IRT Models Using Structural Equation Modeling

Basic Two-Parameter IRT Model

Below I test a confirmatory factor model with binary indicators using the ICAR verbal ability items. I use a maximum likelihood estimation and in Mplus, by default, the results are in the logistic metric (but link = probit can be specified on under ANALYSIS: or probit is the default parameterization with estimator=WLSMV estimator). Remember, however, that even with the logit parameterization some translation is need to get the IRT difficulty parameter values, b. The "IRT parameterization" section gives the difficulty parameters in terms of the IRT values, where $b = \alpha/-a$. In the Mplus output, the intercept (α in our logistic notation) from the logistic models are called "thresholds."

```
Mplus VERSION 8.5
MUTHEN & MUTHEN
05/18/2021  11:11 AM

INPUT INSTRUCTIONS


  title:  ICAR verbal ability items IRT model;

  data:  file=icaritems.dat; format=free;

  variable:  names = v2 v4 v5 v6 v8  sex;


            usevariable=v2 v4 v5 v6 v8;

            categorical=v2 v4 v5 v6 v8;

  analysis:  type=general;  estimator=mlr;
      !estimator=mlr invokes full maximum likelihood with robust adjustments
      ! when categorical variables are identified. Estimates are logit with this
      ! method and odds ratios can be used with predictive paths;

  model: icar by v2* v4 v5 v6 v8;
         icar@1;
      !the second statement icar@1 standardizes the ability factor;

  output: stdyx tech1 tech8;

  plot: type = plot3;



INPUT READING TERMINATED NORMALLY



ICAR verbal ability items IRT model;

SUMMARY OF ANALYSIS

Number of groups                                                 1
Number of observations                                         198

Number of dependent variables                                    5
Number of independent variables                                  0
Number of continuous latent variables                            1

Observed dependent variables

  Binary and ordered categorical (ordinal)
   V2          V4          V5          V6          V8
```

```
Continuous latent variables
   ICAR

Estimator                                                      MLR
Information matrix                                        OBSERVED
Optimization Specifications for the Quasi-Newton Algorithm for
Continuous Outcomes
  Maximum number of iterations                                 100
  Convergence criterion                                  0.100D-05
Optimization Specifications for the EM Algorithm
  Maximum number of iterations                                 500
  Convergence criteria
    Loglikelihood change                                 0.100D-02
    Relative loglikelihood change                        0.100D-05
    Derivative                                           0.100D-02
Optimization Specifications for the M step of the EM Algorithm for
Categorical Latent variables
  Number of M step iterations                                    1
  M step convergence criterion                           0.100D-02
  Basis for M step termination                           ITERATION
Optimization Specifications for the M step of the EM Algorithm for
Censored, Binary or Ordered Categorical (Ordinal), Unordered
Categorical (Nominal) and Count Outcomes
  Number of M step iterations                                    1
  M step convergence criterion                           0.100D-02
  Basis for M step termination                           ITERATION
  Maximum value for logit thresholds                            15
  Minimum value for logit thresholds                           -15
  Minimum expected cell size for chi-square              0.100D-01
Optimization algorithm                                         EMA
Integration Specifications
  Type                                                    STANDARD
  Number of integration points                                  15
  Dimensions of numerical integration                            1
  Adaptive quadrature                                           ON
Link                                                         LOGIT
Cholesky                                                        ON

Input data file(s)
  icaritems.dat
Input data format  FREE


UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

    V2
      Category 1    0.172           34.000
      Category 2    0.828          164.000
    V4
      Category 1    0.293           58.000
      Category 2    0.707          140.000
    V5
      Category 1    0.409           81.000
      Category 2    0.591          117.000
    V6
      Category 1    0.237           47.000
      Category 2    0.763          151.000
    V8
      Category 1    0.374           74.000
      Category 2    0.626          124.000

THE MODEL ESTIMATION TERMINATED NORMALLY
MODEL FIT INFORMATION

Number of Free Parameters                        10

Loglikelihood

          H0 Value                        -527.095
          H0 Scaling Correction Factor      1.0042
            for MLR
```

```
Information Criteria

          Akaike (AIC)                    1074.190
          Bayesian (BIC)                  1107.073
          Sample-Size Adjusted BIC        1075.393
            (n* = (n + 2) / 24)

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

          Pearson Chi-Square

          Value                             17.438
          Degrees of Freedom                    21
          P-Value                           0.6843

          Likelihood Ratio Chi-Square

          Value                             19.670
          Degrees of Freedom                    21
          P-Value                           0.5422


MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 ICAR     BY
    V2                 1.652      0.490      3.370      0.001
    V4                 1.899      0.521      3.649      0.000
    V5                 1.599      0.440      3.635      0.000
    V6                 1.756      0.480      3.659      0.000
    V8                 1.507      0.374      4.030      0.000

 Thresholds
    V2$1              -2.264      0.397     -5.698      0.000
    V4$1              -1.396      0.324     -4.309      0.000
    V5$1              -0.536      0.223     -2.397      0.017
    V6$1              -1.756      0.340     -5.159      0.000
    V8$1              -0.729      0.222     -3.286      0.001

 Variances
    ICAR               1.000      0.000    999.000    999.000


RESULTS IN PROBABILITY SCALE

                    Estimate

 V2
    Category 1         0.172
    Category 2         0.828
 V4
    Category 1         0.293
    Category 2         0.707
 V5
    Category 1         0.409
    Category 2         0.591
 V6
    Category 1         0.237
    Category 2         0.763
 V8
    Category 1         0.374
    Category 2         0.626
```

IRT PARAMETERIZATION

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Item Discriminations				
ICAR BY				
V2	1.652	0.490	3.370	0.001
V4	1.899	0.521	3.649	0.000
V5	1.599	0.440	3.635	0.000
V6	1.756	0.480	3.659	0.000
V8	1.507	0.374	4.030	0.000
Item Difficulties				
V2	-1.370	0.270	-5.071	0.000
V4	-0.735	0.155	-4.742	0.000
V5	-0.335	0.138	-2.421	0.015
V6	-1.000	0.195	-5.140	0.000
V8	-0.484	0.152	-3.191	0.001
Variances				
ICAR	1.000	0.000	0.000	1.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
ICAR BY				
V2	0.673	0.109	6.166	0.000
V4	0.723	0.095	7.649	0.000
V5	0.661	0.102	6.459	0.000
V6	0.696	0.098	7.091	0.000
V8	0.639	0.094	6.810	0.000
Thresholds				
V2$1	-0.923	0.104	-8.891	0.000
V4$1	-0.531	0.092	-5.754	0.000
V5$1	-0.222	0.087	-2.549	0.011
V6$1	-0.696	0.096	-7.245	0.000
V8$1	-0.309	0.087	-3.535	0.000
Variances				
ICAR	1.000	0.000	999.000	999.000

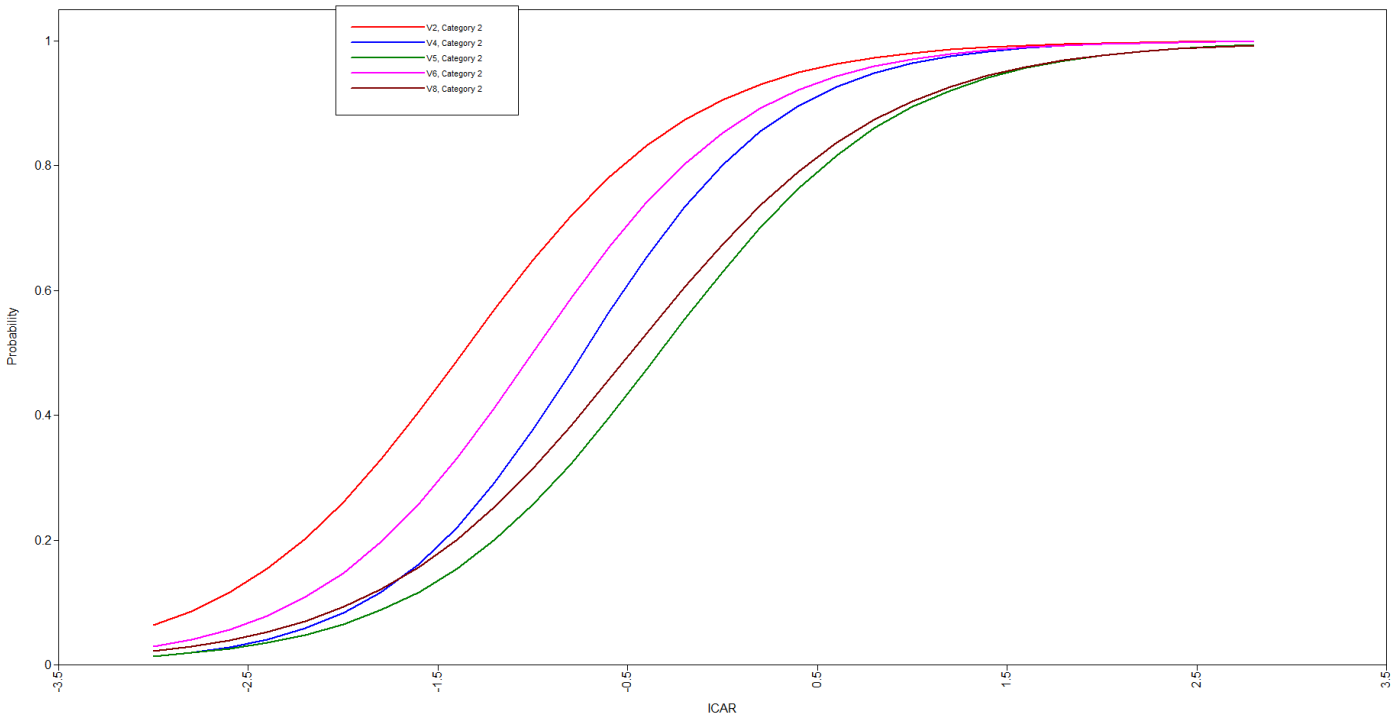
R-SQUARE

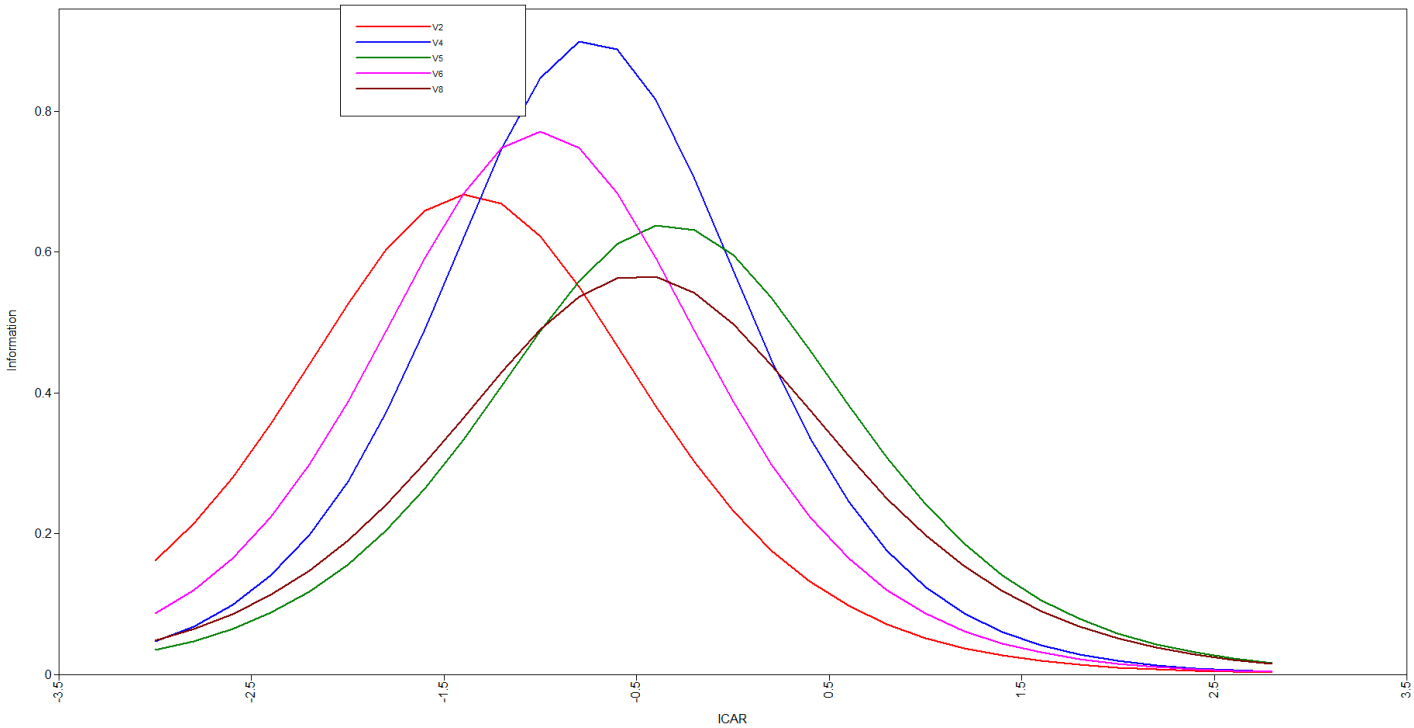
Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
V2	0.453	0.147	3.083	0.002
V4	0.523	0.137	3.824	0.000
V5	0.437	0.135	3.230	0.001
V6	0.484	0.136	3.545	0.000
V8	0.408	0.120	3.405	0.001

QUALITY OF NUMERICAL RESULTS

Condition Number for the Information Matrix 0.642E-01
(ratio of smallest to largest eigenvalue)

Plots for all of the items together can be obtained





Example of a DIF Test using the KNOWNCLASS Specification

Two models are tested to compare the fit of the model with the v8 loading (discrimination parameter) constrained across groups to a model in which the loadings are allowed to differ. The results below are from the model in which the parameters are all allowed to differ (the difficulty parameters match the results from SAS PROC IRT, but the discrimination parameters do not). I omit the results from the constrained model, except for the model fit information

Unconstrained Model

```
Mplus VERSION 8.5
MUTHEN & MUTHEN
05/19/2021  11:08 AM

INPUT INSTRUCTIONS


  title:  ICAR verbal ability items IRT model;

  data:  file=icaritems.dat; format=free;

  variable:  names = v2 v4 v5 v6 v8  sex;
          !sex (0=male,1=female);

          classes = sexgrp (2);
          knownclass = sexgrp (sex=0 sex=1);

          usevariable=v2 v4 v5 v6 v8;

          categorical=v2 v4 v5 v6 v8;

  analysis:  type=mixture; estimator=ml; algorithm=integration;
      link=logit;

  model:
       %overall%
       icar by v2* v4 v5 v6
        v8 ;
       icar@1;


       %sexgrp#1%
       icar by v2* v4 v5 v6
        v8 ;
       icar@1;

      %sexgrp#2%
       icar by v2* v4 v5 v6
        v8 ;
       icar@1;


  output: stdyx tech1 tech8;

  plot: type = plot3;


INPUT READING TERMINATED NORMALLY

ICAR verbal ability items IRT model;

SUMMARY OF ANALYSIS

Number of groups                                                 1
Number of observations                                         198

Number of dependent variables                                    5
Number of independent variables                                  0
Number of continuous latent variables                            1
Number of categorical latent variables                           1
```

```
Observed dependent variables

  Binary and ordered categorical (ordinal)
   V2          V4          V5          V6          V8

Continuous latent variables
   ICAR

Categorical latent variables
   SEXGRP

  Knownclass            SEXGRP


Estimator                                                       ML
Information matrix                                        OBSERVED
Optimization Specifications for the Quasi-Newton Algorithm for
Continuous Outcomes
  Maximum number of iterations                                 100
  Convergence criterion                                  0.100D-05
Optimization Specifications for the EM Algorithm
  Maximum number of iterations                                 500
  Convergence criteria
    Loglikelihood change                                 0.100D-02
    Relative loglikelihood change                        0.100D-05
    Derivative                                           0.100D-02
Optimization Specifications for the M step of the EM Algorithm for
Categorical Latent variables
  Number of M step iterations                                    1
  M step convergence criterion                           0.100D-02
  Basis for M step termination                           ITERATION
Optimization Specifications for the M step of the EM Algorithm for
Censored, Binary or Ordered Categorical (Ordinal), Unordered
Categorical (Nominal) and Count Outcomes
  Number of M step iterations                                    1
  M step convergence criterion                           0.100D-02
  Basis for M step termination                           ITERATION
  Maximum value for logit thresholds                            15
  Minimum value for logit thresholds                           -15
  Minimum expected cell size for chi-square              0.100D-01
Optimization algorithm                                         EMA
Integration Specifications
  Type                                                    STANDARD
  Number of integration points                                   15
  Dimensions of numerical integration                             1
  Adaptive quadrature                                            ON
Link                                                        LOGIT
Cholesky                                                       ON

Input data file(s)
  icaritems.dat
Input data format  FREE

UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES

    V2
      Category 1    0.172          34.000
      Category 2    0.828         164.000
    V4
      Category 1    0.293          58.000
      Category 2    0.707         140.000
    V5
      Category 1    0.409          81.000
      Category 2    0.591         117.000
    V6
      Category 1    0.237          47.000
      Category 2    0.763         151.000
    V8
      Category 1    0.374          74.000
      Category 2    0.626         124.00
```

```
THE MODEL ESTIMATION TERMINATED NORMALLY


MODEL FIT INFORMATION

Number of Free Parameters                       17

Loglikelihood

          H0 Value                        -643.662

Information Criteria

          Akaike (AIC)                    1321.324
          Bayesian (BIC)                  1377.224
          Sample-Size Adjusted BIC        1323.368
            (n* = (n + 2) / 24)

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

          Pearson Chi-Square

          Value                             60.232
          Degrees of Freedom                    46
          P-Value                           0.0776

          Likelihood Ratio Chi-Square

          Value                             51.917
          Degrees of Freedom                    46
          P-Value                           0.2543


FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASSES
BASED ON THE ESTIMATED MODEL

    Latent
   Classes

       1         65.00000          0.32828
       2        133.00000          0.67172


MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Latent Class 1 (0)

 ICAR     BY
    V2                 2.546      0.790      3.222      0.001
    V4                 5.914      4.547      1.300      0.193
    V5                 1.686      0.687      2.453      0.014
    V6                 1.178      0.465      2.536      0.011
    V8                 2.691      1.159      2.322      0.020

 Means
    ICAR               0.046      0.195      0.237      0.813

 Thresholds
    V2$1              -2.036      0.306     -6.660      0.000
    V4$1              -5.495      4.017     -1.368      0.171
    V5$1              -0.444      0.214     -2.070      0.038
    V6$1              -1.639      0.318     -5.152      0.000
    V8$1              -0.686      0.246     -2.789      0.005
```

```
Variances
    ICAR                1.000      0.000    999.000    999.000

Latent Class 2 (1)

 ICAR     BY
    V2                  0.785      0.341      2.306      0.021
    V4                 14.021      9.143      1.533      0.125
    V5                  0.903      0.285      3.172      0.002
    V6                  1.752      0.435      4.027      0.000
    V8                  1.007      0.303      3.321      0.001

 Means
    ICAR                0.000      0.000    999.000    999.000

 Thresholds
    V2$1               -2.036      0.306     -6.660      0.000
    V4$1               -5.495      4.017     -1.368      0.171
    V5$1               -0.444      0.214     -2.070      0.038
    V6$1               -1.639      0.318     -5.152      0.000
    V8$1               -0.686      0.246     -2.789      0.005

 Variances
    ICAR                1.000      0.000    999.000    999.000

Categorical Latent Variables

 Means
    SEXGRP#1           -0.716      0.151     -4.731      0.000


IRT PARAMETERIZATION

Latent Class 1 (0)

 Item Discriminations

 ICAR     BY
    V2                  2.546      0.790      3.222      0.001
    V4                  5.914      4.547      1.300      0.193
    V5                  1.686      0.687      2.453      0.014
    V6                  1.178      0.465      2.536      0.011
    V8                  2.691      1.159      2.322      0.020

 Means
    ICAR                0.000      0.000      0.000      1.000

 Item Difficulties
    V2                 -0.846      0.235     -3.593      0.000
    V4                 -0.975      0.213     -4.582      0.000
    V5                 -0.309      0.214     -1.446      0.148
    V6                 -1.437      0.517     -2.776      0.005
    V8                 -0.301      0.184     -1.640      0.101

 Variances
    ICAR                1.000      0.000      0.000      1.000

Latent Class 2 (1)

 Item Discriminations

 ICAR     BY
    V2                  0.785      0.341      2.306      0.021
    V4                 14.021      9.143      1.533      0.125
    V5                  0.903      0.285      3.172      0.002
    V6                  1.752      0.435      4.027      0.000
    V8                  1.007      0.303      3.321      0.001
```

```
Means
    ICAR               0.000      0.000      0.000      1.000

 Item Difficulties
    V2                -2.592      0.985     -2.632      0.008
    V4                -0.392      0.174     -2.254      0.024
    V5                -0.491      0.261     -1.884      0.060
    V6                -0.935      0.223     -4.193      0.000
    V8                -0.681      0.270     -2.523      0.012

 Variances
    ICAR               1.000      0.000      0.000      1.000


RESULTS IN PROBABILITY SCALE

                    Estimate

Latent Class 1 (0)

 V2
    Category 1         0.243
    Category 2         0.757
 V4
    Category 1         0.175
    Category 2         0.825
 V5
    Category 1         0.414
    Category 2         0.586
 V6
    Category 1         0.206
    Category 2         0.794
 V8
    Category 1         0.400
    Category 2         0.600

Latent Class 2 (1)

 V2
    Category 1         0.139
    Category 2         0.861
 V4
    Category 1         0.349
    Category 2         0.651
 V5
    Category 1         0.407
    Category 2         0.593
 V6
    Category 1         0.252
    Category 2         0.748
 V8
    Category 1         0.362
    Category 2         0.638


STANDARDIZED MODEL RESULTS


STDYX Standardization

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Latent Class 1 (0)

 ICAR     BY
    V2                 0.814      0.085      9.573      0.000
    V4                 0.956      0.063     15.125      0.000
    V5                 0.681      0.149      4.571      0.000
    V6                 0.545      0.151      3.606      0.000
    V8                 0.829      0.112      7.431      0.000
```

```
 Means
    ICAR               0.046      0.195      0.237      0.813

 Thresholds
    V2$1              -0.651      0.149     -4.357      0.000
    V4$1              -0.888      0.204     -4.358      0.000
    V5$1              -0.179      0.087     -2.067      0.039
    V6$1              -0.758      0.145     -5.207      0.000
    V8$1              -0.211      0.097     -2.170      0.030

 Variances
    ICAR               1.000      0.000    999.000    999.000

Latent Class 2 (1)

 ICAR     BY
    V2                 0.397      0.145      2.738      0.006
    V4                 0.992      0.011     93.161      0.000
    V5                 0.446      0.113      3.958      0.000
    V6                 0.695      0.089      7.784      0.000
    V8                 0.485      0.112      4.344      0.000

 Means
    ICAR               0.000      0.000    999.000    999.000

 Thresholds
    V2$1              -1.030      0.133     -7.722      0.000
    V4$1              -0.389      0.172     -2.262      0.024
    V5$1              -0.219      0.104     -2.098      0.036
    V6$1              -0.650      0.115     -5.631      0.000
    V8$1              -0.331      0.114     -2.900      0.004

 Variances
    ICAR               1.000      0.000    999.000    999.000


R-SQUARE

Class 1

    Observed                                           Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value

    V2                 0.663      0.139      4.786      0.000
    V4                 0.914      0.121      7.563      0.000
    V5                 0.463      0.203      2.286      0.022
    V6                 0.297      0.165      1.803      0.071
    V8                 0.688      0.185      3.716      0.000

Class 2

    Observed                                           Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value

    V2                 0.158      0.115      1.369      0.171
    V4                 0.984      0.021     46.581      0.000
    V5                 0.199      0.100      1.979      0.048
    V6                 0.483      0.124      3.892      0.000
    V8                 0.236      0.108      2.172      0.030


QUALITY OF NUMERICAL RESULTS

     Condition Number for the Information Matrix              0.527E-04
       (ratio of smallest to largest eigenvalue)
```

Constrained Model Fit

The (1) behind the v8 item constrains the loading (discrimination parameter) to be equal across groups for item v8.

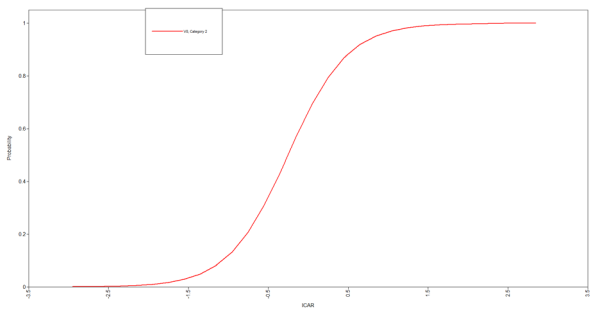
```
  title:  ICAR verbal ability items IRT model;

  data:  file=icaritems.dat; format=free;

  variable:  names = v2 v4 v5 v6 v8  sex;
          !sex (0=male,1=female);

          classes = sexgrp (2);
          knownclass = sexgrp (sex=0 sex=1);

          usevariable=v2 v4 v5 v6 v8;

          categorical=v2 v4 v5 v6 v8;

  analysis:  type=mixture; estimator=ml; algorithm=integration;
      link=logit;

  model:
       %overall%
       icar by v2* v4 v5 v6
        v8 (1);
       icar@1;


       %sexgrp#1%
       icar by v2* v4 v5 v6
        v8 (1);
       icar@1;

      %sexgrp#2%
       icar by v2* v4 v5 v6
        v8 (1);
       icar@1;


  output: stdyx tech1 tech8;

  plot: type = plot3;


MODEL FIT INFORMATION

Number of Free Parameters                       16

Loglikelihood

          H0 Value                        -645.566

Information Criteria

          Akaike (AIC)                    1323.133
          Bayesian (BIC)                  1375.745
          Sample-Size Adjusted BIC        1325.057
            (n* = (n + 2) / 24)

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

          Pearson Chi-Square

          Value                             65.466
          Degrees of Freedom                    47
          P-Value                           0.0386

          Likelihood Ratio Chi-Square
```

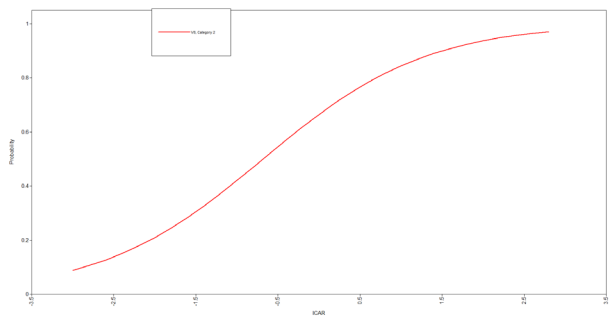
```
Value                              56.263
Degrees of Freedom                     47
P-Value                            0.1668
```

The fit from the unconstrained model was LR χ^2 (47) = 51.917 and the fit from the constrained model was LR χ^2 (46) = 56.263, which has a χ^2 difference of 4.326, which, with 1 df, is significant.

ICC Plots for v8 Item from the unconstrained model Prior version switched plots for Males and Females

Males Females





Information plots *for v8 Item from the unconstrained model*

Males *Females*