PARSCALE Output

In the following we provide two examples using the PARSCALE calibration program. The first example uses the generalized rating scale (GRS) model and the second demonstrates performing a calibration with the generalized partial credit (GPC) model.

PARSCALE 4.1 is a calibration program that uses MMLE for item parameter estimation for the one-, two-, three-parameter models and several polytomous models (e.g., the PC, GPC, GRS, GR models). In this appendix we demonstrate the use of PARSCALE for the GRS and GPC models. One unique aspect of PARSCALE is the use of an item "block" to address the issue of indeterminacy of metric. A block consists of 1 or more items that share the same set of parameters. An instrument's calibration may consist of single or multiple blocks. Below we show the use of single and multiple blocks by re-calibrating the Attitude Towards Condoms scale and the Alike Reasoning instrument.

Example: Generalized Rating Scale Model using PARSCALE

In this example we use PARSCALE for calibrating the six-item Attitude Towards Condoms scale with the GRS model. In the current context, our "block" consists of all 6 items so that they share the same set of threshold parameters. To obtain a unique set of parameter estimates for a block the constraint $\sum_{h=2}^{m_j} \tau_h = 0$ is used to ensure that the mean of the threshold parameters is constant over blocks. Moreover, with PARSCALE it is possible to vary the block composition to compare the relative fit of constraining two or more items to have a common set of thresholds versus allowing these items to have unique threshold estimates. In short, we do not have to assume that all items should have the same threshold values, but rather can test this assumption. However, the discrimination parameter estimates are directly comparable only when the items share the same set of threshold parameters (Muraki, 1992).

Recall that our data use a Likert response format with 1="strongly disagree," 2="disagree more than I agree," 3="agree more than I disagree, " and 4="strongly agree." The input command file for the analysis of the attitudes toward condoms scale is shown in Table D.1. In our example the data conform to PARSCALE's assumption that the data consist of coded values beginning with 1; if this were not the case, then the data would have to be recoded using the `ORIGINAL` keyword. The input command file has a structure similar to those seen with BILOG and MULTILOG. On the `CALIB` line we instruct PARSCALE to perform a generalized partial

credit calibration using the LOGISTIC metric and fifteen quadrature points. In addition, we specify the

maximum number of EM cycles to be 100, the maximum number of Newton iterations to be 20, and to use a

convergence criterion of 0.01. Because only a single BLOCK is specified (see the TEST command line

subcommand NBLOCK=1), a single set of category parameters (τ_h) is estimated for the entire instrument. The

combination of these commands results in the GRS model. In addition, in our analysis we allow the item

discrimination parameter estimates to vary across items. The corresponding output is presented in Table D.2.

Table D.1. Command file for PARSCALE GRS model calibration example.

```
PARSCALE CALIBRATION-RS
DATA ARE CONDOM ATTITUDE INSTRUMENT
>FILES   DFNAME='condoms.dat';
>INPUT   NIDCHAR=4, NTOTAL=6, NTEST=1, LENGTH=6;
(4A1,6A1)
>TEST    TNAME='SCALE1', ITEM=(1(1)6), NBLOCK=1;
>BLOCK1  BNAME='SBLOCK1', NITEMS=6, NCAT=4, ORIGINAL=(1,2,3,4);
>CALIB   PARTIAL, LOGISTIC, NQPTS=15, CYCLE=(100,1,1,1,1), NEWTON=20, CRIT=0.01;
```

Output from PARSCALE is presented in four phases. Phase 0 contains an echo of the input commands

followed by the default and user-specified program parameters (e.g., number of EM cycles, convergence criteria,

etc.) and an echo of the first two cases' response vectors. These echoed vectors should be inspected to verify

that the data were read correctly.

Phase 1 contains traditional item statistics and frequency distributions. For example, from the SUMMARY

ITEM STATISTICS section we see that 3473 individuals responded to item 1 with no omissions. Of these

3473, 1338 (38.5%) "strongly disagree" to the item, 492 (14.2%) "disagreed more than agreed" with the item,

602 (17.3%) "agreed more than disagreed," and 1041 (30%) "strongly agreed "to the item. PARSCALE

provides two indices of the relationship between how individuals respond to an item and their observed scores.

The first of these is the PEARSON correlation between the responses to the item and the observed scores. The

second is the polyserial correlation. The polyserial correlation is the correlation between an ordinal manifest

variable and a continuous latent variable; the manifest and latent variables are assumed to be bivariate normally

distributed. For item 1 the relationship between the Likert response scale and the underlying trait of 0.681 is

somewhat reasonable. The mean response to item 1 (RESPONSE MEAN) is 2.388 on the 4-point response scale.

That is, individuals were more likely to be agreeing than disagreeing with this item.

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

Phase 2 presents the IRT results. This output begins with the iteration history followed by the item parameter estimates. The SLOPE column contains the estimates of the discrimination parameters. For item 1 we have an item discrimination estimate, $\hat{\alpha}_1$, of 0.270. The physical location of the transition locations on the continuum are obtained in a manner similar to that for the RS model. The STEP PARAMTER [sic] line contains the offset values (i.e., τ_hs) and the column labeled LOCATION contain the items' $\hat{\delta}_j$ s. As would be expected, the sum of the thresholds sum to zero. To obtain the transition locations one uses $\delta_{jk} = \delta_j - \tau_k$. For example, for the first item ($\hat{\delta}_1 = 0.210$) the first transition location ($\hat{\delta}_{11}$) is located at $(0.210 - (-1.545)) = 1.755$, the second transition location ($\hat{\delta}_{12}$) is located at $(0.210 - 0.399) = -0.189$, and the third transition location is located at $\hat{\delta}_{13} = (0.210 - 1.146) = -0.936$. The corresponding ORF plot is presented in Figure D.1.

Table D.2. Abridged output file for PARSCALE GRS model calibration example.

```
                              PARSCALE V4.1
:
              ***   POLYTOMOUS MODEL ITEM ANALYSER   ***

                          *** PHASE 0  ***
:
<echo of input commands>
:

 MODEL SPECIFICATIONS
 =====================

 LOGISTIC - PARTIAL CREDIT ITEM RESPONSE MODEL IS SPECIFIED.
             SCALE CONSTANT  1.70 FOR SLOPE PARAMETERS.

 CALIBRATION PARAMETERS
 =====================
 MAXIMUM NUMBER OF EM CYCLES:                   100
:
 ORDER OF INNER EM CYCLES:              CATEGORY - ITEM PARAMETERS
 ESTIMATION ACCELERATOR:               NO (DEFAULT)
 RIDGE METHOD:                         NO (DEFAULT)
:
 INPUT AND RECODED RESPONSE OF FIRST AND SECOND OBSERVATIONS

 OBSERVATION #      1
 GROUP:  1
 ID:     1
 INPUT RESPONSES:     1  1  1  1  1  1
 RECODED RESPONSES:   1  1  1  1  1  1

 OBSERVATION #      2
 GROUP:  1
 ID:     2
 INPUT RESPONSES:     1  1  1  1  1  1
 RECODED RESPONSES:   1  1  1  1  1  1
```

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

```
:
  NORMAL END
:
                            *** PHASE 1  ***
:
 SUMMARY ITEM STATISTICS
 ======================
 BLOCK NO.:   1     NAME: SBLOCK1
 -------------------------------------------------------------
 ITEM  | TOTAL    NOT    OMIT |        CATEGORIES
       |         PRESENT      |
       |                      |   1      2      3      4
 -------------------------------------------------------------
 0001  |                      |
   FREQ.|  3473      0       0|  1338    492    602   1041
   PERC.|           0.0     0.0|  38.5   14.2   17.3   30.0
:
 0006  |                      |
   FREQ.|  3473      0       0|   744    451    774   1504
   PERC.|           0.0     0.0|  21.4   13.0   22.3   43.3
       |                      |
 -------------------------------------------------------------
 CUMMUL.|                     |
   FREQ.|                      |  8718   2608   3131   6381
   PERC.|                      |  41.8   12.5   15.0   30.6
 -------------------------------------------------------------


 -----------------------------------------------------------------------
  BLOCK    | RESPONSE   TOTAL SCORE | PEARSON  & | INITIAL     INITIAL
   ITEM    |   MEAN        MEAN     | POLYSERIAL |  SLOPE      LOCATION
           |   S.D.*       S.D.*    | CORRELATION|
 -----------------------------------------------------------------------
  SBLOCK1  |                        |            |
   1 0001  |   2.388      14.066    |   0.606    |   0.929     -0.028
           |   1.268*      4.366*   |   0.681    |
:
   6 0006  |   2.875      14.066    |   0.555    |   0.799     -0.245
           |   1.185*      4.366*   |   0.624    |
 -----------------------------------------------------------------------
  CATEGORY |  SCORING   |   MEAN    |   S.D.   | PARAMETER
     1     |   1.000    |  11.588   |  3.456   |
     2     |   2.000    |  13.247   |  3.302   |  -0.180
     3     |   3.000    |  14.986   |  3.398   |   0.192
     4     |   4.000    |  17.336   |  4.006   |  -0.012
 -----------------------------------------------------------------------
:
  NORMAL END
:
                            *** PHASE 2  ***
:
 [E-M CYCLES]   PARTIAL CREDIT MODEL

 STEP AND ITEM PARAMETERS AFTER CYCLE    0

   LARGEST CHANGE=  0.000
   -2 LOG LIKELIHOOD =      51969.544

 STEP AND ITEM PARAMETERS AFTER CYCLE    1

   LARGEST CHANGE=  0.415 (  0.732->  0.317) at Slope    of Item:  4 0004
   -2 LOG LIKELIHOOD =      49902.742
:
 STEP AND ITEM PARAMETERS AFTER CYCLE   18

   LARGEST CHANGE=  0.008 ( -1.531-> -1.539) at Category of Block:  1 SBLOCK1

 [NEWTON CYCLES]   PARTIAL CREDIT MODEL
```

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

```
 STEP AND ITEM PARAMETERS AFTER CYCLE     0

   LARGEST CHANGE=  0.000
   -2 LOG LIKELIHOOD =      48524.502

 STEP AND ITEM PARAMETERS AFTER CYCLE     1

   LARGEST CHANGE=  0.007 ( -1.539-> -1.545) at Category of Block:  1 SBLOCK1

   ITEM BLOCK   1  SBLOCK1

   SCORING FUNCTION    :    1.000     2.000     3.000     4.000
   STEP PARAMTER       :    0.000    -1.545     0.399     1.146
   S.E.                :    0.000     0.042     0.052     0.044
+------+-----+---------+---------+---------+---------+---------+---------+
| ITEM |BLOCK|  SLOPE  |  S.E.   |LOCATION |  S.E.   |GUESSING |  S.E.   |
+======+=====+=========+=========+=========+=========+=========+=========+
| 0001 |  1  |  0.270  |  0.010  |  0.210  |  0.038  |  0.000  |  0.000  |
:
| 0006 |  1  |  0.214  |  0.010  | -0.781  |  0.057  |  0.000  |  0.000  |
+------+-----+---------+---------+---------+---------+---------+---------+

     SUMMARY STATISTICS OF PARAMETER ESTIMATES
       +----------+---------+---------+----+
       |PARAMETER |  MEAN   | STN DEV | N  |
       +==========+=========+=========+====+
       |SLOPE     |   0.315 |   0.111 |  6 |
       |LOG(SLOPE)|  -1.210 |   0.367 |  6 |
       |THRESHOLD |   0.146 |   0.673 |  6 |
       |GUESSING  |   0.000 |   0.000 |  0 |
       +----------+---------+---------+----+
:
  NORMAL END
:
                              *** PHASE 3  ***
:
 NO SCORING OF SUBJECTS
:
  NORMAL END
```

The final phase contains information about person location estimation. For this example, we did not request

estimates of the person locations. Therefore, this phase's output states no person locations were estimated (i.e.,

NO SCORING OF SUBJECTS).

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.
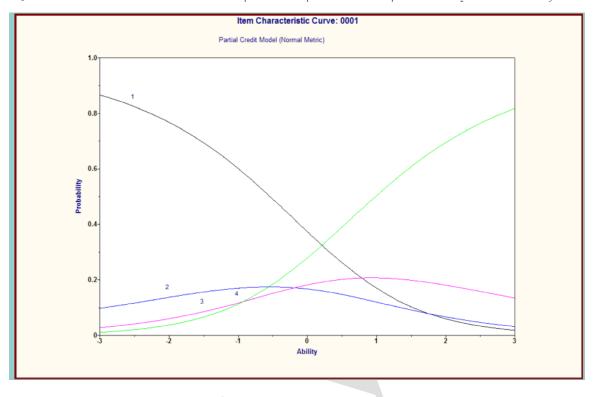
Figure D.1. GRS model ORFs for item 1 ($\hat{\alpha}_1 = 0.270$, $\hat{\delta}_1 = 0.210$ with $\hat{\tau}_1 = -1.545$, $\hat{\tau}_2 = 0.399$, and $\hat{\tau}_3 = 1.146$).

Example: Generalized Partial Credit Model using PARSCALE

To demonstrate a GPC model calibration we revisit the Alike Reasoning data. The input command file is shown in Table D.3. For the GPC model one needs to estimate the thresholds for each item. As a result, each item is a block and the instrument consists multiple blocks. To indicate multiple blocks one includes the NBLOCK subcommand on the TEST command line and then the REPEAT subcommand on the CALIB line. Because PARSCALE's requires that the data consist of coded values beginning with 1 we need to recode our data. That is, our original data defines incorrect as worth 0 points, partially correct as 1 point, and correct as 2 points. Therefore, we need to recode our data so that an incorrect answer garners 1 point, a partially correct response receives 2 points, and a correct answer is worth 3 points. We accomplish this recoding by using the ORIGINAL and MODIFIED keywords on the BLOCK1 line. On the CALIB line we instruct PARSCALE to perform a partial credit calibration (i.e., GPC) using the LOGISTIC metric and fifteen quadrature points. Furthermore, we specify the maximum number of EM cycles to be 100, the maximum number of Newton iterations to be 20, and a convergence criterion of 0.01. The corresponding output is presented in Table D.4.

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

Table D.3. Command file for PARSCALE GPC model calibration example.

```
PARSCALE CALIBRATION-GPC
DATA ARE ALIKE REASONING INSTRUMENT
>FILES   DFNAME='alikeno9.dat';
>INPUT   NIDCHAR=15, NTOTAL=8,NTEST=1, LENGTH=8;
(15A1,T1,8(A1,1X))
>TEST    TNAME='SCALE1', ITEM=(1(1)8), NBLOCK=8;
>BLOCK1  REPEAT=8,NITEMS=1, NCAT=3, ORIGINAL=(0,1,2),MODIFIED=(1,2,3);
>CALIB   PARTIAL, LOGISTIC, NQPTS=15, CYCLE=(100,1,1,1,1), NEWTON=20, CRIT=0.01;
```

As mentioned above the output from PARSCALE is presented in four phases. Phase 0 contains the echo of the input commands followed by the presentation of the default and user-specified program parameters and an echo of the first two cases' response vectors. These echoed response vectors should be inspected to verify that the data are read correctly.

Phase 1 contains traditional item statistics and frequency distributions. As an example, we see from the SUMMARY ITEM STATISTICS section that 2942 individuals responded to item 1 with no omissions. Of these 2942, 201 (6.8%) incorrectly responded to the item, 351 (11.9%) received partial credit, and 2390 (81.2%) responded correctly. In addition, we have the PEARSON and polyserial correlations as measures of the item's discriminatory capacity. For instance, based on the percent correct (or the RESPONSE MEAN of 2.744 on a 3-point response scale) we conclude that item 1 is a relatively easy item that had, based on its PEARSON value, a "moderate" capacity (0.536) to discriminate among different observed scores.

The IRT results (phase 2) begins with the iteration history followed by the item parameter estimates. The SLOPE column contains the estimates of the discrimination parameters. For item 1 we see that its discrimination estimate ($\hat{\alpha}_1$) is 0.766. The physical location of the transition locations on the continuum are obtained in a manner similar to that of the GRS model (i.e., $\delta_{jk} = \delta_j - \tau_k$). The STEP PARAMTER [sic] line contains the item's corresponding offset values (i.e., τ_hs) and the column labeled LOCATION contain the items' $\hat{\delta}_j$s. For example, for the first item the first transition location ($\hat{\delta}_{11}$) is located at (-1.630 - (-0.088)) = -1.542 and the second transition location ($\hat{\delta}_{12}$) is located at (-1.630 - 0.088) = -1.718. As is true with the MULTILOG results, the transition locations are reversed. This item's ORFs are shown in Figure D.2. For item 8 one would use offset values from the ITEM BLOCK 8 BLOCK section. Therefore, for item 8 we have $\hat{\delta}_{81} = (1.829 - (-0.048)) = 1.877$

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

and $\hat{\delta}_{82} = (1.829 - 0.048) = 1.781$. Comparing the PARSCALE results with those of MULTILOG shows Pearson

product-moment correlations of 0.9943 and 0.9862 for the first and second transition locations, respectively.

Table D.4. Abridged output file for PARSCALE GPC model calibration example.

```
:
                      ***   POLYTOMOUS MODEL ITEM ANALYSER   ***

                            *** PHASE 0   ***
:
<echo of input commands>
:

MODEL SPECIFICATIONS
 =====================

 LOGISTIC - PARTIAL CREDIT ITEM RESPONSE MODEL IS SPECIFIED.
                SCALE CONSTANT  1.70 FOR SLOPE PARAMETERS.

 CALIBRATION PARAMETERS
 =====================
 MAXIMUM NUMBER OF EM CYCLES:                 100
:
 ORDER OF INNER EM CYCLES:                 CATEGORY - ITEM PARAMETERS
 ESTIMATION ACCELERATOR:                   NO (DEFAULT)
 RIDGE METHOD:                             NO (DEFAULT)
:
 INPUT AND RECODED RESPONSE OF FIRST AND SECOND OBSERVATIONS

 OBSERVATION #       1
 GROUP:  1
 ID: 1 1 2 0 0 0 0 0
 INPUT RESPONSES:     1  1  2  0  0  0  0  0
 RECODED RESPONSES:   2  2  3  1  1  1  1  1

 OBSERVATION #       2
 GROUP:  1
 ID: 2 2 0 0 0 0 0 0
 INPUT RESPONSES:     2  2  0  0  0  0  0  0
 RECODED RESPONSES:   3  3  1  1  1  1  1  1
:
   NORMAL END
                            *** PHASE 1   ***
:
```

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

```
SUMMARY ITEM STATISTICS
 =======================

 BLOCK NO.:   1    NAME: SBLOCK1
 -------------------------------------------------------
 ITEM  | TOTAL    NOT     OMIT |    CATEGORIES
       |         PRESENT       |
       |                       |   1      2      3
 -------------------------------------------------------
 0001  |                       |
  FREQ.|  2942       0       0 |  201    351   2390
  PERC.|           0.0     0.0 |  6.8   11.9   81.2
:
 0008  |                       |
  FREQ.|  2942       0       0 | 2043    616    283
  PERC.|           0.0     0.0 | 69.4   20.9    9.6
       |                       |
 -------------------------------------------------------
 CUMMUL.|                      |
  FREQ.|                       | 9071   5664   8801
  PERC.|                       | 38.5   24.1   37.4
 -------------------------------------------------------
 -----------------------------------------------------------------------
 BLOCK  | RESPONSE   TOTAL SCORE | PEARSON  &  | INITIAL   INITIAL
  ITEM  |  MEAN        MEAN      | POLYSERIAL  | SLOPE     LOCATION
        |  S.D.*       S.D.*     | CORRELATION |
 -----------------------------------------------------------------------
 BLOCK  |                        |             |
  1 0001|   2.744     15.908     |   0.536     |  1.184     0.000
        |   0.572*     3.359*    |   0.764     |
 -----------------------------------------------------------------------
 CATEGORY | SCORING  |  MEAN    |   S.D.   | PARAMETER
    1     |  1.000   | 10.826   |  2.813   |
    2     |  2.000   | 13.103   |  2.813   |   0.263
    3     |  3.000   | 16.748   |  2.833   |  -0.263
:
 -----------------------------------------------------------------------
 BLOCK  |                        |             |
  8 0008|   1.402     15.908     |   0.475     |  0.750     0.000
        |   0.658*     3.359*    |   0.600     |
 -----------------------------------------------------------------------
 CATEGORY | SCORING  |  MEAN    |   S.D.   | PARAMETER
    1     |  1.000   | 14.839   |  3.044   |
    2     |  2.000   | 17.982   |  2.684   |   0.245
    3     |  3.000   | 19.110   |  2.619   |  -0.245
 -----------------------------------------------------------------------
:
  NORMAL END

                        *** PHASE 2  ***
:
 [E-M CYCLES]   PARTIAL CREDIT MODEL

 STEP AND ITEM PARAMETERS AFTER CYCLE    0

   LARGEST CHANGE=  0.000
   -2 LOG LIKELIHOOD =      40635.580

 STEP AND ITEM PARAMETERS AFTER CYCLE    1

   LARGEST CHANGE=  1.311 (  1.375->  0.064) at Category of Block:  1 BLOCK
   -2 LOG LIKELIHOOD =      42904.124
:

 STEP AND ITEM PARAMETERS AFTER CYCLE   20

   LARGEST CHANGE=  0.009 ( -1.554-> -1.563) at Category of Block:  5 BLOCK
```

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

```
 [NEWTON CYCLES]   PARTIAL CREDIT MODEL

 STEP AND ITEM PARAMETERS AFTER CYCLE     0

   LARGEST CHANGE=  0.000
   -2 LOG LIKELIHOOD =       39293.383

 STEP AND ITEM PARAMETERS AFTER CYCLE     1
   LARGEST CHANGE=  0.007 ( -1.563-> -1.571) at Category of Block:  5 BLOCK

   ITEM BLOCK   1  BLOCK
   SCORING FUNCTION    :     1.000     2.000     3.000
   STEP PARAMTER       :     0.000    -0.088     0.088
   S.E.                :     0.000     0.073     0.047
:
   ITEM BLOCK   8  BLOCK
   SCORING FUNCTION    :     1.000     2.000     3.000
   STEP PARAMTER       :     0.000    -0.048     0.048
   S.E.                :     0.000     0.072     0.111
```

ITEM	BLOCK	SLOPE	S.E.	LOCATION	S.E.	GUESSING	S.E.
0001	1	0.766	0.038	-1.630	0.059	0.000	0.000
0002	2	0.625	0.027	-0.933	0.043	0.000	0.000
:							
0008	8	0.389	0.028	1.829	0.116	0.000	0.000

```
     SUMMARY STATISTICS OF PARAMETER ESTIMATES
```

PARAMETER	MEAN	STN DEV	N
SLOPE	0.509	0.135	8
LOG(SLOPE)	-0.704	0.253	8
THRESHOLD	0.207	1.271	8
GUESSING	0.000	0.000	0

```
:
  NORMAL END
:
                            *** PHASE 3   ***
:
 NO SCORING OF SUBJECTS
:
  NORMAL END
```

Phase 4 contains information about the person location estimates. Because for this example we did not

request estimates of the person locations this phase's output states no person location were estimated (i.e., NO

SCORING OF SUBJECTS).

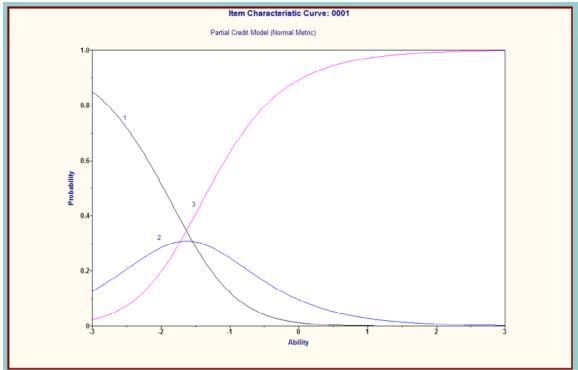de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.

Figure D.2. GPC model ORFs for item 1 ($\hat{\alpha}_1 = 0.766$, $\hat{\delta}_{11} = -1.542$ and $\hat{\delta}_{12} = -1.718$).

de Ayala, R.J. (2009). *The Theory and Application of Item Response Theory*, New York: Guilford Publishing.