



Rater-effect model: multi-record input format with varying numbers of raters per examinee

This example illustrates the parameter estimation for multiple raters. The analysis is based on data in the file **exampl07.dat**. The first few lines of the data are shown below.

```
00001 12 11 32 32
00001 22 21 42 42
00002 12 11 32 32
00002 22 22 43 42
00003 12 12 31 31
00003 23 22 43 41
00004 12 12 33 32
00004 22 22 42 42
00005 11 11 31 31
00005 22 21 41 42
```

The data contain the rating on four items administered to each examinee by four raters. The first 5 columns of each line of data contain the examinee ID. After two blank columns, the rater ID is given, directly followed by the rating on the first item. Similar combinations of rater ID and rating for the other three items follow. As can be seen from the data above, the first line of data is associated with examinee 00001 and contains the ratings for raters 1 and 3. The second line of data, associated with the same examinee, contains the ratings for raters 2 and 4.

The data are read using the format statement

```
(5A1, 4(2X, 2A1))
```

where “5A1” is the format of the examinee ID, and “2X, 2A1” the format for reading of one rater ID/rating combination. The latter is repeated four times, using the notation “4()”. Note that, since the data for each examinee are given on two lines, R-INOPT=2 could have been specified on the INPUT command and the format statement changed to

```
(5A1,4(2X,2A1),/T6,4(2X,2A1)).
```

The MRATER keyword on the INPUT command requests a Rater's-Effect analysis, and indicates the number of raters. The MRATER command provides necessary information about the four raters.

The estimated parameters and scores are saved to external output files using the SAVE option on the FILES command and the PARM and SCORE keywords on the SAVE command.

The command file for a partial credit model based on these data is shown below.

```

EXAMPL07.PSL - ARTIFICIAL EXAMPLE: MONTE CARLO DATA
      GENERALIZED PARTIAL CREDIT MODEL: RATERS' EFFECT MODEL [NESTED DESIGN]
>FILES  DFNAME='EXAMPL07.DAT', SAVE;
>SAVE   PARM='EXAMPL07.PAR', SCORE='EXAMPL07.SCO';
>INPUT  NIDCHAR=5, NTOT=4, LENGTH=4, NTEST=1, NFMT=1, MRATER=4;
      (5A1,4(2X,2A1))
>TEST   TNAME=RATERN, ITEM=(1,2,3,4), NBLOCK=4;
>BLOCK  REPEAT=4, NITEMS=1, NCAT=3, ORIGINAL=('1','2','3'), MODIFIED=(1,2,3);
>MRATER RNAME=(RaterA,RaterB,RaterC,RaterD), RCODE=('1','2','3','4');
>CAL    LOGISTIC, PARTIAL, NQPT=21, CYCLES=(100,1,1,1,1,1), NEWTON=2,
      CRIT=0.05, DIAG=0, ITEMFIT=10, SCALE=1.7;
>SCORE  EAP;

```

Phase 0 output

In addition to the standard Phase 0 output discussed elsewhere, information on the raters' names, codes, and the weight assigned to each is echoed to the output file.

The MRATER command used here only assigns names and codes to the rater. By default, the RATER keyword, not included in the MRATER command shown here, assumes the value (1,1,1,1). The arguments of this keyword are the raters' weights. For the Raters-effect model, the ability score for each respondent is computed for each subtest (or subscale) and each rater separately. A total score of each respondent for each subtest (or subscale) is computed by summing those scores over items within each subtest and all raters who have rated the respondent. The rater weights of this keyword are used to compute the weighted subtest or subscale score for each respondent. Since the number of raters who rated each respondent's responses varies, the weights are normalized (divided by their sum) for each respondent.

```
>MRATER RNAME=(RaterA,RaterB,RaterC,RaterD), RCODE=('1','2','3','4');
```

```
MULTIPLE GROUP MODEL: RATER's EFFECT MODEL
```

```
RATER's NAME, CODE, AND WEIGHT
```

```
=====
```

1	RaterA	1	1.00
2	RaterB	2	1.00
3	RaterC	3	1.00
4	RaterD	4	1.00

Also included in the Phase 0 output is a listing of the first two observations, showing the input and recoded responses. The raters responsible for each rating are also listed. This information is provided so that the user can check that the data are read in correctly. If not, the variable format statement (or the data) should be corrected.

```

INPUT AND RECODED RESPONSE OF FIRST AND SECOND OBSERVATIONS
OBSERVATION #      1

```

```

GROUP: 1
ID: 00001
INPUT RESPONSES: 2 1 2 2
RECODED RESPONSES: 2 1 2 2
RECODED RATERS : 1 1 3 3

```

```

OBSERVATION # 2
GROUP: 1
ID: 00001
INPUT RESPONSES: 2 1 2 2
RECODED RESPONSES: 2 1 2 2
RECODED RATERS : 2 2 4 4

```

The Phase 0 output also reports that 2000 lines of data were read from the data file, and indicates that these 2000 observations are associated with 1000 examinees.

```

[MAIN TEST: RATERN ]
2000 OBSERVATIONS READ FROM FILE: EXAMPL07.DAT
2000 OBSERVATIONS WRITTEN TO FILE: exampl07.MFL

```

```

MULTIPLE RATERS DATA
1000 CASES READ FROM FILE: EXAMPL07.DAT

```

Phase 1 Output

The Phase 1 output file contains no additional information in this type of analysis. As usual, frequencies and percentages for items nested within blocks are reported here. Information for the first block/item is shown below.

```

SUMMARY ITEM STATISTICS
=====
BLOCK NO.: 1 NAME: BLOCK
-----
ITEM | TOTAL NOT OMIT | CATEGORIES
      | | PRESENT | | 1 2 3
-----
0001 | | | |
FREQ. | 2000 0 0 | 235 1145 620
PERC. | | 0.0 0.0 | 11.8 57.2 31.0
      | | | |
-----

```

Phase 2 Output

The Phase 2 output file shows the standard output for category parameters and item parameters at convergence. This is followed by rater parameters and their associated standard errors as shown below.

```

ITEM BLOCK 1 BLOCK
SCORING FUNCTION : 1.000 2.000 3.000
STEP PARAMTER : 0.000 1.243 -1.243
S.E. : 0.000 0.056 0.041

```

```

ITEM BLOCK 2 BLOCK
SCORING FUNCTION : 1.000 2.000 3.000
STEP PARAMTER : 0.000 1.183 -1.183
S.E. : 0.000 0.037 0.049

```

```

ITEM BLOCK 3 BLOCK
SCORING FUNCTION : 1.000 2.000 3.000

STEP PARAMTER : 0.000 1.298 -1.298

S.E. : 0.000 0.079 0.064

```

```

ITEM BLOCK 4 BLOCK
SCORING FUNCTION : 1.000 2.000 3.000
STEP PARAMTER : 0.000 1.273 -1.273
S.E. : 0.000 0.062 0.077

```

ITEM	BLOCK	SLOPE	S.E.	LOCATION	S.E.	GUESSING	S.E.
0001*	1	0.814	0.041	-0.515	0.039	0.000	0.000
0002*	2	0.935	0.047	0.410	0.037	0.000	0.000
0003*	3	0.491	0.027	-0.502	0.051	0.000	0.000
0004*	4	0.505	0.028	0.508	0.050	0.000	0.000

RATER'S EFFECT PARAMETER

RATER'S NAME	PARAMETER	S.E.
RaterA	-0.008	0.030
RaterB	-0.006	0.030
RaterC	0.069	0.044
RaterD	-0.055	0.044

NOTE: RATED ITEMS ARE MARKED BY '*'

From the output above, we see a marked difference between the raters, in particular between RaterC and RaterD. The raters differ appreciably in severity.