

### Graded Response Polytomous IRT Models in Mplus version 5.2

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)
2. Bedmaking
3. Cooking
4. Everyday shopping
5. Getting to places outside of walking distance
6. Handling banking and other business
7. Using the telephone

Item	0=Can't Do It	1=Big Problems	2=Some Problems	3=Can Do It
1	0.09	0.08	0.26	0.58
2	0.07	0.04	0.12	0.77
3	0.09	0.05	0.15	0.72
4	0.10	0.09	0.19	0.62
5	0.06	0.16	0.21	0.57
6	0.06	0.08	0.12	0.74
7	0.01	0.03	0.08	0.88

### Graded Response Model Syntax and Truncated Output:

```

TITLE: Assess polytomous IADL items
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
  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;           ! Standardized solution
              TECH10;        ! Local fit info

SAVEDATA: SAVE = FSCORES;      ! Save factor scores (thetas)
              FILE IS IADL_4Thetas.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
    
```

```

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.
    
```

**And that's it for model fit! There are problems with using the ML chi-square values to assess differences in fit between models. Stay tuned for more on this...**

**3 Different Solutions (2 given, 1 calculated) – all provide the exact same predictions!**

**UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)**

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS = SLOPE RELATION OF LOGIT(Y) TO THETA				
IADL BY				
CIA1	6.846	0.841	8.140	0.000
CIA2	5.200	0.555	9.363	0.000
CIA3	4.613	0.456	10.119	0.000
CIA4	5.701	0.612	9.312	0.000
CIA5	3.556	0.298	11.950	0.000
CIA6	2.897	0.261	11.094	0.000
CIA7	1.778	0.209	8.512	0.000
Thresholds = EXPECTED LOGIT(Y=0) WHEN THETA IS 0 (MEAN OF SAMPLE)				
CIA1\$1	-9.808	1.138	-8.620	0.000
CIA1\$2	-6.460	0.799	-8.088	0.000
CIA1\$3	-1.238	0.384	-3.226	0.001
CIA2\$1	-8.145	0.794	-10.257	0.000
CIA2\$2	-6.313	0.618	-10.219	0.000
CIA2\$3	-3.737	0.441	-8.480	0.000
CIA3\$1	-6.841	0.613	-11.162	0.000
CIA3\$2	-5.194	0.480	-10.810	0.000
CIA3\$3	-2.572	0.330	-7.792	0.000
CIA4\$1	-7.454	0.747	-9.975	0.000
CIA4\$2	-4.635	0.514	-9.026	0.000
CIA4\$3	-1.426	0.327	-4.366	0.000
CIA5\$1	-6.578	0.494	-13.314	0.000
CIA5\$2	-3.041	0.273	-11.155	0.000
CIA5\$3	-0.681	0.203	-3.354	0.001
CIA6\$1	-5.538	0.411	-13.486	0.000
CIA6\$2	-3.583	0.285	-12.554	0.000
CIA6\$3	-2.044	0.219	-9.344	0.000
CIA7\$1	-5.810	0.472	-12.315	0.000
CIA7\$2	-4.398	0.322	-13.673	0.000
CIA7\$3	-2.951	0.237	-12.457	0.000

**STDYX MODEL RESULTS (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				
CIA1	0.967	0.008	124.093	0.000
CIA2	0.944	0.011	86.315	0.000
CIA3	0.931	0.012	75.583	0.000
CIA4	0.953	0.009	101.294	0.000
CIA5	0.891	0.015	57.872	0.000
CIA6	0.848	0.022	39.402	0.000
CIA7	0.700	0.042	16.689	0.000

(rest omitted)

**USING RESULTS FROM IFA MODEL (LEFT PANEL):**

IFA model: Logit(y=1) = -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

For 4-category responses, the sub-models look like this:

Logit(y= 0 vs 123) = -threshold\$1 + loading(Theta)  
 Logit(y= 01 vs 23) = -threshold\$2 + loading(Theta)  
 Logit(y= 012 vs 3) = -threshold\$3 + loading(Theta)

IFA Models:

\$1 Logit(CIA1=0 vs 123)= 9.808 + 6.846(Theta) → if Theta=0, prob=.99994  
 \$2 Logit(CIA1=01 vs 23)= 6.460 + 6.846(Theta) → if Theta=0, prob=.99844  
 \$3 Logit(CIA1=012 vs 3)= 1.238 + 6.846(Theta) → if Theta=0, prob=.77522

\$1 → if Theta=-1, logit= 2.962, prob= .95083  
 \$2 → if Theta=-1, logit= -0.386, prob= .40468  
 \$3 → if Theta=-1, logit= -5.608 prob= .00365

**RESULTS FROM IRT MODEL MUST BE CALCULATED BY YOU!**

IRT model: Logit(y) = 1.7a(theta - difficulty)

a = discrimination (rescaled slope) = loading/1.7  
 b = difficulty (location on latent metric) = threshold/loading

My calculations (see spreadsheet):

CIA1 loading = 6.846/1.7 → a discrimination = 4.027  
 CIA1 threshold\$1 = -9.808/6.846 → b difficulty\$1 = -1.433  
 CIA1 threshold\$2 = -6.460/6.846 → b difficulty\$2 = -0.944  
 CIA1 threshold\$3 = -1.238/6.846 → b difficulty\$3 = -0.181

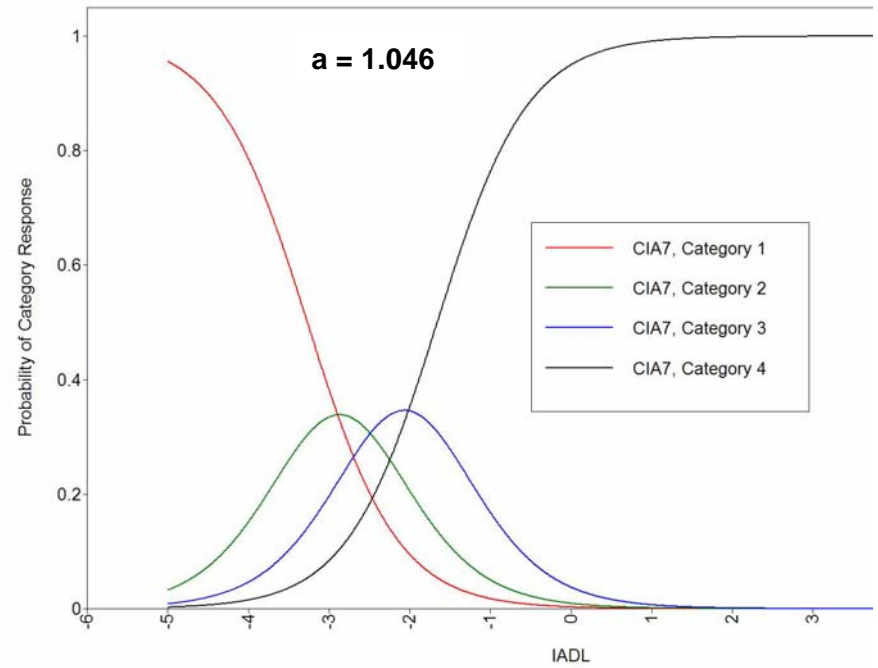
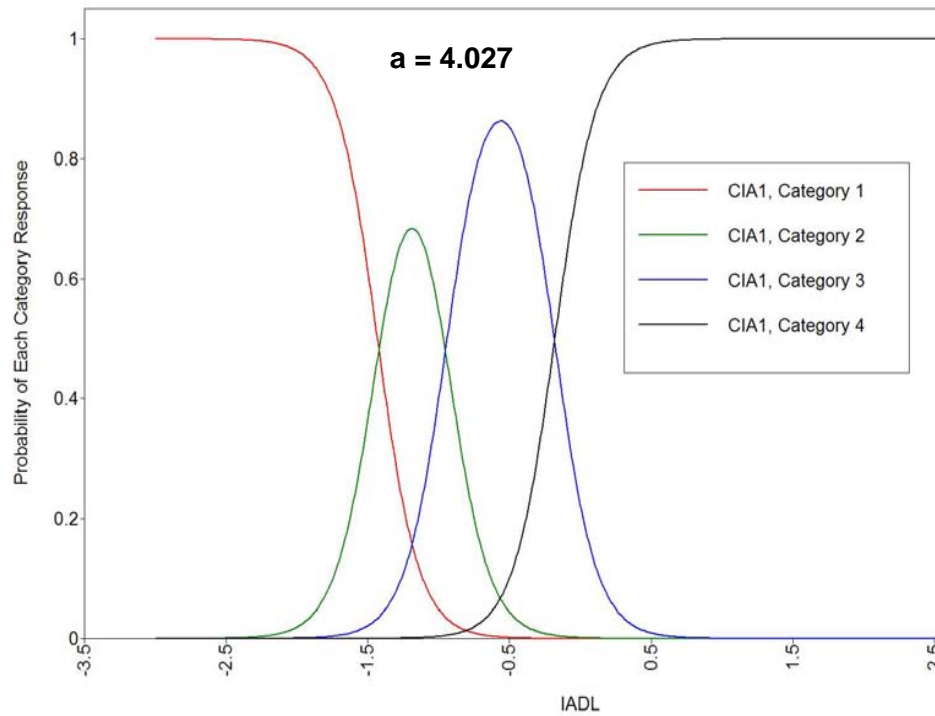
For 4-category responses, the sub-models look like this:

\$1 Logit(y= 0 vs 123) = 1.7a(theta - difficulty\$1)  
 \$2 Logit(y= 01 vs 23) = 1.7a(theta - difficulty\$2)  
 \$3 Logit(y= 012 vs 3) = 1.7a(theta - difficulty\$3)

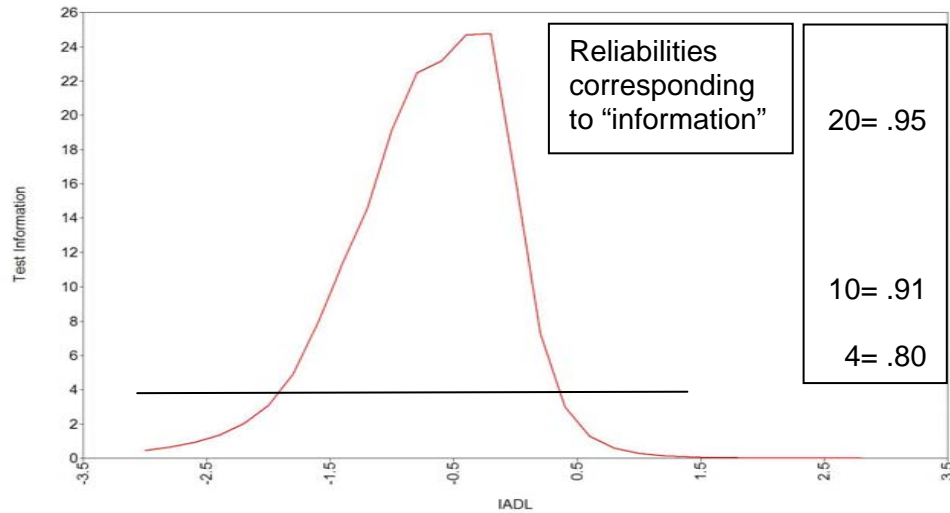
IRT Models:

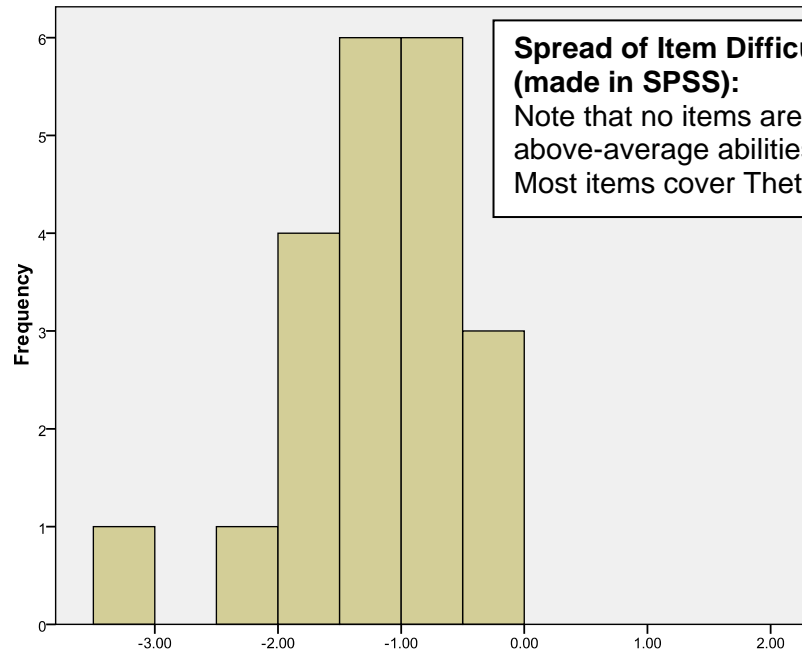
\$1 Logit = 1.7(4.027)(Theta - -1.433)  
 \$2 Logit = 1.7(4.027)(Theta - -0.944)  
 \$3 Logit = 1.7(4.027)(Theta - -0.181)

**Mplus Category Response Curves – Item 1 (good and steep discrimination) and Item 7 (less good because less steep)**



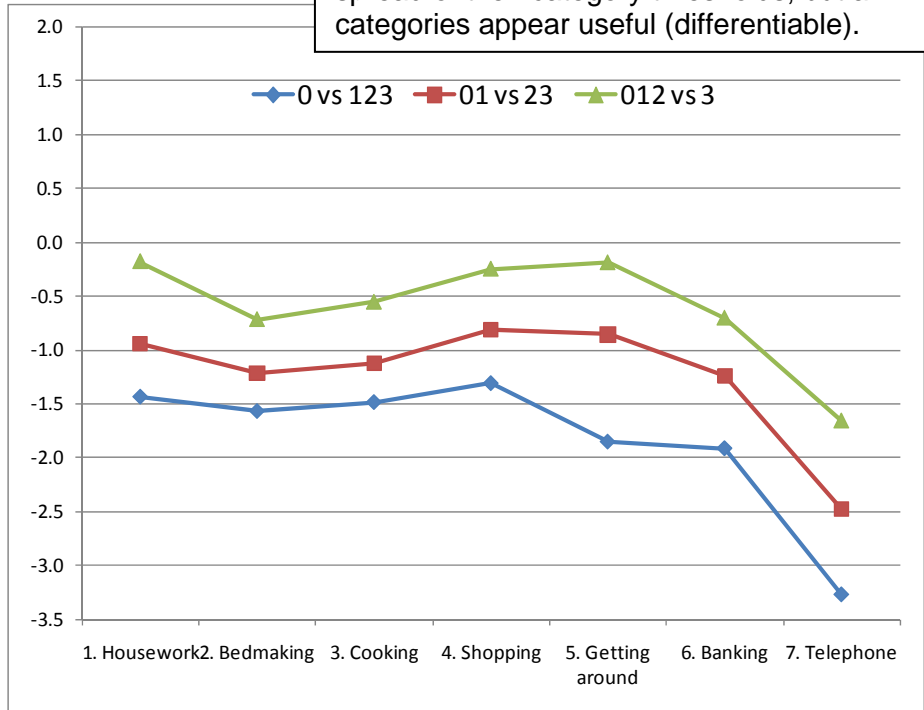
**Mplus Test Information Curve (Y axis is “information”)  
Test is most sensitive near low-range theta**



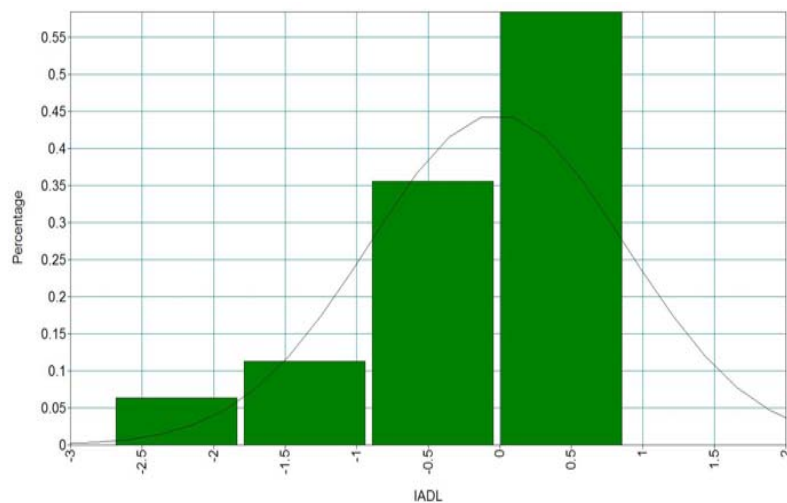


**Spread of Item Difficulties (made in SPSS):**  
 Note that no items are available to measure above-average abilities well!  
 Most items cover Theta of -1.5 to 0.

**Variability in Spread of Item Difficulty (made in excel):**  
 Note that some items (like 5) have a wider spread of their category thresholds, but all categories appear useful (differentiable).



**Distribution of Theta under GRM (made in Mplus)**



Although reliability is above .80 from about -2.0 to 0.4 or so, we still see a huge ceiling effect – most of our sample can do all the tasks.

If we are concerned about measuring the higher end of theta better, then we'd need additional more difficult items for sure!