Example: Application of the BS Model to a Science Test, MMLE, MULTILOG

For this example we explore whether the proportion of “don’t know” individuals is constant and equal to $1/m_j$. To investigate this, the data used in the MC model example are reanalyzed using the BS model and the fit of the models is compared. Because the data involve four 4-option items using (primarily) a multiple-choice response format, $1/m_j$ = 0.25. To perform the BS model calibration we modify the command file for the MC model analysis (MULTILOG_MCMcalibrationEx.pdf’s Table 1). Specifically, the modification is the insertion of the command line `>FIX ITEMS=(1,2,3,4), DK=(1,2,3), VALUE=0.0;` to impose the constraint that the “don’t know” proportions (i.e., DK) are constrained to be $1/m_j$ = 0.25 for all options across the four items. This line is inserted before the `>END;` line. Table 1 contains the corresponding output.

Table 1. Abridged output from BS model calibration example.

```
     :
   MAXIMUM NUMBER OF EM CYCLES PERMITTED: 500
   NUMBER OF PARAMETER-SEGMENTS USED IS:    4
   NUMBER OF FREE PARAMETERS IS:   32
   MAXIMUM NUMBER OF M-STEP ITERATIONS IS  25 TIMES
     :
  FINISHED CYCLE 337
  MAXIMUM INTERCYCLE PARAMETER CHANGE=   0.00051 P(  29)
     :
  ITEM SUMMARY
     :
  ITEM   1:        5 NOMINAL CATEGORIES,  3 HIGH
   CATEGORY(K): 1      2      3      4      5
     A(K)     -0.60   0.82   1.34  -0.73  -0.84
     C(K)      2.44   0.86   1.33  -4.94   0.31
     D(K)             0.25   0.25   0.25   0.25

                  CONTRAST-COEFFICIENTS (STANDARD ERRORS)
   FOR:             A                    C                    D
   CONTRAST P(#)  COEFF.[ DEV.]  P(#)  COEFF.[ DEV.]  P(#)  COEFF.[ DEV.]
       1       1   1.42 (0.48)      5  -1.58 (1.27)     33   0.00 (0.00)
       2       2   1.94 (0.38)      6  -1.12 (1.03)     34   0.00 (0.00)
       3       3  -0.14 (****)      7  -7.39 (****)     35   0.00 (0.00)
       4       4  -0.24 (0.92)      8  -2.14 (1.72)

  @THETA:      INFORMATION:   (Theta values increase in steps of 0.2)
  -3.0 - -1.6  0.004  0.004  0.004  0.004  0.005  0.005  0.006  0.009
  -1.4 -  0.0  0.012  0.018  0.028  0.045  0.073  0.115  0.172  0.241
   0.2 -  1.6  0.313  0.373  0.407  0.409  0.382  0.336  0.282  0.229
   1.8 -  3.0  0.182  0.143  0.112  0.088  0.070  0.056  0.046
```

```
 OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
 CATEGORY(K):  1      2      3      4      5
 OBS. FREQ.       0    456    723    245    375
 OBS. PROP.  0.0000 0.2535 0.4019 0.1362 0.2084
 EXP. PROP.  0.5437 0.2542 0.4011 0.1363 0.2084
   :
ITEM   4:       5 NOMINAL CATEGORIES,  4 HIGH
 CATEGORY(K): 1      2      3      4      5
   A(K)       2.35 -12.61   3.22   2.77   4.27
   C(K)       5.12 -17.46   4.34   4.53   3.47
   D(K)              0.25   0.25   0.25   0.25

                 CONTRAST-COEFFICIENTS (STANDARD ERRORS)
 FOR:              A                    C                    D
 CONTRAST P(#)  COEFF.[ DEV.]  P(#)  COEFF.[ DEV.]  P(#)  COEFF.[ DEV.]
     1      25 -14.95 (****)     29 -22.58 (****)     45   0.00 (0.00)
     2      26   0.87 (0.25)     30  -0.78 (0.20)     46   0.00 (0.00)
     3      27   0.42 (0.23)     31  -0.59 (0.17)     47   0.00 (0.00)
     4      28   1.93 (0.30)     32  -1.65 (0.33)

@THETA:      INFORMATION:   (Theta values increase in steps of 0.2)
-3.0 - -1.6  0.000  0.000  0.000  0.000  0.008  0.161  3.209 34.849
-1.4 -  0.0  8.934  0.097  0.017  0.017  0.021  0.029  0.043  0.066
 0.2 -  1.6  0.102  0.151  0.210  0.274  0.334  0.382  0.412  0.420
 1.8 -  3.0  0.407  0.378  0.339  0.295  0.250  0.208  0.170

 OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
 CATEGORY(K):  1      2      3      4      5
 OBS. FREQ.       0    288    507    551    453
 OBS. PROP.  0.0000 0.1601 0.2818 0.3063 0.2518
 EXP. PROP.  0.3921 0.1609 0.2823 0.3062 0.2506
   :
MARGINAL RELIABILITY:    0.4962
   :
NEGATIVE TWICE THE LOGLIKELIHOOD=       277.7
(CHI-SQUARE FOR SEVERAL TIMES MORE EXAMINEES THAN CELLS)
```

In contrast to the 253 cycles required for convergence with the MC model calibration of these data, the BS model required 337 to converge to a solution. The output format is identical to that seen with the MC model. Because of the imposed constraint, the line labeled D(K) (correctly) shows that each of the $\hat{\phi}_{jk}$s equals 0.25 and the column labeled D contains the constraint's value of 0.0. For item 1 the item parameter estimates are $\underline{\hat{\boldsymbol{\alpha}}}$ = (–0.60, 0.82, 1.34, –0.73, –0.84), $\underline{\hat{\boldsymbol{\gamma}}}$ = (2.44, 0.86, 1.33, –4.94, 0.31), and $\underline{\hat{\phi}}$ = (0.25, 0.25, 0.25, 0.25) with $\sum \hat{\alpha}_{jk}$ = 0, $\sum \hat{\gamma}_{jk}$ = 0, and $\sum \hat{\phi}_{jk}$ = 1.

Item 1's ORFs are presented in the left panel of Figure 1. As is the case with the NR model and the NR–PC mixed model calibrations, this item is primarily functioning in a dichotomous fashion. It may also be seen that if individuals below –2 did not select option 5, then they were equally likely to select among the remaining options. The item's information function (right panel) shows that this item is not providing very much information, and substantially less than that provided by this item using the MC model (cf. MULTILOG_MCMcalibrationEx.pdf). It should be noted that for each calibration the graphing routine imposes a common axis scale across items to facilitate comparisons. Therefore, when we compare figures across calibrations (e.g., Figure 1 with MULTILOG_MCMcalibrationEx.pdf's Figure 1) we need to attend to the scaling differences of the ordinate. In the case of the BS model, item 4's information function has a maximum value of approximately 34.8 and so the maximum for the ordinate scale is set on this item's basis. Although this could potentially invite misleading interpretations of the amount of information an item provides, comparing the values in the `@THETA:INFORMATION` section of the outputs (Table 1and MULTILOG_MCMcalibrationEx.pdf's Table 2) substantiates the above conclusion. The ORFs for all four items are presented in Figure 2. As can be seen, for item 3 there is no single option that is preferred over the other options by individuals of low θ.

The overall BS model fit is –2ln*L* = 277.7 with a BIC = 517.5396 and 32 estimated parameters (i.e., for each item there are 4 α_{jk} s and 4 γ_{jk} s, or 32 parameters for the four items). This deviance statistic is not significantly different from that for the MC model, although the BS model's BIC is about 12% smaller; we assume that the Full model holds for the data. Therefore, the BS model may be preferred to the MC model for these data. Table 2 contains the deviance and BIC statistics for the various polytomous model and model combinations that were applied to these data. According to its –2ln*L* and BIC values, the NR–PC mixed model calibration, where item 2's options 1 and 2 were combined, provides the best fit to these data.

Figure 1. BS model ORFs and item information function for item 1.

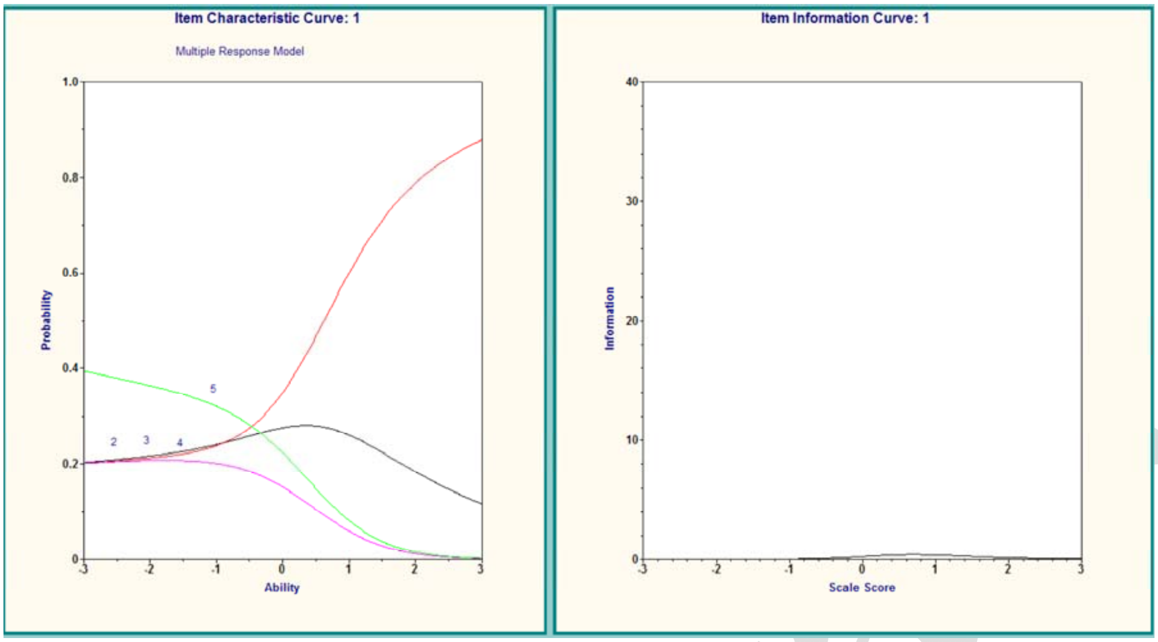

Figure 2. ORFs for all four items, BS model calibration.

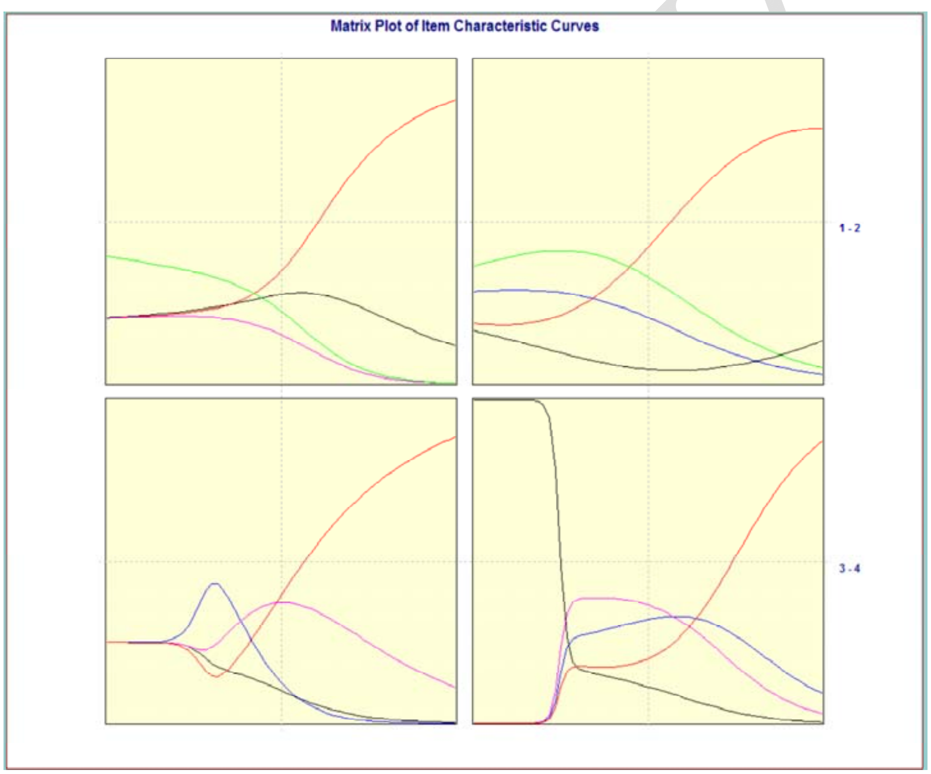

Table 2. Deviance and BIC statistics for polytomous models used in calibrating the Science test.

Model	-2ln*L*	Number of estimated parameters	BIC
NR	288.4	24	468.2797
NR-PC	320.5	22	485.3897
NR-PC[a]	219.6	20	369.4997
NR-GR	322.2	22	487.0897
MC	259.3	44	589.0794
BS	277.7	32	517.5396

[a]Item 2 collapsed from 4 to 3 response categories.