

IRT Model Comparisons in Mplus version 5.2

Here is one approach: a 2PL vs. a 1PL for Binary Responses under ML

```
TITLE: 2PL Binary Model under ML
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
          USEVARIABLES ARE dial-dia7;
          CATEGORICAL ARE dial-dia7;
          MISSING ARE .;

ANALYSIS: ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all estimated in 2PL
  IADL dial-dia7*;
! Item thresholds all free
  [dial$1-dia7$1];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;
(rest of code in 10a binary example)
```

TESTS OF MODEL FIT

Loglikelihood

H0 Value	-1454.634
----------	-----------

Information Criteria

Number of Free Parameters	14
Akaike (AIC)	2937.268
Bayesian (BIC)	2999.619
Sample-Size Adjusted BIC	2955.170
(n* = (n + 2) / 24)	

Chi-Square Test of Model Fit for the Binary and Ordered Categorical (Ordinal) Outcomes

Pearson Chi-Square Value	340.829
Degrees of Freedom	113
P-Value	0.0000
Likelihood Ratio Chi-Square Value	120.273
Degrees of Freedom	113
P-Value	0.3023

Linda Muthén suggests that if these 2 χ^2 values don't match, they should not be used to assess model fit. Further, the possible total df for the χ^2 is calculated based on # possible response patterns. Here, for 7 binary items:

2PL model: $2^{17} = 128$ possible – 7 loadings – 7 thresholds – 1 = 113
 1PL model: $2^{17} = 128$ possible – 1 loading – 7 thresholds – 1 = 119
 However, the 1PL only has df=118 because of the deleted cell.

```
TITLE: 1PL Binary Model under ML
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
          USEVARIABLES ARE dial-dia7;
          CATEGORICAL ARE dial-dia7;
          MISSING ARE .;

ANALYSIS: ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all constrained equal in 1PL
  IADL dial-dia7* (1);
! Item thresholds all free
  [dial$1-dia7$1];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;
(rest of code in 10a binary example)
```

TESTS OF MODEL FIT

Loglikelihood

H0 Value	-1464.457
----------	-----------

Information Criteria

Number of Free Parameters	8
Akaike (AIC)	2944.915
Bayesian (BIC)	2980.544
Sample-Size Adjusted BIC	2955.144
(n* = (n + 2) / 24)	

Chi-Square Test of Model Fit for the Binary and Ordered Categorical (Ordinal) Outcomes**

Pearson Chi-Square Value	296.199
Degrees of Freedom	118
P-Value	0.0000
Likelihood Ratio Chi-Square Value	126.354
Degrees of Freedom	118
P-Value	0.2828

** Of the 630 cells in the latent class indicator table, 1 were deleted in the calculation of chi-square due to extreme values.

This error message indicates that these 2 sets of chi-squares are not on the same scale. We need to test the difference in the -2LL instead:

-1454.634*-2 = 2909.258 difference = 19.946, df = 6, p = .0032
 -1464.457*-2 = 2928.914 AIC (but not BIC) is smaller for 2PL, too

Here is the same approach again: a 2PL vs. a 1PL for Polytomous Responses under ML

```
TITLE: 2PL Graded Response Model under ML
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
          USEVARIABLES ARE cial-cia7;
          CATEGORICAL ARE cial-cia7;
          MISSING ARE .;

ANALYSIS: ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all estimated in 2PL
  IADL cial-cia7*;
! Item thresholds all free
  [cial$1-cia7$1];
  [cial$2-cia7$2];
  [cial$3-cia7$3];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

(rest of code in 10c polytomous example)
```

TESTS OF MODEL FIT

Loglikelihood	H0 Value	-2523.585
---------------	----------	-----------

Information Criteria

Number of Free Parameters	28
Akaike (AIC)	5103.171
Bayesian (BIC)	5227.828
Sample-Size Adjusted BIC	5138.931
(n* = (n + 2) / 24)	

Chi-Square Test of Model Fit for the Binary and Ordered Categorical (Ordinal) Outcomes**

Pearson Chi-Square Value	1876.488
Degrees of Freedom	16317
P-Value	1.0000
Likelihood Ratio Chi-Square Value	676.937
Degrees of Freedom	16317
P-Value	1.0000

** Of the 48600 cells in the latent class indicator table, 38 were deleted in the calculation of chi-square due to extreme values.

```
TITLE: 1PL Graded Response Model under ML
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
          USEVARIABLES ARE cial-cia7;
          CATEGORICAL ARE cial-cia7;
          MISSING ARE .;

ANALYSIS: ESTIMATOR IS ML; LINK IS LOGIT;
MODEL:
! Factor loadings all constrained equal in 1PL
  IADL cial-cia7* (1);
! Item thresholds all free
  [cial$1-cia7$1];
  [cial$2-cia7$2];
  [cial$3-cia7$3];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

(rest of code in 10a polytomous example)
```

TESTS OF MODEL FIT

Loglikelihood	H0 Value	-2591.310
---------------	----------	-----------

Information Criteria

Number of Free Parameters	22
Akaike (AIC)	5226.620
Bayesian (BIC)	5324.565
Sample-Size Adjusted BIC	5254.717
(n* = (n + 2) / 24)	

Chi-Square Test of Model Fit for the Binary and Ordered Categorical (Ordinal) Outcomes**

Pearson Chi-Square Value	2650.119
Degrees of Freedom	16321
P-Value	1.0000
Likelihood Ratio Chi-Square Value	803.028
Degrees of Freedom	16321
P-Value	1.0000

** Of the 48600 cells in the latent class indicator table, 40 were deleted in the calculation of chi-square due to extreme values.

This error message indicates that these 2 sets of chi-squares are not on the same scale. We need to test the -2LL difference instead.

-2523.585*-2 = 5047.170 difference = 135.45, df = 6, p < .0001
 -2591.310*-2 = 5182.620 AIC and BIC are smaller for 2PL, too

Here is another approach: a 2PL vs. a 1PL for Binary Responses under WLSMV

```

TITLE: 2PL Binary Model under WLSMV
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
              USEVARIABLES ARE dial-dia7;
              CATEGORICAL ARE dial-dia7;
              MISSING ARE .;

ANALYSIS: ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
MODEL:
! Factor loadings all estimated in 2PL
  IADL dial-dia7*;
! Item thresholds all free
  [dial$1-dia7$1];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;      ! Standardized and local fit info
SAVEDATA:  DIFFTEST=2PL.dat; ! Save info from bigger model

```

TESTS OF MODEL FIT

Chi-Square Test of Model Fit		
Value		44.910*
Degrees of Freedom		10**
P-Value		0.0000

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus User's Guide.

** The degrees of freedom for MLMV, ULSMV and WLSMV are estimated according to a formula given in the Mplus Technical Appendices at www.statmodel.com. See degrees of freedom in the index of the Mplus User's Guide.

Chi-Square Test of Model Fit for the Baseline Model		
Value		7987.100
Degrees of Freedom		9
P-Value		0.0000

CFI/TLI		
CFI		0.996
TLI		0.996

Number of Free Parameters		14
---------------------------	--	----

RMSEA (Root Mean Square Error Of Approximation)		
Estimate		0.074

```

TITLE: 1PL Binary Model under WLSMV
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
              USEVARIABLES ARE dial-dia7;
              CATEGORICAL ARE dial-dia7;
              MISSING ARE .;

ANALYSIS: ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
              DIFFTEST=2PL.dat; ! Use saved info from bigger model
MODEL:
! Factor loadings all constrained equal in 1PL
  IADL dial-dia7* (1);
! Item thresholds all free
  [dial$1-dia7$1];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;      ! Standardized and local fit info
SAVEDATA:

```

TESTS OF MODEL FIT

Chi-Square Test of Model Fit		
Value		46.132*
Degrees of Freedom		12**
P-Value		0.0000

Chi-Square Test for Difference Testing		
Value		15.845
Degrees of Freedom		5**
P-Value		0.0073

Chi-Square Test of Model Fit for the Baseline Model		
Value		7987.100
Degrees of Freedom		9
P-Value		0.0000

CFI/TLI		
CFI		0.996
TLI		0.997

Number of Free Parameters		8
---------------------------	--	---

RMSEA (Root Mean Square Error Of Approximation)		
Estimate		0.067

The Chi-Square for Difference Testing tells us directly that the 2PL version of the binary model fits significantly better.

Here is the same approach again: a 2PL vs. a 1PL for Polytomous Responses under WLSMV

```

TITLE: 2PL Graded Response Model under WLSMV
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
              USEVARIABLES ARE cial-cia7;
              CATEGORICAL ARE cial-cia7;
              MISSING ARE .;

ANALYSIS: ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
MODEL:
! Factor loadings all estimated in 2PL
  IADL cial-cia7*;
! Item thresholds all free
  [cial$1-cia7$1];
  [cial$2-cia7$2];
  [cial$3-cia7$3];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;      ! Standardized and local fit info
SAVEDATA:   DIFFTEST=2PL.dat;    ! Save info from bigger model

```

TESTS OF MODEL FIT

Chi-Square Test of Model Fit	
Value	75.667*
Degrees of Freedom	9**
P-Value	0.0000

* The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference tests. MLM, MLR and WLSM chi-square difference testing is described in the Mplus Technical Appendices at www.statmodel.com. See chi-square difference testing in the index of the Mplus User's Guide.

** The degrees of freedom for MLMV, ULSMV and WLSMV are estimated according to a formula given in the Mplus Technical Appendices at www.statmodel.com. See degrees of freedom in the index of the Mplus User's Guide.

Chi-Square Test of Model Fit for the Baseline Model	
Value	14637.676
Degrees of Freedom	6
P-Value	0.0000

CFI/TLI	
CFI	0.995
TLI	0.997

Number of Free Parameters	28
---------------------------	----

RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.108

```

TITLE: 1PL Graded Response Model under WLSMV
DATA: FILE IS ADL.dat;
VARIABLE: NAMES ARE case dial-dia7 cial-cia7;
              USEVARIABLES ARE cial-cia7;
              CATEGORICAL ARE cial-cia7;
              MISSING ARE .;

ANALYSIS: ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
              DIFFTEST=2PL.dat; ! Use saved info from bigger model
MODEL:
! Factor loadings all constrained equal in 1PL
  IADL cial-cia7* (1);
! Item thresholds all free
  [cial$1-cia7$1];
  [cial$2-cia7$2];
  [cial$3-cia7$3];
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;      ! Standardized and local fit info
SAVEDATA:

```

TESTS OF MODEL FIT

Chi-Square Test of Model Fit	
Value	134.741*
Degrees of Freedom	9**
P-Value	0.0000

Chi-Square Test for Difference Testing	
Value	75.584
Degrees of Freedom	4**
P-Value	0.0000

Chi-Square Test of Model Fit for the Baseline Model	
Value	14637.676
Degrees of Freedom	6
P-Value	0.0000

CFI/TLI	
CFI	0.991
TLI	0.994

Number of Free Parameters	22
---------------------------	----

RMSEA (Root Mean Square Error Of Approximation)	
Estimate	0.148

The Chi-Square for Difference Testing tells us directly that the 2PL version of the polytomous model fits significantly better.

Here are the parameter estimates under WLSMV Theta Parameterization for the 2PL version of binary (left) and polytomous (right)

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
LOADINGS					
IADL	BY				
	DIA1	2.686	0.317	8.461	0.000
	DIA2	2.941	0.493	5.966	0.000
	DIA3	2.803	0.384	7.290	0.000
	DIA4	3.654	0.575	6.356	0.000
	DIA5	2.486	0.294	8.449	0.000
	DIA6	1.991	0.223	8.940	0.000
	DIA7	1.571	0.299	5.246	0.000

Thresholds

DIA1\$1	-1.004	0.179	-5.607	0.000
DIA2\$1	-3.097	0.481	-6.444	0.000
DIA3\$1	-2.221	0.307	-7.240	0.000
DIA4\$1	-1.581	0.298	-5.312	0.000
DIA5\$1	-1.057	0.174	-6.071	0.000
DIA6\$1	-1.391	0.166	-8.359	0.000
DIA7\$1	-2.946	0.398	-7.401	0.000

IRT PARAMETERIZATION IN TWO-PARAMETER LOGISTIC METRIC
WHERE THE PROBIT IS DISCRIMINATION*(THETA - DIFFICULTY)

Item Discriminations

IADL	BY				
	DIA1	2.686	0.317	8.461	0.000
	DIA2	2.941	0.493	5.966	0.000
	DIA3	2.803	0.384	7.290	0.000
	DIA4	3.654	0.575	6.356	0.000
	DIA5	2.486	0.294	8.449	0.000
	DIA6	1.991	0.223	8.940	0.000
	DIA7	1.571	0.299	5.246	0.000

Item Difficulties

DIA1\$1	-0.374	0.055	-6.743	0.000
DIA2\$1	-1.053	0.069	-15.360	0.000
DIA3\$1	-0.792	0.062	-12.863	0.000
DIA4\$1	-0.433	0.054	-7.982	0.000
DIA5\$1	-0.425	0.056	-7.607	0.000
DIA6\$1	-0.699	0.063	-11.084	0.000
DIA7\$1	-1.875	0.154	-12.191	0.000

Residuals for Covariances/Correlations/Residual Correlations

	DIA1	DIA2	DIA3	DIA4	DIA5	DIA6
DIA1						
DIA2	0.028					
DIA3	0.038	0.029				
DIA4	-0.022	-0.040	-0.046			
DIA5	-0.032	-0.034	-0.103	0.029		
DIA6	-0.052	-0.056	-0.046	0.026	0.032	
DIA7	-0.112	-0.003	0.010	0.031	-0.027	0.064

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
LOADINGS					
IADL	BY				
	CIA1	3.655	0.330	11.083	0.000
	CIA2	3.346	0.388	8.632	0.000
	CIA3	2.923	0.269	10.881	0.000
	CIA4	3.286	0.299	11.008	0.000
	CIA5	2.222	0.159	13.963	0.000
	CIA6	1.907	0.169	11.305	0.000
	CIA7	1.075	0.130	8.279	0.000

Thresholds

CIA1\$1	-5.151	0.424	-12.137	0.000
CIA1\$2	-3.658	0.347	-10.534	0.000
CIA1\$3	-0.734	0.217	-3.383	0.001
CIA2\$1	-5.096	0.497	-10.254	0.000
CIA2\$2	-4.253	0.445	-9.552	0.000
CIA2\$3	-2.620	0.353	-7.425	0.000
CIA3\$1	-4.193	0.327	-12.825	0.000
CIA3\$2	-3.404	0.296	-11.486	0.000
CIA3\$3	-1.761	0.232	-7.592	0.000
CIA4\$1	-4.379	0.342	-12.794	0.000
CIA4\$2	-2.987	0.269	-11.107	0.000
CIA4\$3	-1.024	0.211	-4.863	0.000
CIA5\$1	-3.866	0.233	-16.616	0.000
CIA5\$2	-1.892	0.160	-11.856	0.000
CIA5\$3	-0.425	0.130	-3.277	0.001
CIA6\$1	-3.450	0.235	-14.697	0.000
CIA6\$2	-2.354	0.184	-12.805	0.000
CIA6\$3	-1.400	0.154	-9.072	0.000
CIA7\$1	-3.282	0.249	-13.169	0.000
CIA7\$2	-2.577	0.181	-14.231	0.000
CIA7\$3	-1.757	0.137	-12.840	0.000

IRT RESULTS ARE NOT GIVEN FOR POLYTOMOUS ITEMS;
THEY MUST BE CALCULATED BY YOU!

Residuals for Covariances/Correlations/Residual Correlations

	CIA1	CIA2	CIA3	CIA4	CIA5	CIA6
CIA1						
CIA2	0.013					
CIA3	0.012	0.017				
CIA4	-0.010	-0.025	-0.036			
CIA5	-0.030	-0.045	-0.067	0.032		
CIA6	-0.040	-0.055	-0.025	0.026	0.035	
CIA7	-0.026	-0.007	0.016	0.022	-0.031	0.025