Example 7c: Testing Multiple-Group Measurement Invariance in Item Factor Models using WLSMV in M*plus* v. 8.10 *(complete syntax and output for Mplus and R lavaan is available electronically, but their results don't match)*

This example comes from the Octogenarian Twin Study of Aging in Sweden. The current analysis includes 634 older adults (age 80–100 years) on 7 items measuring the Instrumental Activities of Daily Living (IADL), as also used in Example 6a. We are examining differences for men (*N*=214) versus women (*N*=420) using **four** response options →

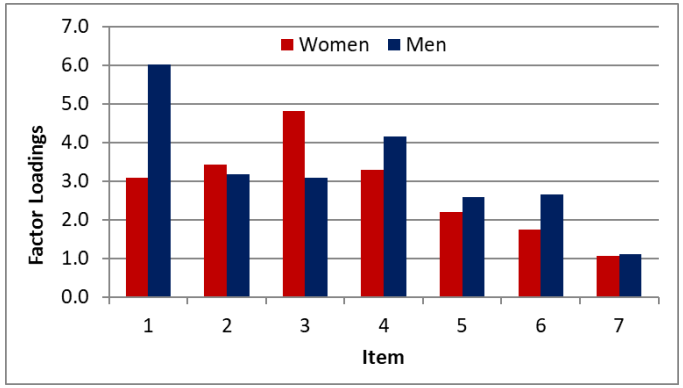
	Item	0=Can't Do It		1=Big Problems		2=Some Problems		3=Can Do It	
		Men	Women	Men	Women	Men	Women	Men	Women
1	Housework	0.11	0.08	0.09	0.08	0.25	0.26	0.56	0.59
2	Bedmaking	0.09	0.06	0.04	0.04	0.12	0.11	0.75	0.78
3	Cooking	0.12	0.07	0.08	0.04	0.24	0.11	0.57	0.78
4	Everyday shopping	0.11	0.10	0.05	0.11	0.14	0.22	0.70	0.58
5	Walking to places	0.05	0.06	0.12	0.18	0.16	0.24	0.67	0.52
6	Handling banking	0.05	0.06	0.07	0.09	0.10	0.13	0.78	0.72
7	Using a telephone	0.02	0.01	0.04	0.02	0.06	0.08	0.88	0.89

Multiple Group IFA Model Syntax and Truncated Output:

```
TITLE:     Assess invariance for polytomous ordinal IADL items using WLSMV
DATA:      FILE = Example7cd.csv; ! Don't need path if data in same folder
           FORMAT = free; TYPE = INDIVIDUAL; ! Defaults

VARIABLE:  NAMES = case female dia1-dia7 cia1-cia7;  ! All vars in data
           USEVARIABLES = cia1-cia7;          ! All vars in model
           CATEGORICAL = cia1-cia7;           ! All ordinal outcomes
           GROUPING = female (0=Men 1=Women); ! Grouping variable
           MISSING = ALL (9999);              ! Missing value code
           IDVARIABLE = case;                 ! Person ID variable
!PLOT:     TYPE = PLOT1 PLOT2 PLOT3;          ! Get all the plots
OUTPUT:    STDYX RESIDUAL MODINDICES(3.841);  ! Constraints to drop p<.05

ANALYSIS:  ESTIMATOR=WLSMV; PARAMETERIZATION=THETA; ! Limited-info version
           CONVERGENCE = 0.0000001;        ! For OS comparability
SAVEDATA:  DIFFTEST=Configural.dat;   ! Save configural info to compare

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

! WOMEN ALTERNATIVE GROUP: 1 CONFIGURAL MODEL
MODEL WOMEN:
! Factor loadings all freely estimated
  IADL BY cia1-cia7*;
! Item thresholds all freely estimated
  [cia1$1-cia7$1*];
  [cia1$2-cia7$2*];
  [cia1$3-cia7$3*];
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       56
Chi-Square Test of Model Fit
          Value                             72.897*
          Degrees of Freedom                   28
          P-Value                           0.0000
Chi-Square Contributions From Each Group
          MEN                               24.965
          WOMEN                             47.932
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.071
          90 Percent C.I.                    0.051  0.091
          Probability RMSEA <= .05           0.040
CFI/TLI
          CFI                                0.999
          TLI                                0.998
SRMR (Standardized Root Mean Square Residual)
          Value                              0.020
```

This will be our baseline configural model.
56 parameters estimated = 2*[7 loadings + 21 thresholds] = 56
Possible parameters = 2* ([7*(7+1)] / 2] + 21 thresholds) = 98
DF = 98 – 56 – 14 "residual variances" = 28



Model 1. Configural Invariance Model
(Everything separate across groups *except* for parameters needed to be constrained for identification)

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

Group MEN

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS: CHANGE IN PROBIT FOR 1-UNIT CHANGE IN THETA				
IADL BY				
CIA1	6.012	1.576	3.815	0.000
CIA2	3.178	0.545	5.828	0.000
CIA3	3.079	0.441	6.986	0.000
CIA4	4.146	0.835	4.966	0.000
CIA5	2.575	0.373	6.901	0.000
CIA6	2.657	0.492	5.401	0.000
CIA7	1.115	0.218	5.119	0.000
Means: MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	0.000	0.000	999.000	999.000
Thresholds: EXPECTED PROBIT OF Y=0 IF THETA=0				
CIA1$1	-7.621	1.979	-3.850	0.000
CIA1$2	-5.251	1.458	-3.601	0.000
CIA1$3	-0.851	0.683	-1.247	0.212
CIA2$1	-4.476	0.635	-7.048	0.000
CIA2$2	-3.842	0.584	-6.576	0.000
CIA2$3	-2.260	0.486	-4.646	0.000
CIA3$1	-3.870	0.480	-8.063	0.000
CIA3$2	-2.802	0.448	-6.249	0.000
CIA3$3	-0.563	0.324	-1.738	0.082
CIA4$1	-5.193	0.930	-5.583	0.000
CIA4$2	-4.199	0.761	-5.517	0.000
CIA4$3	-2.181	0.570	-3.829	0.000
CIA5$1	-4.608	0.572	-8.058	0.000
CIA5$2	-2.620	0.374	-7.013	0.000
CIA5$3	-1.190	0.299	-3.975	0.000
CIA6$1	-4.592	0.705	-6.518	0.000
CIA6$2	-3.333	0.567	-5.884	0.000
CIA6$3	-2.228	0.471	-4.725	0.000
CIA7$1	-3.115	0.383	-8.123	0.000
CIA7$2	-2.376	0.279	-8.516	0.000
CIA7$3	-1.780	0.234	-7.591	0.000
Variances: VARIANCE OF THETA FIXED=1 FOR IDENTIFICATION				
IADL	1.000	0.000	999.000	999.000
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

Group WOMEN

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS: CHANGE IN PROBIT FOR 1-UNIT CHANGE IN THETA				
IADL BY				
CIA1	3.087	0.306	10.104	0.000
CIA2	3.425	0.467	7.330	0.000
CIA3	4.813	0.882	5.459	0.000
CIA4	3.287	0.331	9.927	0.000
CIA5	2.202	0.196	11.258	0.000
CIA6	1.741	0.172	10.100	0.000
CIA7	1.056	0.156	6.786	0.000
Means: MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	0.000	0.000	999.000	999.000
Thresholds: EXPECTED PROBIT OF Y=0 IF THETA=0				
CIA1$1	-4.586	0.401	-11.432	0.000
CIA1$2	-3.292	0.337	-9.783	0.000
CIA1$3	-0.707	0.222	-3.188	0.001
CIA2$1	-5.439	0.628	-8.666	0.000
CIA2$2	-4.465	0.550	-8.119	0.000
CIA2$3	-2.808	0.445	-6.314	0.000
CIA3$1	-7.096	1.136	-6.244	0.000
CIA3$2	-6.051	1.023	-5.912	0.000
CIA3$3	-3.823	0.802	-4.769	0.000
CIA4$1	-4.483	0.378	-11.846	0.000
CIA4$2	-2.806	0.306	-9.156	0.000
CIA4$3	-0.679	0.238	-2.854	0.004
CIA5$1	-3.746	0.281	-13.345	0.000
CIA5$2	-1.687	0.190	-8.881	0.000
CIA5$3	-0.118	0.151	-0.779	0.436
CIA6$1	-3.202	0.243	-13.150	0.000
CIA6$2	-2.115	0.190	-11.134	0.000
CIA6$3	-1.173	0.162	-7.254	0.000
CIA7$1	-3.407	0.325	-10.497	0.000
CIA7$2	-2.712	0.233	-11.636	0.000
CIA7$3	-1.747	0.165	-10.579	0.000
Variances: VARIANCE OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	1.000	0.000	999.000	999.000
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

Model 2a. Metric Invariance Model (IFA loadings held equal across groups – Mplus IRT discriminations still vary via factor variances)

```
ANALYSIS:  DIFFTEST=Configural.dat; ! Compare with 1 configural

SAVEDATA:  DIFFTEST=MetricA.dat;    ! Save 2a full metric info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

! WOMEN ALTERNATIVE GROUP: 2A METRIC MODEL
MODEL WOMEN:
! Factor loadings NOW CONSTRAINED EQUAL TO MEN VIA SAME LABEL
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated
  [cia1$1-cia7$1*];
  [cia1$2-cia7$2*];
  [cia1$3-cia7$3*];
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 for identification
! Factor variance NOW ESTIMATED
  [IADL@0]; IADL*;
```

```
Number of Free Parameters                       50

Chi-Square Test of Model Fit
          Value                             64.667*
          Degrees of Freedom                    34
          P-Value                           0.0012

Chi-Square Contributions From Each Group
          MEN                               30.854
          WOMEN                             33.813

THIS IS THE TEST OF 1 CONFIGURAL VS. 2A FULL METRIC INVARIANCE
Chi-Square Test for Difference Testing
          Value                              9.401
          Degrees of Freedom                     6
          P-Value                           0.1522

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.053
          90 Percent C.I.                    0.033  0.073
          Probability RMSEA <= .05           0.367

CFI/TLI
          CFI                                0.999
          TLI                                0.999

SRMR (Standardized Root Mean Square Residual)
          Value                              0.023
```

Although the DIFFTEST omnibus chi-square is nonsignificant, the modification indices suggest freeing the loading for cia3 between groups:

```
MODEL MODIFICATION INDICES
                                  M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
Group MEN

BY Statements
IADL     BY CIA3                 7.734    -2.379     -2.379       -0.488

Group WOMEN

BY Statements
IADL     BY CIA3                 7.734     1.531      1.275        0.311
```

Model 2b. Metric Invariance Model (IFA loadings held equal across groups except cia3)

```
ANALYSIS:  DIFFTEST=Configural.dat;  ! Compare with 1 configural

SAVEDATA:  DIFFTEST=MetricB.dat;     ! Save 2b partial metric info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 2B PARTIAL METRIC MODEL
MODEL WOMEN:
! Factor loadings NOW CONSTRAINED EQUAL TO MEN EXCEPT 3
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds all freely estimated
  [cia1$1-cia7$1*];
  [cia1$2-cia7$2*];
  [cia1$3-cia7$3*];
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 for identification
! Factor variance NOW ESTIMATED
  [IADL@0]; IADL*;
```

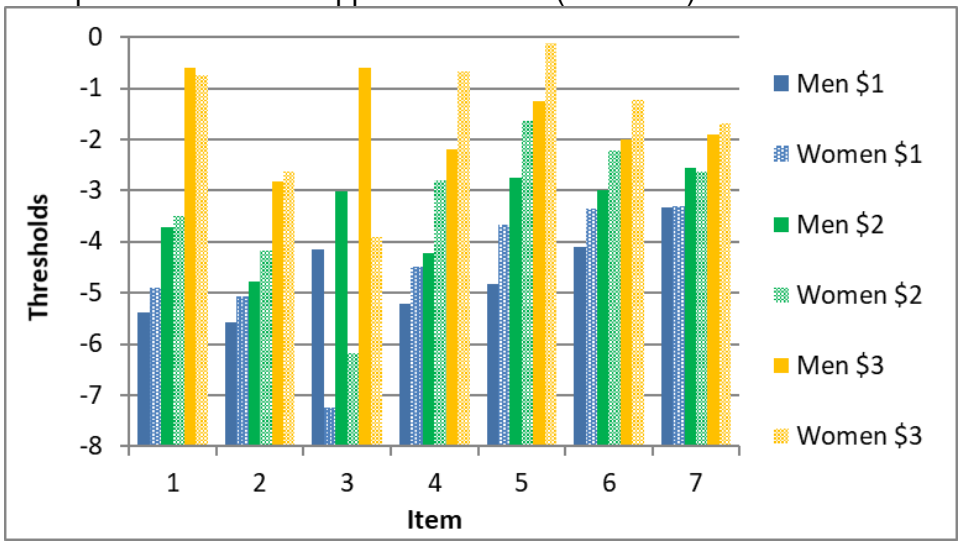
Here is an additional test comparing Metric B to Metric A
to see if freeing the loading of item 3 helped improve fit:
(by running Metric A again, adding these commands):

```
ANALYSIS:   DIFFTEST=MetricB.dat;  ! Compare with 2b partial metric

SAVEDATA:   DIFFTEST=MetricA.dat;  ! Save 2a full metric info
```

```
Chi-Square Test for Difference Testing
          Value                              5.560*
          Degrees of Freedom                    1
          P-Value                           0.0184
```

```
Number of Free Parameters                       51
Chi-Square Test of Model Fit
          Value                             58.607*
          Degrees of Freedom                    33
          P-Value                           0.0039
Chi-Square Contribution From Each Group
          MEN                               24.931
          WOMEN                             33.676

THIS IS THE TEST OF 1 CONFIGURAL VS. 2B PARTIAL METRIC INVARIANCE
Chi-Square Test for Difference Testing
          Value                              5.527
          Degrees of Freedom                     5
          P-Value                           0.3550

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.049
          90 Percent C.I.                    0.028  0.070
          Probability RMSEA <= .05           0.490
CFI/TLI
          CFI                                0.999
          TLI                                0.999
SRMR (Standardized Root Mean Square Residual)
          Value                              0.023
```

No new modification indices, so we are done with metric!
Let's preview what will happen with scalar (threshold) invariance:



2b Partial metric invariance solution (all factor loadings constrained except item 3)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Group MEN				
IADL BY: CHANGE IN PROBIT FOR ONE-UNIT CHANGE IN THETA				
CIA1	4.195	0.456	9.207	0.000
CIA2	4.026	0.556	7.239	0.000
CIA3	**3.327**	**0.518**	**6.428**	**0.000**
CIA4	4.168	0.506	8.240	0.000
CIA5	2.713	0.301	9.012	0.000
CIA6	2.339	0.261	8.946	0.000
CIA7	1.256	0.174	7.227	0.000
Means MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	0.000	0.000	999.000	999.000
Thresholds: EXPECTED PROBIT OF Y=0 IF THETA=0				
CIA1$1	-5.393	0.615	-8.770	0.000
CIA1$2	-3.716	0.552	-6.730	0.000
CIA1$3	-0.602	0.422	-1.429	0.153
CIA2$1	-5.574	0.673	-8.285	0.000
CIA2$2	-4.785	0.644	-7.432	0.000
CIA2$3	-2.814	0.530	-5.313	0.000
CIA3$1	-4.152	0.547	-7.597	0.000
CIA3$2	-3.007	0.504	-5.972	0.000
CIA3$3	-0.604	0.352	-1.718	0.086
CIA4$1	-5.219	0.605	-8.625	0.000
CIA4$2	-4.220	0.555	-7.609	0.000
CIA4$3	-2.192	0.474	-4.625	0.000
CIA5$1	-4.824	0.489	-9.856	0.000
CIA5$2	-2.743	0.352	-7.801	0.000
CIA5$3	-1.245	0.291	-4.285	0.000
CIA6$1	-4.115	0.415	-9.904	0.000
CIA6$2	-2.987	0.339	-8.809	0.000
CIA6$3	-1.996	0.303	-6.594	0.000
CIA7$1	-3.339	0.355	-9.394	0.000
CIA7$2	-2.547	0.253	-10.085	0.000
CIA7$3	-1.908	0.219	-8.716	0.000
Variances: VARIANCE OF THETA FIXED=1 FOR IDENTIFICATION				
IADL	1.000	0.000	999.000	999.000
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Group WOMEN				
IADL BY				
CIA1	4.195	0.456	9.207	0.000
CIA2	4.026	0.556	7.239	0.000
CIA3	**6.227**	**1.253**	**4.969**	**0.000**
CIA4	4.168	0.506	8.240	0.000
CIA5	2.713	0.301	9.012	0.000
CIA6	2.339	0.261	8.946	0.000
CIA7	1.256	0.174	7.227	0.000
Means: MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	0.000	0.000	999.000	999.000
Thresholds				
CIA1$1	-4.889	0.429	-11.389	0.000
CIA1$2	-3.510	0.361	-9.735	0.000
CIA1$3	-0.754	0.236	-3.187	0.001
CIA2$1	-5.080	0.486	-10.462	0.000
CIA2$2	-4.170	0.430	-9.691	0.000
CIA2$3	-2.622	0.360	-7.275	0.000
CIA3$1	-7.242	1.230	-5.888	0.000
CIA3$2	-6.176	1.103	-5.598	0.000
CIA3$3	-3.902	0.854	-4.569	0.000
CIA4$1	-4.488	0.367	-12.214	0.000
CIA4$2	-2.809	0.301	-9.331	0.000
CIA4$3	-0.679	0.236	-2.880	0.004
CIA5$1	-3.662	0.259	-14.126	0.000
CIA5$2	-1.649	0.179	-9.200	0.000
CIA5$3	-0.115	0.148	-0.780	0.435
CIA6$1	-3.349	0.256	-13.090	0.000
CIA6$2	-2.212	0.197	-11.205	0.000
CIA6$3	-1.227	0.168	-7.321	0.000
CIA7$1	-3.300	0.287	-11.506	0.000
CIA7$2	-2.626	0.200	-13.111	0.000
CIA7$3	-1.692	0.142	-11.901	0.000
Variances: VARIANCE OF THETA NOW ESTIMATED				
IADL	**0.623**	**0.139**	**4.471**	**0.000**
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

Model 3a. Full Threshold Invariance Model (all IFA thresholds held equal across groups – Mplus IRT difficulties will still vary)

```
ANALYSIS: DIFFTEST=MetricB.dat; ! Compare with 2b partial metric

SAVEDATA: DIFFTEST=ScalarA.dat; ! Save 3a full scalar info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3A FULL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds NOW CONSTRAINED EQUAL TO MEN BY OMISSION



! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean NOW ESTIMATED
! Factor variance still estimated
  [IADL*]; IADL*;
```

The DIFFTEST chi-square is significant, and the modification indices suggest that item 3 threshold 3 is the biggest problem. This is not surprising given that its loading differed, too (but I thought I'd try to constrain item 3's thresholds anyway given that there are only seven items).

Btw, the imbalance in sample size (more women than men) weighted the common item 3 threshold 3 towards that of women (not shown).

```
MODEL FIT INFORMATION
Number of Free Parameters                       31

Chi-Square Test of Model Fit
          Value                             134.477*
          Degrees of Freedom                    53
          P-Value                           0.0000

Chi-Square Contributions From Each Group
          MEN                               88.178
          WOMEN                             46.299

THIS IS THE TEST OF 2B PARTIAL METRIC VS. 3A FULL SCALAR INVARIANCE
Chi-Square Test for Difference Testing
          Value                             91.663
          Degrees of Freedom                    20
          P-Value                           0.0000

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.070
          90 Percent C.I.                    0.055  0.084
          Probability RMSEA <= .05           0.014
CFI/TLI
          CFI                                0.998
          TLI                                0.998
SRMR (Standardized Root Mean Square Residual)
          Value                              0.032

MODEL MODIFICATION INDICES
                                    M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
Group MEN
Means/Intercepts/Thresholds
[ IADL     ]                       40.777    -4.610     -4.610       -4.610
[ CIA1$3   ]                        6.825     0.628      0.628        0.159
[ CIA3$1   ]                        6.313    -2.566     -2.566       -0.351
[ CIA3$3   ]                       26.583     3.043      3.043        0.417
[ CIA4$3   ]                        4.572    -0.544     -0.544       -0.130
[ CIA5$2   ]                        4.494    -0.409     -0.409       -0.146
[ CIA5$3   ]                        8.185    -0.460     -0.460       -0.164
```

Note: As of Mplus v 7, there is an "automatic" option to do invariance testing. For WLSMV, at this step (scalar), their version simultaneously constrains the thresholds while freeing the residual variances in the alternative group (which are now possible to be separately identified). I disagree with this logic, and thus these results will not match those of the automated testing. Instead, I test scalar first, and then test (added estimated) residual variances in a separate step. Reasonable people may disagree with this choice, but to me the residual variances are of much less importance given that they are not even estimated parameters ever in the full-information estimation version of these IFA/IRT models.

Model 3b. Partial Threshold Invariance Model (freeing all item 3 thresholds between groups)

```
ANALYSIS: DIFFTEST=MetricB.dat; ! Compare with 2b partial metric

SAVEDATA: DIFFTEST=ScalarB.dat; ! Save 3b partial scalar info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3B PARTIAL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds otherwise CONSTRAINED EQUAL TO MEN BY OMISSION
  [cia3$1* cia3$2* cia3$3*];  ! 3b: Free item 3


! Item residual variances all fixed=1
  cia1-cia7@1;
! Factor mean NOW ESTIMATED
! Factor variance still estimated
  [IADL*]; IADL*;
```

```
Threshold for Men…
    CIA3$1            -4.112      0.526     -7.814      0.000
    CIA3$2            -2.978      0.487     -6.114      0.000
    CIA3$3            -0.599      0.347     -1.727      0.084
Threshold for Women…
    CIA3$1            -8.823      1.576     -5.598      0.000
    CIA3$2            -7.745      1.448     -5.349      0.000
    CIA3$3            -5.446      1.208     -4.508      0.000
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       34

Chi-Square Test of Model Fit
          Value                             96.673*
          Degrees of Freedom                   50
          P-Value                           0.0001

Chi-Square Contributions From Each Group
          MEN                               53.401
          WOMEN                             43.271

THIS IS THE TEST OF 2B PARTIAL METRIC VS. 3B PARTIAL SCALAR INVARIANCE
Chi-Square Test for Difference Testing
          Value                             44.798
          Degrees of Freedom                   17
          P-Value                           0.0003

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.054
          90 Percent C.I.                    0.038  0.070
          Probability RMSEA <= .05           0.315

CFI/TLI
          CFI                                0.999
          TLI                                0.999

SRMR (Standardized Root Mean Square Residual)
          Value                              0.026

MODEL MODIFICATION INDICES
                                    M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
Group MEN
Means/Intercepts/Thresholds
[ CIA1$3    ]                     10.582      0.855       0.855        0.196
[ CIA5$3    ]                      4.999     -0.373      -0.373       -0.127
```

The DIFFTEST chi-square is still significant, and the modification indices suggest that item 1 threshold 3 is the next to go…

Model 3c. Partial Threshold Invariance Model (also freeing item 1 threshold 3 between groups)

```
ANALYSIS: DIFFTEST=MetricB.dat; ! Compare with 2b partial metric

SAVEDATA: DIFFTEST=ScalarC.dat; ! Save 3c partial scalar info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3C PARTIAL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds otherwise CONSTRAINED EQUAL TO MEN BY OMISSION
  [cia3$1* cia3$2* cia3$3*];  ! 3b: Free item 3
  [cia1$3*]; ! 3c: Free item 1 threshold 3

! Item residual variances all fixed=1
  cia1-cia7@1;
! Factor mean NOW ESTIMATED
! Factor variance still estimated
  [IADL*]; IADL*;
```

Thresholds for Men…

CIA1$1	-5.906	0.583	-10.127	0.000
CIA1$2	-4.468	0.530	-8.426	0.000
CIA1$3	**-0.620**	**0.434**	**-1.429**	**0.153**

Thresholds for Women…

CIA1$1	-5.906	0.583	-10.127	0.000
CIA1$2	-4.468	0.530	-8.426	0.000
CIA1$3	**-1.936**	**0.462**	**-4.193**	**0.000**

MODEL FIT INFORMATION

Number of Free Parameters 35

Chi-Square Test of Model Fit

Value	87.761*
Degrees of Freedom	49
P-Value	0.0006

Chi-Square Contributions From Each Group

MEN	46.385
WOMEN	41.376

THIS IS THE TEST OF 2B PARTIAL METRIC VS. 3C PARTIAL SCALAR INVARIANCE

Chi-Square Test for Difference Testing

Value	32.462
Degrees of Freedom	16
P-Value	0.0087

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.050	
90 Percent C.I.	0.033	0.067
Probability RMSEA <= .05	0.480	

CFI/TLI

CFI	0.999
TLI	0.999

SRMR (Standardized Root Mean Square Residual)

Value	0.025

MODEL MODIFICATION INDICES – **after re-running with 3.841 (p<.05) cutoff:**

	M.I.	E.P.C.	Std E.P.C.	StdYX E.P.C.
Group MEN				
Means/Intercepts/Thresholds				
[CIA2$3]	**5.146**	**0.635**	**0.635**	**0.147**

The DIFFTEST chi-square is still significant, and the modification indices suggest that item 2 threshold 3 is the next to go…

Model 3d. Partial Threshold Invariance Model (also freeing item 2 threshold 3 between groups)

```
ANALYSIS: DIFFTEST=MetricB.dat; ! Compare with 2b partial metric

SAVEDATA: DIFFTEST=ScalarD.dat; ! Save 3d partial scalar info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3D PARTIAL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds otherwise CONSTRAINED EQUAL TO MEN BY OMISSION
  [cia3$1* cia3$2* cia3$3*];  ! 3b: Free item 3
  [cia1$3*]; ! 3c: Free item 1 threshold 3
  [cia2$3*]; ! 3d: Free item 2 threshold 3

! Item residual variances all fixed=1
  cia1-cia7@1;
! Factor mean NOW ESTIMATED
! Factor variance still estimated
  [IADL*]; IADL*;
```

```
Thresholds for Men…
    CIA2$1            -6.162      0.683     -9.026      0.000
    CIA2$2            -5.290      0.640     -8.269      0.000
    CIA2$3            -2.927      0.554     -5.279      0.000

Thresholds for Women…
    CIA2$1            -6.162      0.683     -9.026      0.000
    CIA2$2            -5.290      0.640     -8.269      0.000
    CIA2$3            -3.852      0.587     -6.566      0.000
```

```
MODEL FIT INFORMATION

Number of Free Parameters                       36

Chi-Square Test of Model Fit
          Value                             82.974*
          Degrees of Freedom                   48
          P-Value                           0.0013

Chi-Square Contributions From Each Group
          MEN                               42.552
          WOMEN                             40.422

THIS IS THE TEST OF 2B PARTIAL METRIC VS. 3D PARTIAL SCALAR INVARIANCE
Chi-Square Test for Difference Testing
          Value                             26.168
          Degrees of Freedom                   15
          P-Value                           0.0363

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.048
          90 Percent C.I.                    0.030  0.065
          Probability RMSEA <= .05           0.556

CFI/TLI
          CFI                                0.999
          TLI                                0.999

SRMR (Standardized Root Mean Square Residual)
          Value                              0.025

MODEL MODIFICATION INDICES
                                    M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
Group MEN
Means/Intercepts/Thresholds
[ CIA1$2    ]                     4.629      0.721       0.721        0.162
```

So close… but the DIFFTEST chi-square is still significant, and the modification indices suggest that cia1 threshold 2 is the next to go…

Model 3e. Partial Threshold Invariance Model (also freeing item 1 threshold 2 between groups)

```
ANALYSIS: DIFFTEST=MetricB.dat; ! Compare with 2b partial metric

SAVEDATA: DIFFTEST=ScalarE.dat; ! Save 3e partial scalar info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3D PARTIAL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds otherwise CONSTRAINED EQUAL TO MEN BY OMISSION
  [cia3$1* cia3$2* cia3$3*];  ! 3b: Free item 3
  [cia1$3*]; ! 3c: Free item 1 threshold 3
  [cia2$3*]; ! 3d: Free item 2 threshold 3
  [cia1$2*]; ! 3e: Free item 1 threshold 2

! Item residual variances all fixed=1
  cia1-cia7@1;
! Factor mean NOW ESTIMATED
! Factor variance still estimated
  [IADL*]; IADL*;
```

Thresholds for Men…

CIA1$1	-6.028	0.589	-10.228	0.000
CIA1$2	**-3.826**	**0.569**	**-6.729**	**0.000**
CIA1$3	-0.620	0.434	-1.429	0.153

Thresholds for Women…

CIA1$1	-6.028	0.589	-10.228	0.000
CIA1$2	**-4.817**	**0.543**	**-8.877**	**0.000**
CIA1$3	-2.102	0.467	-4.496	0.000

MODEL FIT INFORMATION

Number of Free Parameters 37

Chi-Square Test of Model Fit

Value	78.436*
Degrees of Freedom	47
P-Value	0.0027

Chi-Square Contributions From Each Group

MEN	38.881
WOMEN	39.555

THIS IS THE TEST OF 2B PARTIAL METRIC VS. 3E PARTIAL SCALAR INVARIANCE

Chi-Square Test for Difference Testing

Value	20.322*
Degrees of Freedom	14
P-Value	0.1203

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.046	
90 Percent C.I.	0.027	0.063
Probability RMSEA <= .05	0.627	

CFI/TLI

CFI	0.999
TLI	0.999

SRMR (Standardized Root Mean Square Residual)

Value	0.025

No more significant modification indices for specific thresholds!
Scalar invariance is finally done!

The next step is to see if the “residual” variances need to differ between groups. This is only possible (i.e., the model with freely estimated residual variances is only identified) given at least partial metric and scalar invariance per item.

So far the residual variances have been fixed=1 in both groups for identification. We now proceed backwards: We first need to estimate a “bigger” model (“ResidualFreeA”), in which the residual variances are freed in the alternative women group for the items with at least partial scalar invariance (so all but item 3 here), and compare it back to our Scalar E model (in which they were fixed=1), to be renamed as “ResidualFixedB” for consistency with its purpose.

Partial Scalar E invariance solution

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Group MEN				
IADL BY: CHANGE IN PROBIT FOR ONE-UNIT CHANGE IN THETA				
CIA1	4.326	0.475	9.102	0.000
CIA2	4.192	0.585	7.165	0.000
CIA3	**3.242**	**0.476**	**6.814**	**0.000**
CIA4	4.331	0.524	8.274	0.000
CIA5	2.820	0.316	8.933	0.000
CIA6	2.403	0.271	8.855	0.000
CIA7	1.307	0.182	7.184	0.000
Means: MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	**0.000**	**0.000**	**999.000**	**999.000**
Thresholds: EXPECTED PROBIT OF Y=0 IF THETA=0				
CIA1$1	-6.028	0.589	-10.228	0.000
CIA1$2	**-3.826**	**0.569**	**-6.729**	**0.000**
CIA1$3	**-0.620**	**0.434**	**-1.429**	**0.153**
CIA2$1	-6.214	0.686	-9.057	0.000
CIA2$2	-5.343	0.643	-8.307	0.000
CIA2$3	**-2.924**	**0.554**	**-5.278**	**0.000**
CIA3$1	**-4.055**	**0.510**	**-7.957**	**0.000**
CIA3$2	**-2.937**	**0.473**	**-6.202**	**0.000**
CIA3$3	**-0.590**	**0.341**	**-1.729**	**0.084**
CIA4$1	-5.716	0.563	-10.151	0.000
CIA4$2	-4.195	0.505	-8.299	0.000
CIA4$3	-2.085	0.459	-4.546	0.000
CIA5$1	-4.622	0.368	-12.555	0.000
CIA5$2	-2.590	0.314	-8.239	0.000
CIA5$3	-1.066	0.285	-3.744	0.000
CIA6$1	-4.090	0.337	-12.132	0.000
CIA6$2	-2.961	0.298	-9.929	0.000
CIA6$3	-1.981	0.277	-7.151	0.000
CIA7$1	-3.593	0.295	-12.171	0.000
CIA7$2	-2.879	0.234	-12.324	0.000
CIA7$3	-2.058	0.194	-10.602	0.000
Variances: VARIANCE OF THETA FIXED=1 FOR IDENTIFICATION				
IADL	1.000	0.000	999.000	999.000
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Group WOMEN				
IADL BY: CHANGE IN PROBIT FOR ONE-UNIT CHANGE IN THETA				
CIA1	4.326	0.475	9.102	0.000
CIA2	4.192	0.585	7.165	0.000
CIA3	**6.722**	**1.377**	**4.882**	**0.000**
CIA4	4.331	0.524	8.274	0.000
CIA5	2.820	0.316	8.933	0.000
CIA6	2.403	0.271	8.855	0.000
CIA7	1.307	0.182	7.184	0.000
Means: MEAN OF THETA ESTIMATED				
IADL	**-0.314**	**0.097**	**-3.239**	**0.001**
Thresholds				
CIA1$1	-6.028	0.589	-10.228	0.000
CIA1$2	**-4.817**	**0.543**	**-8.877**	**0.000**
CIA1$3	**-2.102**	**0.467**	**-4.496**	**0.000**
CIA2$1	-6.214	0.686	-9.057	0.000
CIA2$2	-5.343	0.643	-8.307	0.000
CIA2$3	**-3.923**	**0.590**	**-6.645**	**0.000**
CIA3$1	**-9.561**	**1.790**	**-5.341**	**0.000**
CIA3$2	**-8.463**	**1.653**	**-5.121**	**0.000**
CIA3$3	**-6.125**	**1.392**	**-4.401**	**0.000**
CIA4$1	-5.716	0.563	-10.151	0.000
CIA4$2	-4.195	0.505	-8.299	0.000
CIA4$3	-2.085	0.459	-4.546	0.000
CIA5$1	-4.622	0.368	-12.555	0.000
CIA5$2	-2.590	0.314	-8.239	0.000
CIA5$3	-1.066	0.285	-3.744	0.000
CIA6$1	-4.090	0.337	-12.132	0.000
CIA6$2	-2.961	0.298	-9.929	0.000
CIA6$3	-1.981	0.277	-7.151	0.000
CIA7$1	-3.593	0.295	-12.171	0.000
CIA7$2	-2.879	0.234	-12.324	0.000
CIA7$3	-2.058	0.194	-10.602	0.000
Variances: VARIANCE OF THETA ESTIMATED				
IADL	**0.567**	**0.125**	**4.531**	**0.000**
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

Model 4a. Partial Residual Variance Invariance Model (thresholds unconstrained between groups from Scalar E)

This last step for testing measurement invariance proceeds backwards. Because freeing the residual variances is adding parameters, we must estimate model 4a with free residual variances FIRST without comparing its fit to any other nested model. However, we cannot estimate the residual variance for item 3 because its factor loading differs between groups (i.e., separate residual variances are not identified for item 3).

```
ANALYSIS: ! Nothing compare to model 4a yet

SAVEDATA: DIFFTEST= ResidualFreeA.dat; ! Save 4a free residual info

MODEL: ! MEN REFERENCE GROUP MODEL (WILL STAY SAME)
! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;


! WOMEN ALTERNATIVE GROUP: 3D PARTIAL SCALAR MODEL
MODEL WOMEN:
! Factor loadings still constrained equal to men (except 3)
  IADL BY cia1-cia7* (L_I1-L_I2 L_I3w L_I4-L_I7); ! 2b: Free item 3
! Item thresholds otherwise constrained equal to men by omission
  [cia3$1* cia3$2* cia3$3*];  ! 3b: Free item 3
  [cia1$3*]; ! 3c: Free item 1 threshold 3
  [cia2$3*]; ! 3d: Free item 2 threshold 3
  [cia1$2*]; ! 3e: Free item 1 threshold 2
! 4A Item residual variances NOW FREE EXCEPT ITEM 3
  cia1* cia2* cia3@1 cia4* cia5* cia6* cia7*;
! Factor mean still estimated
! Factor variance still estimated
  [IADL*]; IADL*;
```

```
(Number of Free Parameters                       43

Chi-Square Test of Model Fit
          Value                             72.066*
          Degrees of Freedom                   41
          P-Value                           0.0019

Chi-Square Contributions From Each Group
          MEN                               29.078
          WOMEN                             42.988

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.049
          90 Percent C.I.                    0.029  0.067
          Probability RMSEA <= .05           0.515

CFI/TLI
          CFI                                0.999
          TLI                                0.999

SRMR (Standardized Root Mean Square Residual)
          Value                              0.021

Residual Variances for Men…
    CIA1               1.000      0.000    999.000    999.000
    CIA2               1.000      0.000    999.000    999.000
    CIA3               1.000      0.000    999.000    999.000
    CIA4               1.000      0.000    999.000    999.000
    CIA5               1.000      0.000    999.000    999.000
    CIA6               1.000      0.000    999.000    999.000
    CIA7               1.000      0.000    999.000    999.000

Residual Variances for Women…
    CIA1               1.706      0.962      1.772      0.076
    CIA2               0.388      0.143      2.720      0.007
    CIA3               1.000      0.000    999.000    999.000
    CIA4               0.720      0.308      2.333      0.020
    CIA5               0.700      0.245      2.855      0.004
    CIA6               1.115      0.391      2.855      0.004
    CIA7               0.552      0.146      3.784      0.000
```

Model 4b. Residual Variance Invariance Model (testing all items' residual variances except item 3)
We now constrain the residual variances to be equal between groups and test the decrease in fit (left side is re-labeled Scalar E).

```
ANALYSIS: DIFFTEST=ResidualFreeA.dat; ! Compare with 4a free residual

SAVEDATA: DIFFTEST=ResidualFixedB.dat; ! Save 4b fixed residual info

! 4B RESIDUAL FIXED MODEL IS THE SAME AS 3E SCALAR
! (needs to re-run for DIFFTEST comparison against 4A RESIDUAL FREE)

MODEL FIT INFORMATION
Number of Free Parameters                       37

Chi-Square Test of Model Fit
          Value                             78.436*
          Degrees of Freedom                    47
          P-Value                           0.0027

Chi-Square Contribution From Each Group
          MEN                               38.881
          WOMEN                             39.555

THIS IS THE TEST OF 4B FULL RESIDUAL VS. 3E PARTIAL SCALAR INVARIANCE
Chi-Square Test for Difference Testing
          Value                             11.016
          Degrees of Freedom                     6
          P-Value                           0.0879

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.046
          90 Percent C.I.                    0.027  0.063
          Probability RMSEA <= .05           0.627

CFI/TLI
          CFI                                0.999
          TLI                                0.999

SRMR (Standardized Root Mean Square Residual)
          Value                              0.025


MODEL MODIFICATION INDICES
                        M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
Group MEN
Variances/Residual Variances
CIA2                   4.998     1.004      1.004        0.054
CIA7                   9.395     0.632      0.632        0.233
```

Although DIFFTEST is nonsignificant, it looks like the residual variance for items 7 and 2 want to be free, too… I will free both at the same time to skip a step here (results testing them separately were the same):

```
ANALYSIS: DIFFTEST=ResidualFreeA.dat;  ! Compare with 4a free residual

SAVEDATA: DIFFTEST=ResidualPartialC.dat; ! Save 4c partially fixed
                                         ! residual info

! 4C PARTIAL RESIDUAL MODEL has residual vars partially constrained
  cia1@1 cia2* cia3-cia6@1 cia7*;

MODEL FIT INFORMATION for 4c
Number of Free Parameters                       39
Chi-Square Test of Model Fit
          Value                             69.757*
          Degrees of Freedom                    45
          P-Value                           0.0104
Chi-Square Contribution From Each Group
          MEN                               30.125
          WOMEN                             39.632

THIS IS THE TEST OF 4C PARTIAL RESIDUAL VS. 3E PARTIAL SCALAR INVARIANCE
Chi-Square Test for Difference Testing
          Value                              2.800*
          Degrees of Freedom                     4
          P-Value                           0.5918
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.042
          90 Percent C.I.                    0.021  0.060
          Probability RMSEA <= .05           0.754
CFI/TLI
          CFI                                0.999
          TLI                                0.999
SRMR (Standardized Root Mean Square Residual)
          Value                              0.021
```

No more modification indices were significant, so residual variance is now done! The last step is to test structural invariance to see if the factor variance can be constrained across groups:

```
ANALYSIS: DIFFTEST=ResidualPartialC.dat; ! Compare 4c partial residual

SAVEDATA: DIFFTEST=FactorVariance.dat;  ! Save 5 fixed variance info

! MODEL IS THE SAME AS 4C PARTIAL RESIDUAL EXCEPT:
    IADL@1;

Chi-Square Test for Difference Testing: Keep the factor variance
different!
          Value                             11.498
          Degrees of Freedom                     1
          P-Value                           0.0007
```

The factor mean was significantly lower for women than mean already (from the last partial measurement invariance model 4C), so we are done!

Final Model: Partial Measurement Invariance (solution from Model 4c)

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

Group MEN

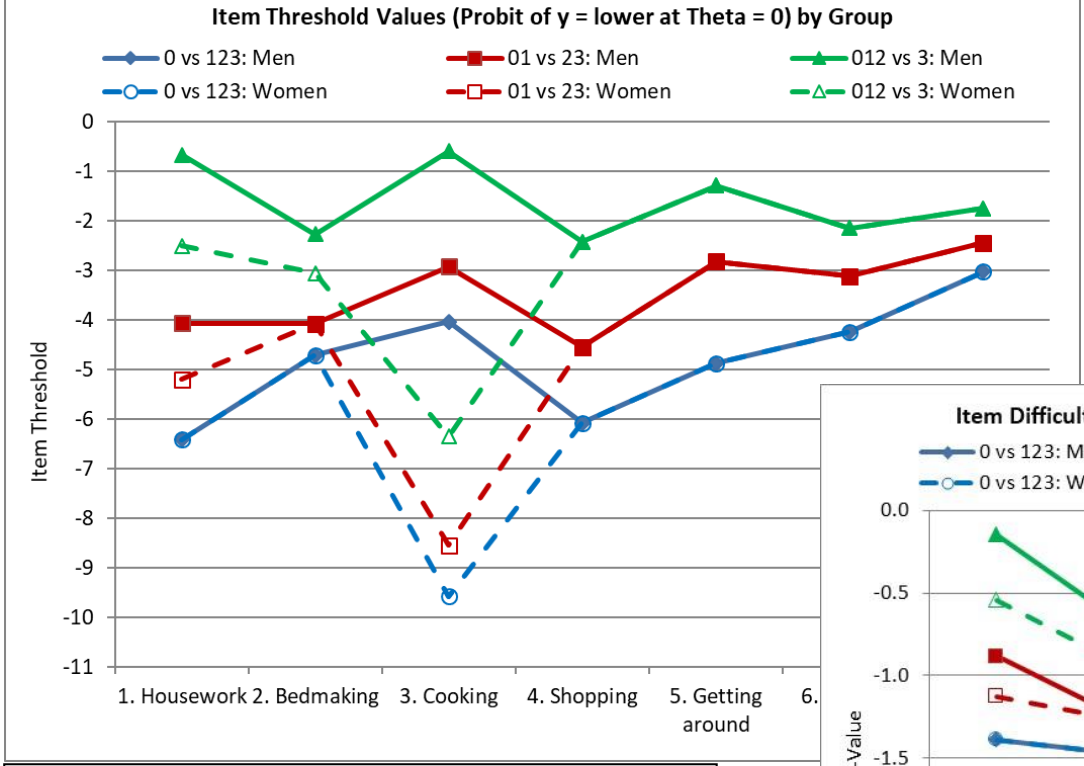
	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
FACTOR LOADINGS: CHANGE IN PROBIT FOR 1-UNIT CHANGE IN THETA				
IADL BY				
CIA1	4.614	0.532	8.674	0.000
CIA2	3.191	0.506	6.309	0.000
CIA3	**3.227**	**0.467**	**6.912**	**0.000**
CIA4	4.607	0.571	8.062	0.000
CIA5	3.009	0.352	8.548	0.000
CIA6	2.519	0.292	8.614	0.000
CIA7	1.123	0.174	6.471	0.000
Means: MEAN OF THETA FIXED=0 FOR IDENTIFICATION				
IADL	0.000	0.000	999.000	999.000
Thresholds: EXPECTED PROBIT OF LOWER CATEGORY WHEN THETA=0				
CIA1$1	-6.411	0.651	-9.855	0.000
CIA1$2	**-4.068**	**0.617**	**-6.594**	**0.000**
CIA1$3	**-0.660**	**0.461**	**-1.431**	**0.152**
CIA2$1	-4.709	0.665	-7.081	0.000
CIA2$2	-4.078	0.597	-6.836	0.000
CIA2$3	**-2.268**	**0.470**	**-4.823**	**0.000**
CIA3$1	**-4.037**	**0.501**	**-8.062**	**0.000**
CIA3$2	**-2.924**	**0.467**	**-6.263**	**0.000**
CIA3$3	**-0.588**	**0.339**	**-1.732**	**0.083**
CIA4$1	-6.079	0.613	-9.919	0.000
CIA4$2	-4.544	0.553	-8.217	0.000
CIA4$3	-2.423	0.504	-4.805	0.000
CIA5$1	-4.868	0.409	-11.895	0.000
CIA5$2	-2.822	0.348	-8.106	0.000
CIA5$3	-1.284	0.317	-4.055	0.000
CIA6$1	-4.240	0.356	-11.908	0.000
CIA6$2	-3.117	0.318	-9.794	0.000
CIA6$3	-2.143	0.298	-7.199	0.000
CIA7$1	-3.024	0.357	-8.472	0.000
CIA7$2	-2.438	0.287	-8.490	0.000
CIA7$3	-1.747	0.225	-7.782	0.000
Variances: VARIANCE OF THETA FIXED=1 FOR IDENTIFICATION				
IADL	1.000	0.000	999.000	999.000
Residual Variances ALL FIXED=1 FOR IDENTIFICATION				
CIA1	1.000	0.000	999.000	999.000
CIA2	1.000	0.000	999.000	999.000
CIA3	1.000	0.000	999.000	999.000
CIA4	1.000	0.000	999.000	999.000
CIA5	1.000	0.000	999.000	999.000
CIA6	1.000	0.000	999.000	999.000
CIA7	1.000	0.000	999.000	999.000

UNSTANDARDIZED MODEL RESULTS (IFA MODEL SOLUTION)

Group WOMEN

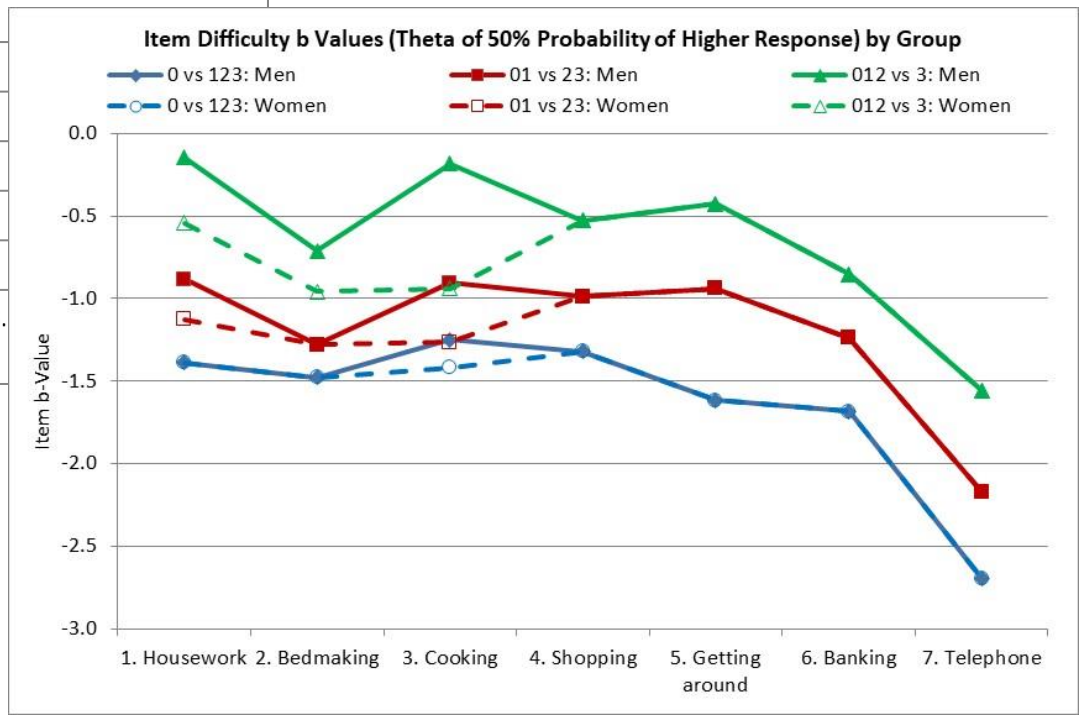
	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value	
FACTOR LOADINGS					
IADL BY					
CIA1	4.614	0.532	8.674	0.000	
CIA2	3.191	0.506	6.309	0.000	
CIA3	**6.758**	**1.277**	**5.294**	**0.000**	**cooking**
CIA4	4.607	0.571	8.062	0.000	
CIA5	3.009	0.352	8.548	0.000	
CIA6	2.519	0.292	8.614	0.000	
CIA7	1.123	0.174	6.471	0.000	
Means: MEAN DIFFERENCE OF THETA IN WOMEN					
IADL	**-0.382**	**0.095**	**-4.014**	**0.000**	
Thresholds:					
CIA1$1	-6.411	0.651	-9.855	0.000	housework
CIA1$2	**-5.198**	**0.602**	**-8.633**	**0.000**	
CIA1$3	**-2.499**	**0.520**	**-4.808**	**0.000**	
CIA2$1	-4.709	0.665	-7.081	0.000	bed making
CIA2$2	-4.078	0.597	-6.836	0.000	
CIA2$3	**-3.053**	**0.497**	**-6.143**	**0.000**	
CIA3$1	**-9.576**	**1.602**	**-5.976**	**0.000**	**cooking**
CIA3$2	**-8.545**	**1.498**	**-5.705**	**0.000**	
CIA3$3	**-6.349**	**1.307**	**-4.857**	**0.000**	
CIA4$1	-6.079	0.613	-9.919	0.000	shopping
CIA4$2	-4.544	0.553	-8.217	0.000	
CIA4$3	-2.423	0.504	-4.805	0.000	
CIA5$1	-4.868	0.409	-11.895	0.000	get around
CIA5$2	-2.822	0.348	-8.106	0.000	
CIA5$3	-1.284	0.317	-4.055	0.000	
CIA6$1	-4.240	0.356	-11.908	0.000	banking
CIA6$2	-3.117	0.318	-9.794	0.000	
CIA6$3	-2.143	0.298	-7.199	0.000	
CIA7$1	-3.024	0.357	-8.472	0.000	telephone
CIA7$2	-2.438	0.287	-8.490	0.000	
CIA7$3	-1.747	0.225	-7.782	0.000	
Variances: VARIANCE OF THETA					
IADL	0.492	0.111	4.424	0.000	
Residual Variances (NOT ALL FIXED=1)					
CIA1	1.000	0.000	999.000	999.000	
CIA2	**0.425**	**0.154**	**2.769**	**0.006**	
CIA3	1.000	0.000	999.000	999.000	
CIA4	1.000	0.000	999.000	999.000	
CIA5	1.000	0.000	999.000	999.000	
CIA6	1.000	0.000	999.000	999.000	
CIA7	**0.564**	**0.158**	**3.558**	**0.000**	

Figures from Partial Measurement Invariance Model 4c (see next page for Mplus code and output for the IRT a_i and b_i parameters <u>using theta mean=0 and variance=1 in both groups to maintain comparable invariance</u>):



The **thresholds** (shown on the left) predict the probit of y=lower response for someone with Theta=0. At theta=0, women (dashed lines) have a *lower* probability than men of not being able to do housework, make the bed, or cook (items 1–3).

The **b-values** (=threshold/loading; shown below) predict the trait level at which the next response becomes 50%. So in women, it takes a lesser amount of the latent trait to be more able to perform these items.



The difference in the loading for item 3 (not pictured) also indicates that cooking is less discriminating (lower a-value) in men than in women (i.e., it is less useful in men for measuring ability to live independently at its b_i locations).

In addition, the women in this sample are less variable and less able to live independently than the men, as given by the factor variance and factor mean differences, respectively.

```
MODEL:  !!! FINAL MODEL 4c FOR MEN REFERENCE GROUP

! Factor loadings all freely estimated, just labeled
  IADL BY cia1-cia7* (L_I1-L_I7);
! Item thresholds all freely estimated, just labeled
  [cia1$1-cia7$1*] (T1_I1-T1_I7);
  [cia1$2-cia7$2*] (T2_I1-T2_I7);
  [cia1$3-cia7$3*] (T3_I1-T3_I7);
! Item residual variances all fixed=1 for identification
  cia1-cia7@1;
! Factor mean=0 and variance=1 for identification
  [IADL@0]; IADL@1;

MODEL WOMEN: !!! FINAL MODEL 4c FOR WOMEN ALTERNATIVE GROUP
! Factor loadings CONSTRAINED EQUAL TO MEN EXCEPT ITEM 3
  IADL BY cia1-cia7* (L_I1 L_I2 L_I3w L_I4 L_I5 L_I6 L_I7);
! Item thresholds CONSTRAINED EQUAL TO MEN BY LABELS EXCEPT "w"
  [cia1$1-cia7$1*] (T1_I1  T1_I2  T1_I3w T1_I4 T1_I5 T1_I6 T1_I7);
  [cia1$2-cia7$2*] (T2_I1w T2_I2  T2_I3w T2_I4 T2_I5 T2_I6 T2_I7);
  [cia1$3-cia7$3*] (T3_I1w T3_I2w T3_I3w T3_I4 T3_I5 T3_I6 T3_I7);
! Item residual variances HELD EQUAL EXCEPT 2 AND 7
  cia1@1 cia2* cia3-cia6@1 cia7*;
! Factor mean and variance ESTIMATED
  [IADL*]; IADL*;

MODEL CONSTRAINT: ! IRT parms using current factor mean and variance
! DO (begin, end), replace # with index
! A = discrimination, B1=y>0, B2=y>1, B3=y>2
! Ref group (men) parms
  NEW(Am_I1-Am_I7 B1m_I1-B1m_I7 B2m_I1-B2m_I7 B3m_I1-B3m_I7);
  ! Discriminations
    DO (1,7) Am_I# = L_I#;
  ! Difficulties
    DO (1,7) B1m_I# = T1_I# / L_I#;
    DO (1,7) B2m_I# = T2_I# / L_I#;
    DO (1,7) B3m_I# = T3_I# / L_I#;
! Alt group (women) parms
  NEW(Aw_I1-Aw_I7 B1w_I1-B1w_I7 B2w_I1-B2w_I7 B3w_I1-B3w_I7);
  ! Discriminations
    DO (1,2) Aw_I# = L_I#;
             Aw_I3 = L_I3w; ! w marks diff loading for women
    DO (4,7) Aw_I# = L_I#;
  ! Difficulties for Thresh 1
    DO (1,2) B1w_I# = T1_I#  / L_I#;
             B1w_I3 = T1_I3w / L_I3w; ! w marks diff thresh for women
    DO (4,7) B1w_I# = T1_I#  / L_I#;
  ! Difficulties for Thresh 2
             B2w_I1 = T2_I1w / L_I1;
             B2w_I2 = T2_I2  / L_I2;
             B2w_I3 = T2_I3w / L_I3w;
    DO (4,7) B2w_I# = T2_I#  / L_I#;
  ! Difficulties for Thresh 3
             B3w_I1 = T3_I1w / L_I1;
             B3w_I2 = T3_I2w / L_I2;
             B3w_I3 = T3_I3w / L_I3w;
    DO (4,7) B3w_I# = T2_I#  / L_I#;
```

New/Additional Parameters

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
AM_I1	4.614	0.532	8.674	0.000
AM_I2	3.191	0.506	6.309	0.000
AM_I3	3.227	0.467	6.912	0.000
AM_I4	4.607	0.571	8.062	0.000
AM_I5	3.009	0.352	8.548	0.000
AM_I6	2.519	0.292	8.614	0.000
AM_I7	1.123	0.174	6.471	0.000
B1M_I1	-1.390	0.117	-11.877	0.000
B1M_I2	-1.476	0.125	-11.816	0.000
B1M_I3	-1.251	0.136	-9.224	0.000
B1M_I4	-1.320	0.113	-11.628	0.000
B1M_I5	-1.618	0.138	-11.698	0.000
B1M_I6	-1.683	0.142	-11.891	0.000
B1M_I7	-2.693	0.287	-9.370	0.000
B2M_I1	-0.882	0.111	-7.977	0.000
B2M_I2	-1.278	0.110	-11.587	0.000
B2M_I3	-0.906	0.115	-7.871	0.000
B2M_I4	-0.986	0.096	-10.239	0.000
B2M_I5	-0.938	0.095	-9.893	0.000
B2M_I6	-1.238	0.109	-11.341	0.000
B2M_I7	-2.171	0.218	-9.946	0.000
B3M_I1	-0.143	0.096	-1.492	0.136
B3M_I2	-0.711	0.101	-7.051	0.000
B3M_I3	-0.182	0.098	-1.865	0.062
B3M_I4	-0.526	0.085	-6.186	0.000
B3M_I5	-0.427	0.086	-4.952	0.000
B3M_I6	-0.851	0.091	-9.331	0.000
B3M_I7	-1.555	0.150	-10.365	0.000
AW_I1	4.614	0.532	8.674	0.000
AW_I2	3.191	0.506	6.309	0.000
AW_I3	6.758	1.277	5.294	0.000
AW_I4	4.607	0.571	8.062	0.000
AW_I5	3.009	0.352	8.548	0.000
AW_I6	2.519	0.292	8.614	0.000
AW_I7	1.123	0.174	6.471	0.000
B1W_I1	-1.390	0.117	-11.877	0.000
B1W_I2	-1.476	0.125	-11.816	0.000
B1W_I3	-1.417	0.120	-11.836	0.000
B1W_I4	-1.320	0.113	-11.628	0.000
B1W_I5	-1.618	0.138	-11.698	0.000
B1W_I6	-1.683	0.142	-11.891	0.000
B1W_I7	-2.693	0.287	-9.370	0.000
B2W_I1	-1.127	0.103	-10.935	0.000
B2W_I2	-1.278	0.110	-11.587	0.000
B2W_I3	-1.264	0.110	-11.444	0.000
B2W_I4	-0.986	0.096	-10.239	0.000
B2W_I5	-0.938	0.095	-9.893	0.000
B2W_I6	-1.238	0.109	-11.341	0.000
B2W_I7	-2.171	0.218	-9.946	0.000
B3W_I1	-0.542	0.088	-6.126	0.000
B3W_I2	-0.957	0.097	-9.914	0.000
B3W_I3	-0.939	0.095	-9.922	0.000
B3W_I4	-0.986	0.096	-10.239	0.000
B3W_I5	-0.938	0.095	-9.893	0.000
B3W_I6	-1.238	0.109	-11.341	0.000
B3W_I7	-2.171	0.218	-9.946	0.000

Example write-up of these IFA analyses:

The extent to which an item factor model for a trait of independent daily living (with seven observed items) exhibited measurement and structural invariance between men and women was examined using M*plus* v. 8.10 (Muthén & Muthén, 1998–2017). WLSMV limited-information estimation (i.e., diagonally weighted least squares) using a probit link and the THETA parameterization was used for all models. Thus, model fit statistics describe the fit of the item factor model to the item polychoric correlation matrix for each group. Nested model comparisons were conducted using the M*plus* DIFFTEST procedure. A configural invariance model was initially specified in which a single factor was estimated simultaneously in each group. The factor variance was fixed to 1 and the factor mean was fixed to 0 in each group for identification, such that all item factor loadings (one per item) and thresholds (three per item given four response options) were then estimated. The residual variances are not uniquely identified in the configural invariance model and as such were all constrained to 1 in both groups. As shown in Table 1, the configural invariance model had good fit. The analysis proceeded by applying parameter constraints in successive models to examine potential decreases in fit resulting from measurement or structural non-invariance between men and women, with men as the reference group.

Equality of the unstandardized item factor loadings between groups was then examined in a metric invariance model. The factor variance was fixed to 1 in men for identification but was freely estimated in women; the factor mean was fixed to 0 in both groups for identification. All factor loadings were constrained equal across groups, all item thresholds were estimated, and all residual variances were constrained to 1 across groups. Although the metric invariance model did not fit significantly worse than the configural invariance model, DIFFTEST (6) = 9.40, $p = .15$, the modification indices suggested localized misfit for the constrained loading of item 3 (cooking). After freeing its loading between groups, the partial metric invariance model also did not fit significantly worse than the configural invariance model, DIFFTEST (5) = 5.53, $p = .36$, and no loadings were indicated as individually problematically different. The fact that partial metric invariance (i.e., “weak invariance”) held indicates that the items were related to the latent factor equivalently across groups, or more simply, that the same latent factor was being measured in each group. The exception is for the cooking item 3, which was less related to the trait of independent daily living in men than in women.

Equality of the unstandardized item thresholds across groups was then examined in a scalar invariance model. The factor mean and variance were fixed to 0 and 1, respectively, in men for identification, but the factor mean and variance were then estimated for women. All factor loadings (except item 3) and all item thresholds were constrained equal across groups; all residual variances were still constrained equal to 1 in both groups. The full scalar invariance model A fit significantly worse than the partial metric invariance model, DIFFTEST (20) = 91.66, $p < .01$. Not surprisingly given its differential loading across groups, the modification indices suggested that item 3 was the largest source of the misfit. After freeing item 3’s thresholds, the partial scalar invariance model B still fit significantly worse than the partial metric invariance model, DIFFTEST (17) = 44.80, $p < 01$. The modification indices suggested that item 1 threshold 3 was the next largest remaining source of the misfit and should be freed. After doing so, the new partial scalar invariance model C (with all item 3 thresholds and item 1 threshold 3 freed) still fit significantly worse than the partial metric invariance model, DIFFTEST (16) = 32.46, $p < .01$. The modification indices suggested that item 2 threshold 3 was the largest remaining source of the misfit and should be freed. After doing so, the new partial scalar invariance model D (with all item 3 thresholds, item 2 threshold 3, and item 1 threshold 3 freed) still fit significantly worse than the partial metric invariance model, DIFFTEST (15) = 26.17, $p = .04$. The modification indices suggested that item 1 threshold 2 was the largest remaining source of the misfit and should be freed. After doing so, the new partial scalar invariance model E (with all thresholds for item 3, thresholds 2 and 3 for item 1, and threshold 3 for item 2 freed) did not fit significantly worse than the partial metric invariance model, DIFFTEST (14) = 20.32, $p = .12$. The factor that partial scalar invariance (i.e., “strong invariance”) held indicates that items 4, 5, 6, and 7 have the same expected response at the same level of the trait, or more simply, that the observed differences in the proportion of responses in each category for those items was due to factor mean differences only. However, at the same absolute level of the IADL factor, thresholds 2 and 3 for item 1, threshold 3 for item 2, and all thresholds for item 3 were more difficult for men, indicating that men had a greater probability of not being able to do housework, make the bed, or cook, respectively, at the same level of the latent trait.

Equality of the unstandardized residual variances across groups was then examined in a residual variance invariance model. The model comparison at this step proceeded backwards, such that fitted first was a residual variance A model with all residual variances freely estimated in the women (except for item 3, whose residual variance could not be estimated given that its factor loading and thresholds differed across groups). This was then compared with a model in which all residual variances were fixed to 1 in the women (residual variance B; in which all model parameters were estimated as described for the last partial scalar invariance model E). Although the model B with the residual variances constrained to 1 (to be equal to the men) fit nonsignificantly worse than the A

model with those residual variances freed, DIFFTEST (6) = 11.02, p = .09, the modification indices suggested that the residual variances for items 2 and 7 (bed-making and answering the telephone) were a source of local misfit and should be freed. After doing so, the new partial residual variance invariance model C did not fit significantly worse than the model with free residual variances in the women group, DIFFTEST (4) = 2.80, p = .59, indicating that residual variances for items 2 and 7 were significantly smaller for women than men. The fact that partial residual variance invariance (i.e., “strict invariance”) held indicates that the amount of item variance not attributable to the factor was the same across groups in the five other items.

After achieving partial measurement invariance as was just described, structural invariance was then tested with one additional model. The factor variance in the women (which had been estimated freely) was constrained to 1 (i.e., to be equal to the factor variance in men), resulting in a significant decrease in fit relative to the last partial residual invariance model C, DIFFTEST (1) = 11.50, p < .01. Thus, women showed significantly less variability in ability to live independently (factor variance of 0.49) than did men (factor variance fixed = 1 for identification). The factor mean for women in the partial measurement invariance model was significantly different from 0 as indicated by a Wald test (difference = −0.38, SE = 0.10, p < .01), indicating that women on average were significantly less able to live independently on average than men (factor mean fixed = 0 for identification).

In conclusion, these analyses showed that partial measurement invariance was obtained across men and women—that is, the relationships of four items to the latent factor of independent living were equivalent between men and women. However, items 1, 2, and 3 (housework, bedmaking, and cooking) were systematically more difficult for men than women at the same level of the latent trait. Structural invariance was not obtained, such that women were less variable and less able on average than men in their latent traits. Model parameters from the final model are given in Table 2.

(see Example 7c spreadsheet for Table 1 and figures; Table 2 would have unstandardized and standardized estimates and their SEs)

Reference: Muthén, L. K., & Muthén, B.O. (1998–2017). *Mplus user’s guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.