Binary IFA-IRT Models in Mplus version 7.4

Example data: 635 older adults (age 80-100) self-reporting on 7 items assessing the Instrumental Activities of Daily Living (IADL) as follows:

1. Housework (cleaning and laundry): 1=64%
2. Bedmaking: 1=84%
3. Cooking: 1=77%
4. Everyday shopping: 1=66%
5. Getting to places outside of walking distance: 1=65%
6. Handling banking and other business: 1=73%
7. Using the telephone 1=94%

Two versions of a response format were available:

Binary → 0 = "needs help", 1 = "does not need help"
Categorical → 0 = "can't do it", 1="big problems", 2="some problems", 3="no problems"

Higher scores indicate greater function. We will look at each response format in turn.

Binary 2-PL Model Syntax (left) and 1-PL Model Syntax (right) using ML and a logit scale:

```
TITLE:   Assess binary IADL items using 2PL
DATA:    FILE IS ADL.dat;

VARIABLE:    NAMES ARE case dia1-dia7 cia1-cia7;
             USEVARIABLES ARE dia1-dia7;
             CATEGORICAL ARE dia1-dia7;
             MISSING ARE .;
             IDVARIABLE IS case;

ANALYSIS:    ESTIMATOR IS ML;
             LINK IS LOGIT;

MODEL:
! Factor loadings all estimated in 2PL
    IADL BY dia1-dia7*;
! Item thresholds all estimated
    [dia1$1-dia7$1*];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

OUTPUT:      STDYX;               ! Standardized solution
             RESIDUAL TECH10;     ! Local fit info

SAVEDATA:    SAVE = FSCORES;            ! Save factor scores (thetas)
             FILE = IADL_2PLThetas.dat; ! File factor scores saved to

PLOT:    TYPE IS PLOT1;     ! PLOT1 gets you sample descriptives
         TYPE IS PLOT2;     ! PLOT2 gets you the IRT-relevant curves
         TYPE IS PLOT3;     ! PLOT3 gets you descriptives for theta
```

```
TITLE:   Assess binary IADL items using 1PL
DATA:    FILE IS ADL.dat;

VARIABLE:    NAMES ARE case dia1-dia7 cia1-cia7;
             USEVARIABLES ARE dia1-dia7;
             CATEGORICAL ARE dia1-dia7;
             MISSING ARE .;
             IDVARIABLE IS case;

ANALYSIS:    ESTIMATOR IS ML;
             LINK IS LOGIT;

MODEL:
! Factor loadings all held equal in 1PL
    IADL BY dia1-dia7* (loading);
! Item thresholds all estimated
    [dia1$1-dia7$1*];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

OUTPUT:      STDYX;               ! Standardized solution
             RESIDUAL TECH10;     ! Local fit info

SAVEDATA:    SAVE = FSCORES;            ! Save factor scores (thetas)
             FILE = IADL_1PLThetas.dat; ! File factor scores saved to

PLOT:    TYPE IS PLOT1;     ! PLOT1 gets you sample descriptives
         TYPE IS PLOT2;     ! PLOT2 gets you the IRT-relevant curves
         TYPE IS PLOT3;     ! PLOT3 gets you descriptives for theta
```

Binary 2-PL Model Fit (left) and 1-PL Model Fit (right) using ML logit:

MODEL FIT INFORMATION – 2PL

Number of Free Parameters 14

Loglikelihood

H0 Value -1454.634

Information Criteria

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

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: 2^7 = 128 possible – 7 loadings – 7 thresholds – 1 = 113
1PL model: 2^7 = 128 possible – 1 loading – 7 thresholds – 1 = 119

However, the 1PL only has df=118 because of the deleted cell.

MODEL FIT INFORMATION – 1 PL

Number of Free Parameters 8

Loglikelihood

H0 Value -1464.457

Information Criteria

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

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 for the 2-PL and 1-PL are not on the same scale because they are not based on the same data. So we can't compare the chi-squares to test the difference in model fit, but we can still compare LL values.

Does the 2-PL fit better than the 1-PL?

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

3 differently scaled 2-PL solutions from ML logit provided by Mplus – all provide the exact same model predictions!

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS = CHANGE IN LOGIT(Y=1) PER UNIT CHANGE IN THETA				
IADL BY				
DIA1	4.328	0.560	7.725	0.000
DIA2	4.978	0.808	6.159	0.000
DIA3	4.323	0.570	7.579	0.000
DIA4	7.511	1.696	4.429	0.000
DIA5	4.248	0.527	8.062	0.000
DIA6	3.451	0.401	8.600	0.000
DIA7	3.283	0.601	5.467	0.000
THRESHOLDS = EXPECTED LOGIT(Y=0) WHEN THETA IS 0				
DIA1$1	-1.629	0.295	-5.516	0.000
DIA2$1	-5.202	0.770	-6.754	0.000
DIA3$1	-3.462	0.441	-7.842	0.000
DIA4$1	-3.120	0.744	-4.193	0.000
DIA5$1	-1.833	0.298	-6.158	0.000
DIA6$1	-2.442	0.292	-8.368	0.000
DIA7$1	-5.962	0.858	-6.951	0.000

STDYX MODEL RESULTS (STANDARDIZED IFA MODEL SOLUTION)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS IN STANDARDIZED METRIC = loading*SD(Theta)/SD(Y)				
IADL BY				
DIA1	0.922	0.018	51.712	0.000
DIA2	0.940	0.018	52.557	0.000
DIA3	0.922	0.018	50.622	0.000
DIA4	0.972	0.012	80.380	0.000
DIA5	0.920	0.018	52.291	0.000
DIA6	0.885	0.022	39.729	0.000
DIA7	0.875	0.037	23.380	0.000
THRESHOLDS IN STANDARDIZED METRIC = threshold/SD(Y)				
DIA1$1	-0.347	0.048	-7.303	0.000
DIA2$1	-0.982	0.056	-17.409	0.000
DIA3$1	-0.739	0.051	-14.373	0.000
DIA4$1	-0.404	0.045	-8.928	0.000
DIA5$1	-0.397	0.048	-8.348	0.000
DIA6$1	-0.626	0.050	-12.558	0.000
DIA7$1	-1.590	0.080	-19.949	0.000
R-SQUARE = standardized loading^2				
DIA1	0.851	0.033	25.856	0.000
DIA2	0.883	0.034	26.278	0.000
DIA3	0.850	0.034	25.311	0.000
DIA4	0.945	0.024	40.190	0.000
DIA5	0.846	0.032	26.145	0.000
DIA6	0.784	0.039	19.865	0.000
DIA7	0.766	0.066	11.690	0.000

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

Item Discriminations = SLOPE OF ICC AT P=.50				
IADL BY				
DIA1	4.328	0.560	7.725	0.000
DIA2	4.978	0.808	6.159	0.000
DIA3	4.323	0.570	7.579	0.000
DIA4	7.511	1.696	4.429	0.000
DIA5	4.248	0.527	8.062	0.000
DIA6	3.451	0.401	8.600	0.000
DIA7	3.283	0.601	5.467	0.000
Item Difficulties = LOCATION OF ITEM ON LATENT TRAIT at P=.50, LOGIT=0				
DIA1$1	-0.376	0.052	-7.298	0.000
DIA2$1	-1.045	0.065	-15.978	0.000
DIA3$1	-0.801	0.059	-13.562	0.000
DIA4$1	-0.415	0.047	-8.849	0.000
DIA5$1	-0.432	0.052	-8.296	0.000
DIA6$1	-0.708	0.060	-11.889	0.000
DIA7$1	-1.816	0.126	-14.454	0.000

USING RESULTS FROM IFA MODEL (LEFT PANEL):

IFA model: Logit(y) = -threshold + loading(Theta)
Threshold = expected logit of (y=0) for someone with Theta=0
When *-1, threshold becomes intercept: expected logit for (y=1) instead

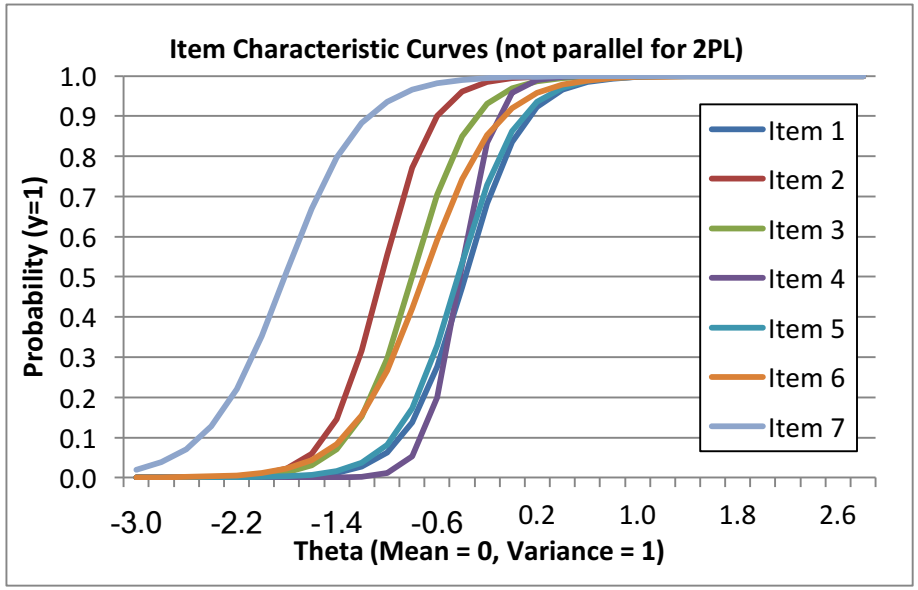
Loading = regression of item logit on Theta
= change in logit(y) for a one-unit change in Theta

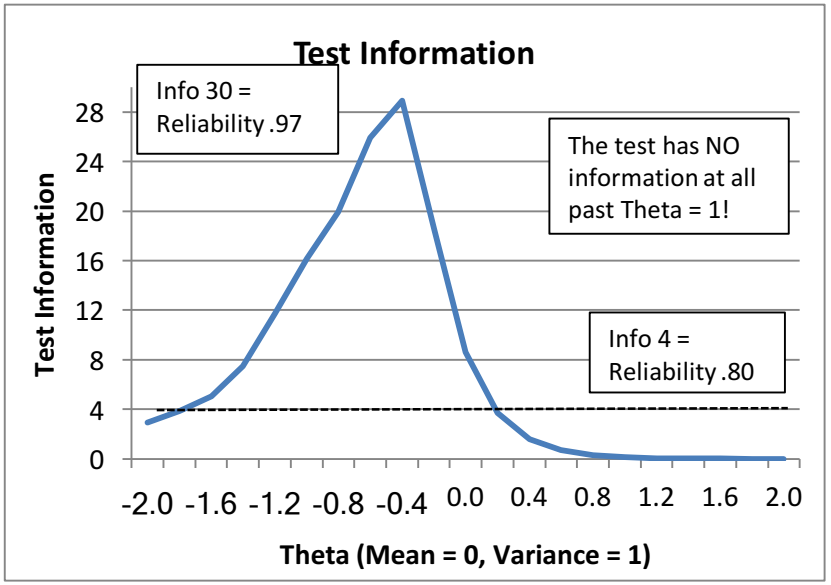
IFA Models:
Logit (DIA1=1) = 1.629 + 4.328(Theta) → if Theta=0, prob(y=1)= .836
Logit (DIA7=1) = 5.962 + 3.283(Theta) → if Theta=0, prob(y=1)= .997

USING RESULTS FROM IRT MODEL (RIGHT PANEL):

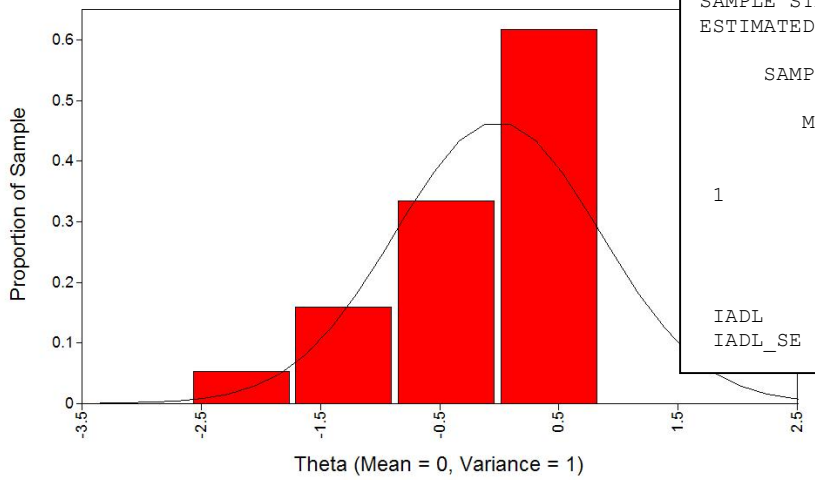
IRT model: Logit(y=1) = a(theta – difficulty)
a = discrimination (rescaled slope) = loading/1.7
b = difficulty (location on latent metric) = threshold/loading

IRT Models:
Logit (DIA1=1) = 4.328*(Theta - -0.376)→ if Theta=0, prob(y=1)= .836
Logit (DIA7=1) = 3.283*(Theta - -1.816)→ if Theta=0, prob(y=1)= .997





Distribution of Theta under 2 PL (made in Mplus): Although reliability is > .80 from −1.5 to 0.3 or so, we see a huge ceiling effect: most respondents can do all the tasks.



SAMPLE STATISTICS FOR ESTIMATED FACTOR SCORES

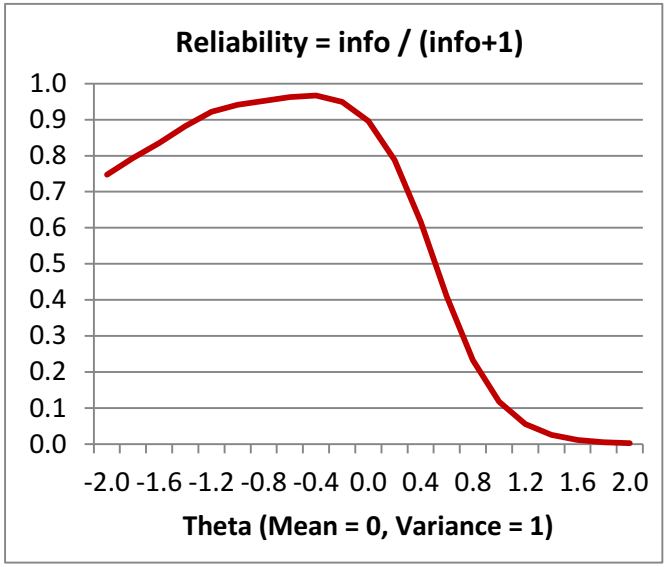
SAMPLE STATISTICS

Means

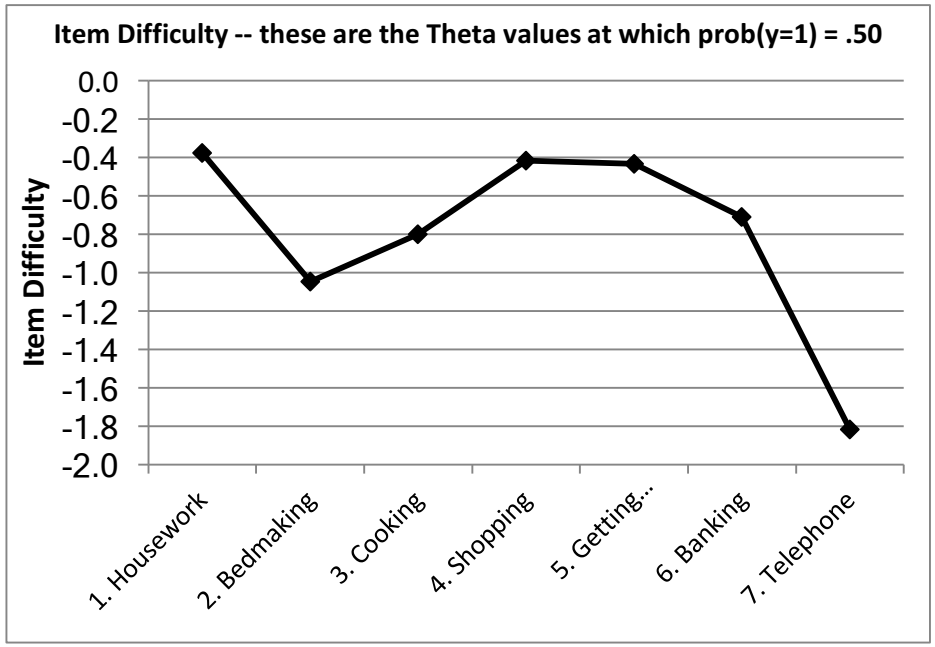
	IADL	IADL_SE
1	**-0.009**	0.471

Covariances

	IADL	IADL_SE
IADL	**0.741**	
IADL_SE	0.124	0.038



The estimated theta scores are supposed to have a mean of 0 and a variance of 1, but this table shows that they have a variance of only .741 instead. Such shrinkage is why it can be problematic to use these estimated theta scores as observed variables in other analyses.



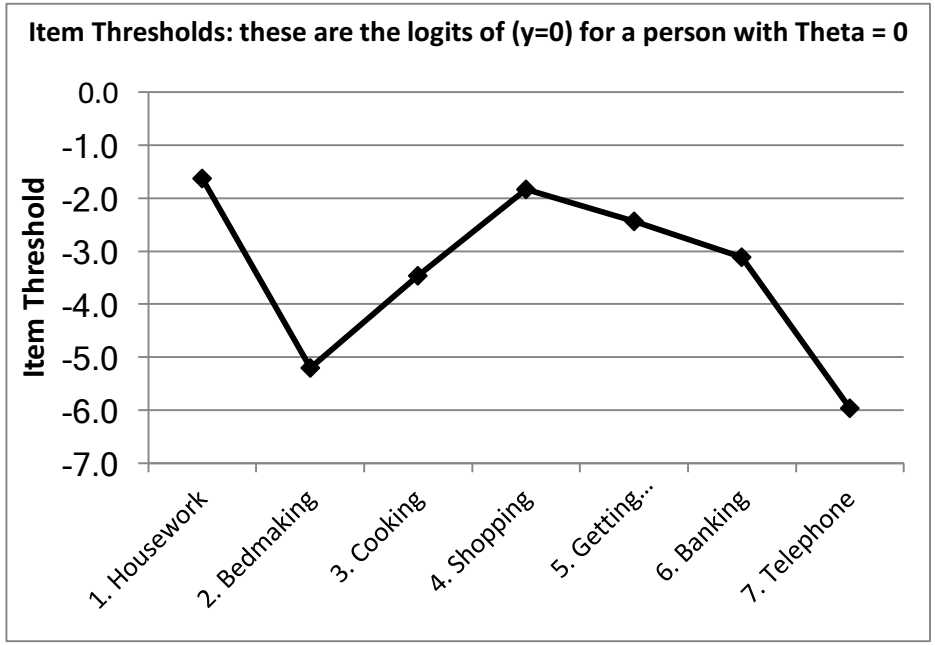
Plots of item parameters and predicted probabilities of item responses (made in excel):

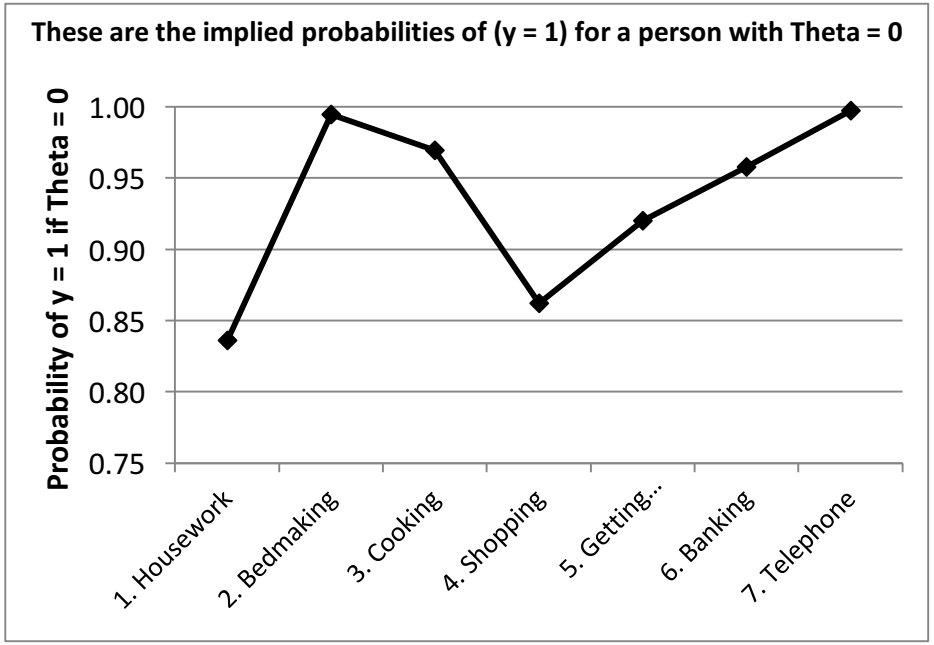
Top Left: Note that no items are available to measure above-average abilities well! The item difficulty for most items covers values of Theta between −1.0 to −0.5.

Bottom Left: These are the thresholds for each item, or the logit of (y=0) if Theta=0. These are hard to interpret as is….

Bottom Right: These are the probability of y=1 if Theta=0, as given by 1 – [exp(threshold) / (1+(exp(threshold))]

See excel workbook for calculations and plots





Here is another estimation approach: a 2PL vs. a 1PL for Binary Responses using WLSMV Probit model

```
TITLE:   2PL Binary Model under WLSMV
DATA:    FILE IS ADL.dat;
VARIABLE:    NAMES ARE case dia1-dia7 cia1-cia7;
             USEVARIABLES ARE dia1-dia7;
             CATEGORICAL ARE dia1-dia7;
             MISSING ARE .; IDVARIABLE is case;
ANALYSIS:    ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
MODEL:
! Factor loadings all estimated in 2PL
    IADL BY dia1-dia7*;
! Item thresholds all estimated
    [dia1$1-dia7$1*];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;    ! Standardized solution, local fit
SAVEDATA:    DIFFTEST=2PL.dat;  ! Save info from bigger model
             SAVE = FSCORES;           ! Save factor scores (thetas)
             FILE IS IADL_2PLThetas.dat; ! File factor scores saved to

PLOT:    TYPE IS PLOT1 PLOT2 PLOT3;    ! Get IRT plots
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       14

Chi-Square Test of Model Fit
          Value                             54.820*
          Degrees of Freedom                   14
          P-Value                           0.0000

*   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV
cannot be used for chi-square difference testing in the regular way.
MLM, MLR and WLSM chi-square difference testing is described on the
Mplus website.  MLMV, WLSMV, and ULSMV difference testing is done
using the DIFFTEST option.

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.068
          90 Percent C.I.                    0.049  0.087
          Probability RMSEA <= .05           0.055

CFI/TLI
          CFI                                0.997
          TLI                                0.995

Chi-Square Test of Model Fit for the Baseline Model
          Value                          12351.798
          Degrees of Freedom                   21
          P-Value                           0.0000
WRMR (Weighted Root Mean Square Residual)
          Value                              1.160
```

```
TITLE:   1PL Binary Model under WLSMV
DATA:    FILE IS ADL.dat;
VARIABLE:    NAMES ARE case dia1-dia7 cia1-cia7;
             USEVARIABLES ARE dia1-dia7;
             CATEGORICAL ARE dia1-dia7;
             MISSING ARE .; IDVARIABLE is case;

ANALYSIS:    ESTIMATOR IS WLSMV; PARAMETERIZATION IS THETA;
             DIFFTEST=2PL.dat;  ! Use saved info from bigger model
MODEL:
! Factor loadings all equal in 1PL
    IADL BY dia1-dia7* (loading);
! Item thresholds all estimated
    [dia1$1-dia7$1*];
! Factor mean=0 and variance=1 for identification
    [IADL@0]; IADL@1;

OUTPUT:      STDYX Residual;    ! Standardized solution, local fit
SAVEDATA:
             SAVE = FSCORES;           ! Save factor scores (thetas)
             FILE IS IADL_1PLThetas.dat; ! File factor scores saved to

PLOT:    TYPE IS PLOT1 PLOT2 PLOT3;    ! Get IRT plots
```

```
MODEL FIT INFORMATION

Number of Free Parameters                        8

Chi-Square Test of Model Fit
          Value                             64.889*
          Degrees of Freedom                   20
          P-Value                           0.0000

Chi-Square Test for Difference Testing
          Value                             18.450
          Degrees of Freedom                     6
          P-Value                           0.0052

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.059
          90 Percent C.I.                    0.044  0.076
          Probability RMSEA <= .05           0.154

CFI/TLI
          CFI                                0.996
          TLI                                0.996
WRMR (Weighted Root Mean Square Residual)
          Value                              1.501
```

The Chi-Square for Difference Testing tells us directly that the 2PL version of the binary model fits significantly better (now under WLSMV, same as it did under ML).

Here are the parameter estimates under WLSMV Theta Parameterization (Probit) for the 2PL version of binary items

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS = CHANGE IN PROBIT(Y=1) PER UNIT CHANGE IN THETA				
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 = EXPECTED PROBIT(Y=0) WHEN THETA IS 0				
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

STDYX MODEL RESULTS (STANDARDIZED IFA MODEL SOLUTION)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS IN STANDARDIZED METRIC = loading*SD(Theta)/SD(Y)				
IADL BY				
DIA1	0.937	0.013	69.487	0.000
DIA2	0.947	0.016	57.551	0.000
DIA3	0.942	0.015	64.551	0.000
DIA4	0.965	0.011	91.196	0.000
DIA5	0.928	0.015	60.671	0.000
DIA6	0.894	0.020	44.371	0.000
DIA7	0.844	0.046	18.195	0.000
Thresholds IN STANDARDIZED METRIC = threshold/SD(Y)				
DIA1$1	-0.350	0.052	-6.790	0.000
DIA2$1	-0.997	0.061	-16.474	0.000
DIA3$1	-0.746	0.056	-13.326	0.000
DIA4$1	-0.417	0.052	-8.041	0.000
DIA5$1	-0.395	0.051	-7.676	0.000
DIA6$1	-0.624	0.054	-11.648	0.000
DIA7$1	-1.582	0.081	-19.628	0.000
R-SQUARE = standardized loading^2				
DIA1	0.878	0.025	34.744	0.000
DIA2	0.896	0.031	28.775	0.000
DIA3	0.887	0.027	32.276	0.000
DIA4	0.930	0.020	45.598	0.000
DIA5	0.861	0.028	30.336	0.000
DIA6	0.799	0.036	22.185	0.000
DIA7	0.712	0.078	9.097	0.000

IRT PARAMETERIZATION IN TWO-PARAMETER PROBIT 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

Logit = 1.7*probit, or Probit = Logit/1.7

IFA model: PROBIT(y) = –threshold + loading(Theta)
Threshold = expected probit of (y=0) for someone with Theta=0
When *-1, threshold → intercept: expected probit for (y=1) instead
Loading = regression of item probit on Theta

IRT model: Probit(y=1) = a(theta – difficulty)
a = discrimination (rescaled slope) = loading/1
b = difficulty (location on latent metric) = threshold/loading

LOCAL FIT VIA STANDARDIZED RESIDUAL CORRELATIONS LEFTOVER TETRACHORIC CORRELATION (HOW FAR OFF FROM DATA)

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**