CONTROL1	<---	Lcontrol	1.000			
CONTROL2	<---	Lcontrol	.995	.007	151.301	***
CONTROL3	<---	Lcontrol	.992	.007	146.326	***

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
Lcontrol	<---	cuban	.119
Lcontrol	<---	PR	.044
CONTROL1	<---	Lcontrol	.980
CONTROL2	<---	Lcontrol	.980
CONTROL3	<---	Lcontrol	.977

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P
cuban	<-->	PR	-.130	.005	-24.678	***

Correlations: (Group number 1 - Default model)

			Estimate
cuban	<-->	PR	-.662

Squared Multiple Correlations: (Group number 1 - Default model)

	Estimate
Lcontrol	.009
CONTROL3	.955
CONTROL2	.960
CONTROL1	.960

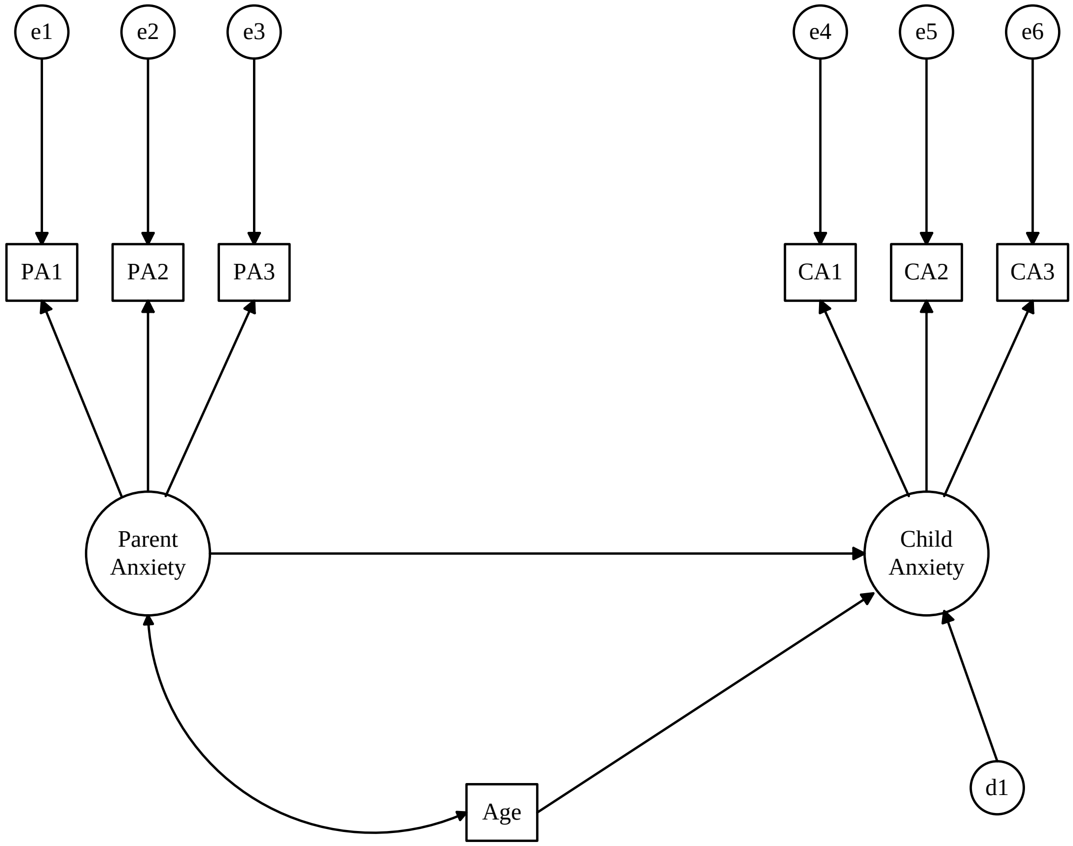
Holm Modified Bonferroni Test

Contrast	p Value	Critical p Value
Cuban vs. Mexican	<.001	.05/3 = 0.0167
Cuban vs. Puerto Rican	.015	.05/2 = 0.025
Mexican vs. Puerto Rican	.145	.05/1 = 0.05

Example Write-Up of Categorical Variable Model

The model that was fit is in Figure 1. Ethnicity was reflected by two dummy variables using dummy coding with Cubans serving as the reference group. The fit of the model was evaluated with AMOS 17.0. The model is statistically overidentified. Prior to analysis, the data for the continuous variables were evaluated for multivariate outliers by examining leverage indices for each individual and defining an outlier as a leverage score four times greater than the mean leverage. Two outliers were identified and checked for accurate coding. There were no coding errors and the outliers proved to be inconsequential for the analysis (i.e., all major conclusions remained intact when they were omitted from the analysis). The analysis reported includes the outlier cases. Model based outliers were evaluated using a limited information approach in which each indicator for the latent endogenous variable was separately regressed onto the dummy variables representing the variables influencing the endogenous variable. The analysis used ordinary least squares regression. Standardized dfbetas were examined for each individual and each predictor as well as the intercept. An influential outlier was defined as any individual with an absolute standardized dfbeta greater than 1 for a given coefficient. No outliers were evident. Multivariate normality was evaluated using Mardia's index. The multivariate test was not statistically significant ($p > 0.05$). Examination of univariate indices of skewness and kurtosis revealed no absolute skewness value greater than 0.66 nor absolute kurtosis values greater than 0.20. There were no missing data. A variety of indices of model fit was evaluated. The overall chi square test of model fit was statistically non-significant (X^2 (4) = 5.35, $p < 0.26$). The Root Mean Square Error of Approximation (RMSEA) was 0.013. The p value for the test of close fit was 0.99. The Comparative Fit index was 1.00 and the standardized root mean square residual was 0.02. The indices uniformly point towards good model fit. Inspection of the residuals and modification indices revealed no theoretically meaningful and significant points of ill-fit in the model. Figure 1 presents the parameter estimates for the model. The latent variable mean difference between Mexican Americans and Cubans was statistically significant (mean difference = -3.98, $z = 3.97$, $p < .05$), indicating that Cuban adolescents perceive themselves as having higher levels of control than Mexican adolescents. The latent mean difference for Puerto Ricans and Cubans was -2.24 ($z = 2.24$, $p < .05$), indicating that Cuban adolescents perceive themselves as having higher levels of control than Puerto Rican adolescents. The latent mean difference between Puerto Ricans and Mexicans was not statistically significant (mean difference = 1.74, $z = 1.46$, ns). This pattern of significant differences maintained when a Holm modified Bonferroni procedure was used to control the familywise error rate at 0.05.

A researcher is interested in the effects of maternal anxiety on child anxiety and if the effects vary as a function of gender. She collects three indicators of child anxiety and three indicators of parent anxiety. Each measure ranges from 0 to 100, with higher scores indicating higher anxiety. The indicators came from different sources (e.g., for the child, a self report, the child's mother, and a clinician; for the mother, a self report, a relative of the mother, and a clinician). The researcher hypothesized that the effect of maternal anxiety on child anxiety would be stronger for girls than for boys. The sample included children ranging from 6 to 17, so age was included as a covariate.


e1
e2
e3
e4
e5
e6
PA1
PA2
PA3
CA1
CA2
CA3
Parent Anxiety
Child Anxiety
Age
d1

Equal Form Model

Assessment of normality (Males)

Variable	min	max	skew	c.r.	kurtosis	c.r.
age	6.000	17.000	.338	2.392	-.406	-1.435
ca3	21.504	74.509	-.301	-2.129	-.116	-.408
ca2	27.871	74.251	.026	.187	-.358	-1.266
ca1	20.860	72.717	.072	.508	-.255	-.902
pa3	24.846	80.983	.096	.678	-.034	-.121
pa2	22.275	79.229	.132	.934	.208	.735
pa1	24.803	73.584	.074	.527	-.209	-.740
Multivariate					-1.556	-1.200

Assessment of normality (Females)

Variable	min	max	skew	c.r.	kurtosis	c.r.
age	6.000	17.000	.342	2.416	-.687	-2.430
ca3	16.583	72.268	-.425	-3.003	.119	.420
ca2	27.554	72.356	-.031	-.218	-.253	-.893
ca1	20.467	76.603	.018	.126	.174	.614
pa3	24.846	80.983	.096	.678	-.034	-.121
pa2	22.275	79.229	.132	.934	.208	.735
pa1	24.803	73.584	.074	.527	-.209	-.740
Multivariate					-2.334	-1.801

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	32	24.282	24	.446	1.012
Saturated model	56	.000	0		
Independence model	14	1901.915	42	.000	45.284

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	2.161	.989	.973	.424
Saturated model	.000	1.000		
Independence model	33.300	.478	.304	.358

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.987	.978	1.000	1.000	1.000
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.004	.000	.034	.999
Independence model	.272	.262	.283	.000

Modification Indices (Males - Default model)

Covariances: (Males - Default model)

	M.I.	Par Change

Variances: (Males - Default model)

	M.I.	Par Change

Regression Weights: (Males - Default model)

	M.I.	Par Change

Modification Indices (Females - Default model)

Covariances: (Females - Default model)

	M.I.	Par Change

Variances: (Females - Default model)

	M.I.	Par Change

Regression Weights: (Females - Default model)

	M.I.	Par Change

Standardized Residual Covariances (Males - Default model)

	age	ca3	ca2	ca1	pa3	pa2	pa1
age	.000						
ca3	-.122	.000					
ca2	.144	-.010	.000				
ca1	-.014	-.025	.038	.000			
pa3	-.301	.366	-.512	-.093	.000		
pa2	.247	1.460	.490	.824	-.052	.000	
pa1	.045	.246	-1.467	-.865	.041	-.001	.000

Standardized Residual Covariances (Females - Default model)

	age	ca3	ca2	ca1	pa3	pa2	pa1
age	.000						
ca3	.011	.000					
ca2	-.175	.010	.000				
ca1	.154	-.077	.084	.000			
pa3	-.352	-.032	-.055	.047	.000		
pa2	.760	.686	.593	.610	-.175	.000	
pa1	-.382	-.148	-.951	-.670	.152	.004	.000

Estimates (Males - Default model)

Regression Weights: (Males - Default model)

			Estimate	S.E.	C.R.	P
Child Anx	<---	age	.793	.170	4.670	***
Child Anx	<---	Parent Anx	.157	.059	2.675	.007
pa1	<---	Parent Anx	1.000			
pa2	<---	Parent Anx	1.001	.060	16.665	***
pa3	<---	Parent Anx	1.044	.063	16.708	***
ca1	<---	Child Anx	1.000			
ca2	<---	Child Anx	.977	.073	13.353	***
ca3	<---	Child Anx	1.058	.079	13.444	***

Standardized Regression Weights: (Males - Default model)

			Estimate
Child Anx	<---	age	.286
Child Anx	<---	Parent Anx	.172
pa1	<---	Parent Anx	.869
pa2	<---	Parent Anx	.835
pa3	<---	Parent Anx	.838
ca1	<---	Child Anx	.798
ca2	<---	Child Anx	.801
ca3	<---	Child Anx	.812

Covariances: (Males - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	.375	1.344	.279	.780	

Correlations: (Males - Default model)

			Estimate
Parent Anx	<-->	age	.017

Squared Multiple Correlations: (Males - Default model)

	Estimate
Child Anx	.113
ca3	.659
ca2	.641
ca1	.637
pa3	.701
pa2	.697
pa1	.755

Estimates (Females - Default model)

Regression Weights: (Females - Default model)

			Estimate	S.E.	C.R.	P
Child Anx	<---	age	.612	.138	4.422	***
Child Anx	<---	Parent Anx	.537	.059	9.138	***
pa1	<---	Parent Anx	1.000			
pa2	<---	Parent Anx	1.025	.060	17.074	***
pa3	<---	Parent Anx	1.058	.062	16.933	***
ca1	<---	Child Anx	1.000			
ca2	<---	Child Anx	.950	.070	13.645	***
ca3	<---	Child Anx	1.098	.078	14.121	***

Standardized Regression Weights: (Females - Default model)

			Estimate
Child Anx	<---	age	.235
Child Anx	<---	Parent Anx	.585
pa1	<---	Parent Anx	.859
pa2	<---	Parent Anx	.845
pa3	<---	Parent Anx	.838
ca1	<---	Child Anx	.797
ca2	<---	Child Anx	.789
ca3	<---	Child Anx	.825

Covariances: (Females - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	2.066	1.415	1.460	.144	

Correlations: (Females - Default model)

			Estimate
Parent Anx	<-->	age	.090

Squared Multiple Correlations: (Females - Default model)

	Estimate
Child Anx	.422
ca3	.681
ca2	.623
ca1	.636
pa3	.703
pa2	.714
pa1	.737

Measurement Invariance Model

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	28	24.760	28	.641	.884
Saturated model	56	.000	0		
Independence model	14	1901.915	42	.000	45.284

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	2.224	.988	.977	.494
Saturated model	.000	1.000		
Independence model	33.300	.478	.304	.358

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.987	.980	1.002	1.003	1.000
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.000	.000	.027	1.000

Estimates (Males - Default model)

Regression Weights: (Males - Default model)

			Estimate	S.E.	C.R.	P	Label
Child Anx	<---	age	.790	.168	4.717	***	
Child Anx	<---	Parent Anx	.159	.059	2.723	.006	
pa1	<---	Parent Anx	1.000				
pa2	<---	Parent Anx	1.013	.042	23.866	***	c
pa3	<---	Parent Anx	1.051	.044	23.802	***	d
ca1	<---	Child Anx	1.000				
ca2	<---	Child Anx	.963	.050	19.106	***	a
ca3	<---	Child Anx	1.078	.055	19.472	***	b

Standardized Regression Weights: (Males - Default model)

			Estimate
Child Anx	<---	age	.286
Child Anx	<---	Parent Anx	.174
pa1	<---	Parent Anx	.867
pa2	<---	Parent Anx	.838
pa3	<---	Parent Anx	.838
ca1	<---	Child Anx	.797
ca2	<---	Child Anx	.793
ca3	<---	Child Anx	.820

Covariances: (Males - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	.373	1.336	.279	.780	

Correlations: (Males - Default model)

			Estimate
Parent Anx	<-->	age	.017

Squared Multiple Correlations: (Males - Default model)

	Estimate
Child Anx	.114
ca3	.672
ca2	.629
ca1	.635
pa3	.702
pa2	.702
pa1	.751

Estimates (Females - Default model)

Regression Weights: (Females - Default model)

			Estimate	S.E.	C.R.	P	Label
Child Anx	<---	age	.614	.138	4.461	***	
Child Anx	<---	Parent Anx	.535	.055	9.719	***	
pa1	<---	Parent Anx	1.000				
pa2	<---	Parent Anx	1.013	.042	23.866	***	c
pa3	<---	Parent Anx	1.051	.044	23.802	***	d
ca1	<---	Child Anx	1.000				
ca2	<---	Child Anx	.963	.050	19.106	***	a
ca3	<---	Child Anx	1.078	.055	19.472	***	b

Standardized Regression Weights: (Females - Default model)

			Estimate
Child Anx	<---	age	.235
Child Anx	<---	Parent Anx	.584
pa1	<---	Parent Anx	.861
pa2	<---	Parent Anx	.842
pa3	<---	Parent Anx	.838
ca1	<---	Child Anx	.798
ca2	<---	Child Anx	.796
ca3	<---	Child Anx	.818

Covariances: (Females - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	2.069	1.422	1.455	.146	

Correlations: (Females - Default model)

			Estimate
Parent Anx	<-->	age	.090

Squared Multiple Correlations: (Females - Default model)

	Estimate
Child Anx	.421
ca3	.670
ca2	.634
ca1	.638
pa3	.703
pa2	.710
pa1	.741

Path Equivalence Model

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	27	47.658	29	.016	1.643
Saturated model	56	.000	0		
Independence model	14	1901.915	42	.000	45.284

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	8.226	.978	.957	.506
Saturated model	.000	1.000		
Independence model	33.300	.478	.304	.358

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.975	.964	.990	.985	.990
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.033	.014	.049	.960
Independence model	.272	.262	.283	.000

Estimates (Males - Default model)

Regression Weights: (Males - Default model)

			Estimate	S.E.	C.R.	P	Label
Child Anx	<---	age	.778	.171	4.549	***	
Child Anx	<---	Parent Anx	.375	.042	9.037	***	e
pa1	<---	Parent Anx	1.000				
pa2	<---	Parent Anx	1.014	.043	23.809	***	c
pa3	<---	Parent Anx	1.052	.044	23.739	***	d
ca1	<---	Child Anx	1.000				
ca2	<---	Child Anx	.962	.050	19.286	***	a
ca3	<---	Child Anx	1.080	.055	19.701	***	b

Standardized Regression Weights: (Males - Default model)

			Estimate
Child Anx	<---	age	.261
Child Anx	<---	Parent Anx	.375
pa1	<---	Parent Anx	.860
pa2	<---	Parent Anx	.836
pa3	<---	Parent Anx	.834
ca1	<---	Child Anx	.818
ca2	<---	Child Anx	.813
ca3	<---	Child Anx	.842

Covariances: (Males - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	.375	1.323	.283	.777	

Correlations: (Males - Default model)

			Estimate
Parent Anx	<-->	age	.017

Squared Multiple Correlations: (Males - Default model)

	Estimate
Child Anx	.212
ca3	.709
ca2	.660
ca1	.669
pa3	.695
pa2	.699
pa1	.739

Estimates (Females - Default model)

Regression Weights: (Females - Default model)

			Estimate	S.E.	C.R.	P	Label
Child Anx	<---	age	.655	.139	4.699	***	
Child Anx	<---	Parent Anx	.375	.042	9.037	***	e
pa1	<---	Parent Anx	1.000				
pa2	<---	Parent Anx	1.014	.043	23.809	***	c
pa3	<---	Parent Anx	1.052	.044	23.739	***	d
ca1	<---	Child Anx	1.000				
ca2	<---	Child Anx	.962	.050	19.286	***	a
ca3	<---	Child Anx	1.080	.055	19.701	***	b

Standardized Regression Weights: (Females - Default model)

			Estimate
Child Anx	<---	age	.267
Child Anx	<---	Parent Anx	.442
pa1	<---	Parent Anx	.867
pa2	<---	Parent Anx	.844
pa3	<---	Parent Anx	.842
ca1	<---	Child Anx	.779
ca2	<---	Child Anx	.777
ca3	<---	Child Anx	.800

Covariances: (Females - Default model)

			Estimate	S.E.	C.R.	P	Label
Parent Anx	<-->	age	2.057	1.435	1.434	.152	

Correlations: (Females - Default model)

			Estimate
Parent Anx	<-->	age	.088

Squared Multiple Correlations: (Females - Default model)

	Estimate
Child Anx	.288
ca3	.640
ca2	.604
ca1	.607
pa3	.709
pa2	.713
pa1	.751

Example Write-Up of Multi-Group Model

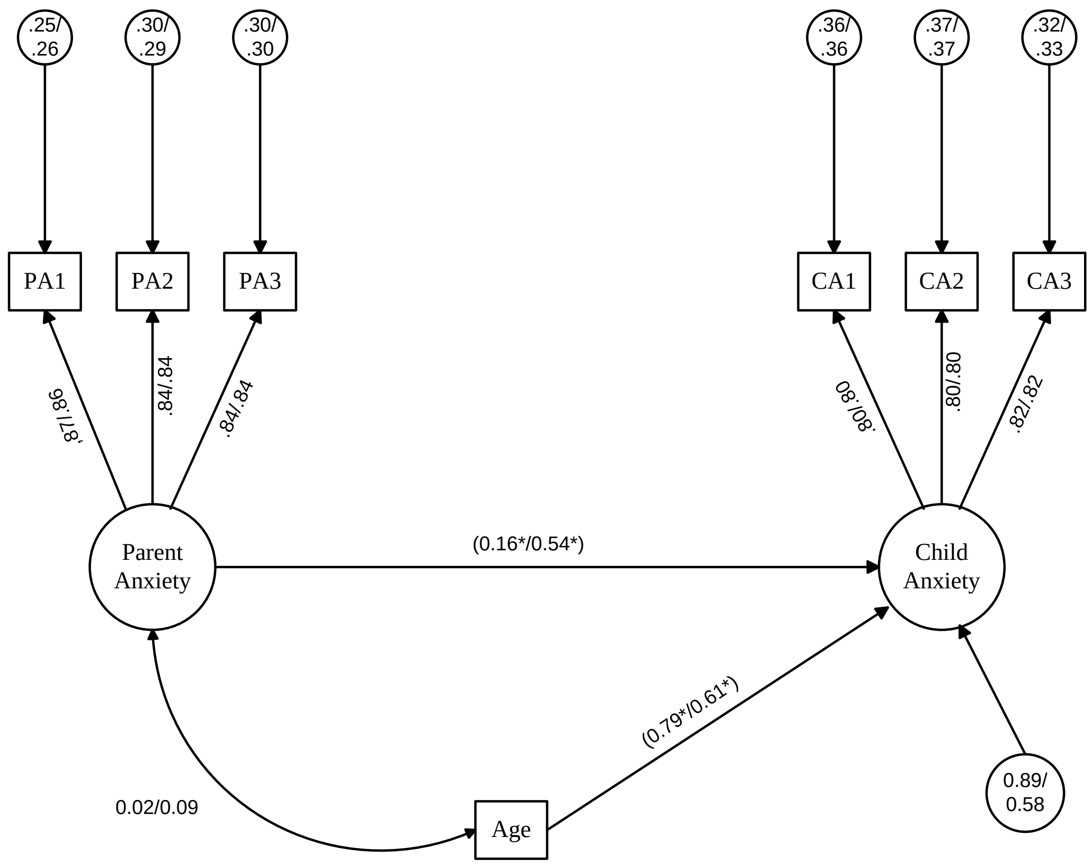
The model that was fit is in Figure 1. The fit of the model was evaluated with AMOS 17.0 with a maximum likelihood solution. A multi-group solution was pursued with males and females representing the two groups. The model is statistically overidentified. Prior to analysis and within each group, the data for the continuous variables were evaluated for multivariate outliers by examining leverage indices for each individual and defining an outlier as a leverage score four times greater than the mean leverage. No outliers were found. Model based outliers were evaluated using a limited information approach in which a single indicator for a given latent endogenous variable was regressed onto a single indicator of all variables that directly influence that variable in the model. The endogenous variable was regressed onto the predictor variables and the standardized dfbetas were examined for each individual, using ordinary least squares regression. An influential outlier was defined as any individual with an absolute standardized dfbeta greater than 1 for a given coefficient. No outliers were present based on this analysis. There were no missing data. Multivariate normality was evaluated using Mardia's index. The multivariate test was statistically non-significant ($p > 0.05$) for both males and females. Examination of univariate indices of skewness and kurtosis within each group revealed no absolute skewness values greater than 0.42 nor absolute kurtosis values greater than 0.69.

The first model tested had no equality constraints across groups. The purpose of this analysis was to establish a common model form in the two groups. We refer to this model as the equal-form model. This model yielded a good fit to the data. The overall chi square test of model fit was statistically non-significant (X^2 (24) = 24.28, $p < 0.45$). The Root Mean Square Error of Approximation (RMSEA) was 0.004. The p value for the test of close fit was 0.99. The Comparative Fit Index was 1.00 and the traditional GFI was 0.99. The standardized root mean square residual was 0.02. More focused tests of fit revealed no theoretically meaningful nor sizeable modification indices, nor were any of the absolute standardized residuals larger than 1.96.

The second model tested for measurement invariance across the groups for the latent variables. This model was equivalent to the equal form model but with the constraint that the factor loading for an indicator of a latent variable in one group had to be equal in value to the corresponding factor loading in the other group. We refer to this model as the measurement invariance model. This model also yielded a good fit to the data. The overall chi square test of model fit was statistically non-significant (X^2 (28) = 24.76, $p < 0.64$). The Root Mean Square Error of Approximation (RMSEA) was < 0.001. The p value for the test of close fit was 0.99. The Comparative Fit Index was 1.00 and the traditional GFI was 0.99. The standardized root mean square residual was 0.02. More focused tests of fit revealed no theoretically meaningful and sizeable modification indices, nor were any of the absolute standardized residuals larger than 1.96. Importantly, the nested chi square test comparing this model to the equal form model yielded a statistically non-significant chi square difference (X^2 diff (4) = 0.48, ns), a result that is consistent with measurement invariance across groups. .

The final model tested the measurement invariance model but with an additional across group equality constraint, namely that the path coefficient from the latent parent anxiety variable to the latent child anxiety variable was constrained to be equal in the groups. We refer to this model as the path equivalence model. This model yielded an overall chi square test that was statistically significant (X^2 (29) = 47.66, $p < 0.02$). The Root Mean Square Error of Approximation (RMSEA) was < 0.013. The p value for the test of close fit was 0.96. The Comparative Fit Index was 0.99 and the traditional GFI was 0.98. The standardized root mean square residual was 0.05. We performed a nested chi square test by calculating the chi square difference between this model and the measurement invariance model (X^2 diff (1) = 22.9, $p < 0.01$). The chi square difference was statistically significant. This result leads us to reject the null hypothesis of equal path coefficients for males and females.

Figure 2 presents relevant coefficients for males and females for the measurement invariance model. Standardized coefficients for the factor loadings for the measurement model vary, even though the unstandardized coefficients were held constant across groups It can be seen that the effect of parent anxiety on child anxiety (holding age constant) was stronger for females (path = 0.537, $p < 0.01$) than it was for males (path = 0.157, $p < 0.01$).



Note: Unstandardized coefficients are in parentheses. Male coefficient is listed first, then female coefficient; * = p < 0.05 for structural coefficients

Two Factor Measurement Model Each With 3 Indicators

```
TITLE:  Example of CFA
DATA: FILE IS c:\temp\mplus\ex5.1.dat;
VARIABLE: NAMES ARE y1-y6;
          MISSING ALL (-999);
MODEL: f1 BY y1-y3;
       f2 BY y4-y6;
OUTPUT: Samp Stand Residual Mod(3.84) Cinterval Tech4
```

Notes: First indicator is by default fixed at 1.

Two Factor Measurement Model Each With 3 Indicators and Correlated Error

```
TITLE:  Example of CFA with correlated error
DATA: FILE IS c:\temp\mplus\ex5.1.dat;
VARIABLE: NAMES ARE y1-y6;
      MISSING ALL (-999);
      MODEL: f1 BY y1-y3;
       f2 BY y4-y6;
       y2 WITH y5;
OUTPUT: Samp Stand Residual Mod(3.84) Cinterval Tech4
```

Notes: Correlated error for the y2 and y5 indicators is introduced by y2 WITH y5 statement.

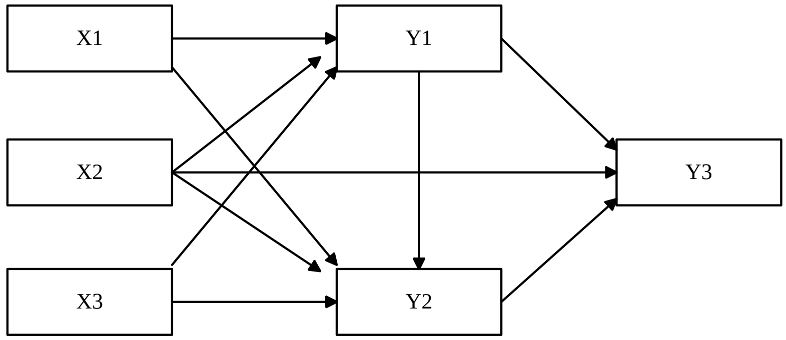
Two Factor Measurement Model With Binary/Ordinal Indicators

```
TITLE:  Example of CFA with binary or ordinal indicators
DATA: FILE IS c:\temp\mplus\ex5.2.dat;
VARIABLE: NAMES ARE y1-y6;
          CATEGORICAL ARE y1-y6;
          MISSING ARE ALL (-999);
MODEL: f1 BY y1-y3;
       f2 BY y4-y6;
OUTPUT: Samp Stand Residual Mod(3.84) Cinterval Tech4
```

Logistic Regression

```
TITLE:      Logistic regression
DATA: FILE IS c:\temp\mplus\ex3.5.dat;
VARIABLE:   NAMES ARE y1 x1 x2;
            CATEGORICAL IS y1;
ANALYSIS:   ESTIMATOR = ML;
MODEL:      y1 ON x1 x2;
OUTPUT:  SAMP STAND CINT ;
```

Single Indicator Model




```
TITLE: Example of single indicator model
DATA: FILE IS c:\temp\mplus\ex3.11.dat;
VARIABLE: NAMES ARE y1-y3 x1-x3;
ANALYSIS: Type = Mean;
MODEL:  y1 ON x1 x2 x3;
        y2 ON y1 x1 x2 x3;
        y3 ON y1 y2 x2;
MODEL INDIRECT:  y3 IND x1; y3 IND x2; y3 IND x3; y3 IND y1;
OUTPUT: Samp Stand Residual Mod(3.84) Cinterval Tech4;
```

Introducing Constraints

Parameters can be constrained to be equal by placing the same number in parentheses at the end of a line and continuing (if necessary) onto the next line. For example

```
y1 ON x1 x2 (1)
      x3;
```

regresses y1 onto x1, x2 and x3 and forces the x1 and x2 coefficients to be equal. Or,

```
y1 ON x1 (1)
      x2;
y2 ON x3 (1)
      x4;
```

regresses y1 onto x1 and x2 and y2 onto x3 and x4 and forces the coefficients for x1 and x3 to be equal.

As another example,

```
f1 BY y1 y2 y3 y4 y5 y6 (1)
      y7 y8 y9  (1);
```

forces factor loadings for y2 to y9 to be equal (the loading for y1 is set to 1.0 by default). But

```
f1 BY y1 y2 y3 y4 y5 y6
      y7 y8 y9  (1);
```

forces the loadings for y7-y9 to be equal but not the others (again with the y1 loading being fixed at 1.0).

```
f1 BY y1-y4 (1)
      y5-y6 (2)
      y7-y9 (3);
```

sets y1 as the marker variable and then forces y2-y4 to be equal, y5-y6 to be equal, y7-y9 to be equal.

[y1 y2 y3](1); forces the means (or intercepts) of y1 y2 and y3 to be equal.

To label a parameter, put the label after the parameter in parentheses

```
y on x1(p1)
     X2(p2)
     X3(p3);
```

This example labels the path for x1 to y as p1, x2 to y as p2 and x3 to y as p3. Only one label can appear per line.

MODEL CONSTRAINT LINE (Model Constraint:)

For non-linear constraints on parameters, first label the parameters, as above. Then add the statement to impose the constraint:

```
MODEL CONSTRAINT:
    p1 = p2*p3;
```

For constraints in multiple group solutions, see the tutorial on multiple group solutions.

GENERAL SYNTAX

1. M Plus correlates all exogenous variables by default

2. M Plus estimates all residual variances by default

3. For most analyses, M Plus uses ML solution by default, but this will change for certain kinds of analyses)

4. No line can be more than 60 characters long, no variable name more than 8 characters.

5. The ; symbol is used to terminate commands.

TITLE LINE (Title:)

This line is optional. Can use as many lines as want. Just hit return and keep typing. A line can not be more than 60 characters.

DATA LINE (Data:)

This line defines where to get the input data.

FILE IS c:\temp\mplus\ex3.1.dat; tells M Plus where to get the data file. The data are in free format in an ascii file.

Type=Individual; is the default and reads in raw data, respondents by variables
Type=COVARIANCE; reads in lower triangular covariance matrix (with diagonals)
Type=CORRELATION; reads in lower triangular correlation matrix (with diagonals)
Type=FULLCOV; reads in full covariance matrix
Type=FULLCORR; reads in full correlation matrix
Type=MEANS; reads in means
Type=STDEVIATIONS; reads in standard deviations
Type=IMPUTATION; reads in imputed data sets to be analyzed

The number of variables it reads in the above commands is linked to the number of names on the names command below. Five names means it reads in five variables per case.

NOBS=2000; is the number of observations to set N at when summary data are analyzed. Not necessary to specify this when the input is for raw data

NGROUPS=3; specifies 3 groups in a multi-group solution. Default is 1 group and do not need to specify if you are analyzing only one group. See **GROUPING IS** command in the variable line below for how to specify the groups themselves.

VARIABLE LINE (Variable:)

This line defines variable names and variable types.

NAMES ARE y x1 x2; Gives the names of the variables to be read in, in the order they occur in the input data set. A name can not be longer than 8 characters. Can not use commas in a name or blanks.
NAMES ARE y1-y5; will generate the names y1 y2 y3 y4 y5.
NAMES ARE jima-jimd; will generate the names jima jimb jimc jimd.

USEVARIABLES ARE y x1; Specifies the subset of variables to use if you only want to work with a subset of the variables input in the previous command. Default is to use all variables that were input.

USEOBSERVATIONS = ethnic EQ 1 AND gender EQ 2; uses only observations whose scores are 1 on ethnic and 2 on gender. Can invoke AND, OR, NOT (logical not) EQ (equal to) NE (not equal to) GE (greater than or equal to) LE (less than or equal to) GT (greater than) LT (less than). For example,
USEOBSERVATIONS = ethnic GT 1

MISSING ARE .; indicates a period in the entry as indicating a missing value for all variables

MISSING ARE ALL(9999); indicates a value of 9999 is a missing value for all variables

MISSING ARE ethnic(7 99) y1(8); indicates that the values 7 and 99 represent missing values for the variable ethnic and the value 8 represents a missing value for variable y1.

MISSING ARE ALL(7 99) indicates that the values 7 and 99 represent missing values for all variables.

CENSORED ARE y(a) x1(b); Specifies which variables are censored, with the letter a or b in the parentheses indicating censoring above or below. The censoring points are determined empirically.

CENSORED ARE y(ai) x1(bi); Specifies which variables are censored inflated, with the letter ai or bi in the parentheses indicating censoring above or below with inflation. The censoring points are determined empirically.

CATEGORICAL ARE x1 x2; Specifies x1 and x2 are categorical variables. For binary or ordered variables, zero must be the lowest category. If they are not, then they will automatically be recoded this way by M Plus. If the values of an ordered variable are 2, 5, 8, and 9, M Plus will recode these to 0, 1, 2, 3.

NOMINAL ARE x1 x2; Specifies which endogenous variables are treated as unordered, categorical variables. Limit is 10 categories per variable. The values are automatically renumbered by M Plus to begin with the number 0. In the Model command, you sometimes must refer to a specific category on a nominal variable. This is done with the # sign. So x1#1 refers to category 1 on the variable x1.

COUNT are x1 x2; Specifies which outcome variables are to be treated as count variables for Poisson like regression.

COUNT are x1(1) x2(i); Specifies which outcome variables are to be treated as count variables for Poisson like regression with zero inflated models.

GROUPING IS gender(1=male 2=female); For multiple group analysis specifies the grouping variable and the values and labels to use for that variable.

WEIGHT is x1; Specifies the weight variable in a complex sampling design to use in the analysis. Weights are not allowed with bootstrapping.

CLUSTER is x2; Specifies, in a complex sample design, the cluster information (e.g., school, household).

CENTERING = GRANDMEAN (x1 x2 x3); Specifies group mean centering for variables x1, x2, x3 and x4. Can use in any program (e.g., linear regression) to mean center the variables.

CENTERING = GROUPMEAN (x1 x2 x3); Specifies group mean centering for variables x1, x2, x3 and x4. Used only in multilevel analyses.

WITHIN = y1 y2 x1; Used in a multilevel design and specifies level 1 variables. They cannot be used as level 2 variables.

BETWEEN = y1 y2 x1; Used in a multilevel design and specifies level 2 variables. They cannot be used as level 1 variables.

DEFINE LINE (Define:)

This line allows for transformations of data that are input into M Plus

x1 = log(x1); takes the log of x1 and outs it in x1.

Can use the logical operators in if statements AND, OR, NOT (logical not) EQ (equal to) NE (not equal to) GE (greater than or equal to) LE (less than or equal to) GT (greater than) LT (less than). For example,
If (gender EQ 1 and ses EQ 2) then group = 1.

Can use addition (=), subtraction (-), multiplication (*), division(/), exponenitation(**) and remainder (%), i.e., the remainder of y/x. Also have log, log10, exp, sqrt, abs, sin, cos, tan, asin, acos, atan.

CUT y1(30,40); takes the continuous variable y1 and breaks it into 3 groups, (1) less than or equal to 30, (2) greater than 30 but less than 40 and (3) greater than or equal to 40. Can do multiple variables at a time with the same cutpoints
CUT y1-y7(30,40);
CUT y1 y2 y3 (30,40);

ANALYSIS Line (Analysis:)

Esti=ML; specifies a maximum likelihood analysis
Esti=MLM; specifies the Satorra-Bentler scaled chi square analysis

Bootstrap=2000; specifies 2000 bootstrap samples. You can not get modification indices with bootstrap. Also, to get standard z tests must do a separate run eliminating the bootstrap option. Once you specify bootstrap, all significance tests and CIs are bootstrap based.

Type=Mean; Specifies you want means and intercepts in the analysis using FIML

Type=Missing; Specifies you want to analyze data with missing values. If you do not specify this command here, then M Plus uses listwise deletion instead of FIML.

Type=Missing Mean; Specifies you want to analyze data with missing values using FIML and you want means and intercepts in the analysis. Note you can specify multiple types of analysis by placing all the types on the same line.

Type=Random; Specifies random intercept and random coefficient models.

Type=Complex; Specifies you want to compute standard error s and chi squares for cluster and complex sampling designs.

Type=Mixture; Specifies you want to compute models with a combination of continuous and categorical latent variables.

Type=TwoLevel; Specifies you want to analyze a multilevel model.

Type=Logistic; Specifies you want to conduct a logistic analysis.

Stseed=2365; Specifies seed for random number generator (in this case it is 2365).

MODEL LINE (Model:)

This line defines the model to be analyzed.

BY is short for "measured by" and defines the latent variable. For example, f1 BY y1 y2 y3;. First variable listed is set as marker variable. If pout a * after the observed variable, then this says to estimate the loading. A @ says to fix the loading at a value. For example,
f1 BY y1@1 y2* y3*;
The defaults of M Plus make this statement equal to f1 BY y1 y2 y3;.

ON describes a regression equation. For example, y on x1 x2; regresses y onto x1 and x2. Default is to allow exogenous variables to be correlated and to estimate an error variance for an endogenous variable.

WITH describes a correlation or covariance. f1 WITH f2; means correlate f1 with f2. Default is that all exogenous variables are correlated and the correlations are estimated.

PWITH pairs the variables on the left side with those on the right side for correlations. For example

y1 y2 y3 PWITH y4 y5 y6;

means correlate y1 WITH y4; y2 WITH y5; y3 WITH y6;

To **fix the variance** of an exogenous variable, use the @ convention. For example
f1@1 f2@1 fixes the variances of the two factors f1 and f2 to be 1.0.

A list of observed or latent variables refers to the **variances or residual variances** of those variables. For example

y1 y2 y3;

refers to the variance of y1, y2 and y3 if they are exogenous variables and the residual variances if they are endogenous variables.

y1 WITH y2 allows **correlated error** between y1 and y2 of they are both exogenous variables

Means and intercepts are referred to by variables in brackets. For example

```
[y1 y2 y3];
```

refers to the means of y1 y2 and y3 if they are exogenous and the intercepts if they are endogenous. Default is to estimate all means and intercepts of observed variables and to fix latent variables means and intercepts at 0. In multiple group analysis, default is to fix means and intercepts of first group at 0 and estimate the other groups.

Parameters can be constrained to be equal by placing the same number in parentheses at the end of a line and continuing (if necessary) onto the next line. For example

```
y1 ON x1 x2 (1)
      x3;
```

regresses y1 onto x1, x2 and x3 and forces the x1 and x2 coefficients to be equal. Or,

```
y1 ON x1 (1)
      x2;
y2 ON x3 (1)
      x4;
```

regresses y1 onto x1 and x2 and y2 onto x3 and x4 and forces the coefficients for x1 and x3 to be equal.

As another example,

```
f1 BY y1 y2 y3 y4 y5 y6 (1)
      y7 y8 y9  (1);
```

forces factor loadings for y2 to y9 to be equal (the loading for y1 is set to 1.0
by default). But

```
f1 BY y1 y2 y3 y4 y5 y6
      y7 y8 y9  (1);
```

forces the loadings for y7-y9 to be equal but not the others (again with the y1 loading being fixed at 1.0).

```
f1 BY y1-y4 (1)
      y5-y6 (2)
      y7-y9 (3);
```

sets y1 as the marker variable and then forces y2-y4 to be equal, y5-y6 to be equal, y7-y9 to be equal.

[y1 y2 y3](1); **forces the means** (or intercepts) of y1 y2 and y3 **to be equal.**

To **label a parameter**, put the label after the parameter in parentheses

```
y on x1(p1)
     X2(p2)
     X3(p3);
```

This example labels the path for x1 to y as p1, x2 to y as p2 and x3 to y as p3. Only one label can appear per line.

MODEL CONSTRAINT LINE (Model Constraint:)

For **non-linear constraints on parameters**, first label the parameters, as above. Then add the statement to impose the constraint:

```
MODEL CONSTRAINT:
    p1 = p2*p3;
```

MODEL INDIRECT LINE (Model Indirect:)

MODEL INDIRECT: y3 IND x1; does an indirect effect and total effect analysis using y3 as the outcome variable and x1 as the distal variable. Must specify for each distal-outcome variable you want to examine. For example

MODEL INDIRECT: y3 IND x1; y3 IND x2; y3 IND x3

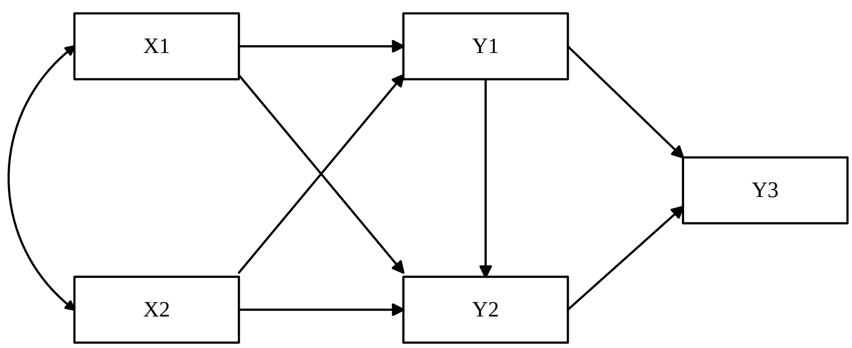
will do mediation analysis using y3 as outcome and x1 as distal variable; y3 as outcome and x2 as distal variable; and y3 as outcome and x2 as the distal variable.

If use with bootstrap, then get bootstrapped CIs

If you want to focus on specific mediated pathways among multiple mediators, specify them using VIA:

MODEL INDIRECT: y3 VIA y1 x1

This command requests all indirect effects from x1 to y3 that are mediated by y1. These include x1 to y1 to y3, and x1 to y1 to y2 to y3 in the following model:



SAVEDATA Line (Savedata:)

Saves raw data used in the analysis plus whatever else you request to be saved. Missing values are saved as a *.

FILE IS c:\temp\jim.txt; saves the results in jim.txt. Save in ascii file

FORMAT IS FREE; saves data in free format

SAVE = CPROB; saves in a mixture model the probability that an individual is in each class as well as the class number that has the highest probability for that individual.

SAVE = FSCORES; saves factor scores for latent variables.

OUTPUT Line (Output:)

Samp prints basic descriptive statistics of input variables

Stand prints standardized solution (in column StdYX); can't be used in TYPE = LOGISTIC

Residual prints the predicted covariances and the (unstandardized) residuals for the predicted and observed covariances

Mod(4) prints modification indices greater than 4.0

Cinterval prints 95% and 99% confidence intervals

Cinterval(bcbootstrap) prints 95% and 99% bootstrapped confidence intervals. To use this, you must specify bootstrap in the analysis command. You can not get modification indices with bootstrap. Can't get both regular CIS and bootstrap CIs in the same run.

Tech4 shows estimated means and covariances and correlations among the latent variables

Tech1 numbers the parameters and is good for determining how many free parameters there are in a model and where they are.