

# IRTPRO™

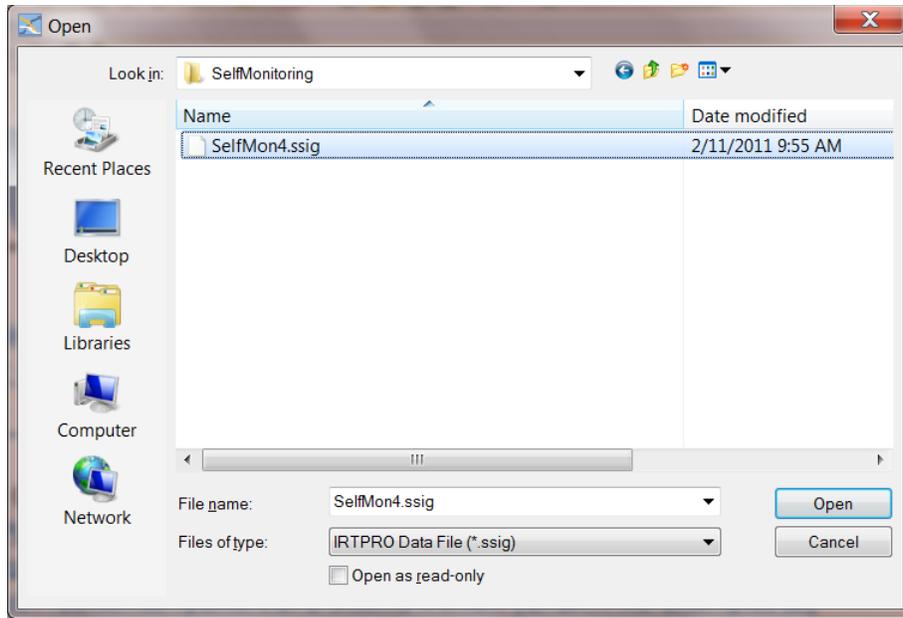
## Unidimensional analysis of four "self monitoring" items

Thissen & Steinberg (2009) describe IRT model fitting for the responses of 393 undergraduate students to four items of the Self-Monitoring Scale (SMS; Snyder, 1974; modified by Snyder and Gangestad, 1986). The data used is from the 1988 "Self-Monitoring Scale (CAPS-SELFMON, SELF\_MONIT and SELF\_MONIT\_PAPER module)", hdl:1902.29/CAPS-SELFMON Odum Institute Dataverse. They consider a subset of the data for the following four items:

- SelfMon8: I have considered being an entertainer. (T)
- SelfMon13: I have never been good at games like charades or improvisational acting. (F)
- SelfMon18: I would probably make a good actor. (T)
- SelfMon20: In different situations and with different people, I often act like a very different person. (T)

The high self-monitoring response to each of the items above (T or F in parentheses after each item) is coded 1 and the other response is coded 0.

To begin, we use the **Open** file dialog under the **File** menu of IRTPRO and navigate to the **C:\IRTPRO Examples\By Dataset\SelfMonitoring** folder. However, here we change the **Files of type**: selection from its default **IRTPRO Command File (\*.irtpro)** to **IRTPRO Data File (\*.ssig)** in the **Open** file dialog, and open the file **SelfMon4.ssig**.

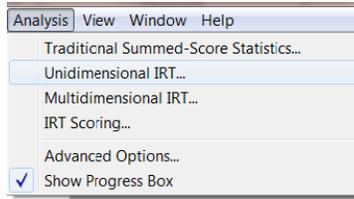


The first 15 cases of these data are shown below.

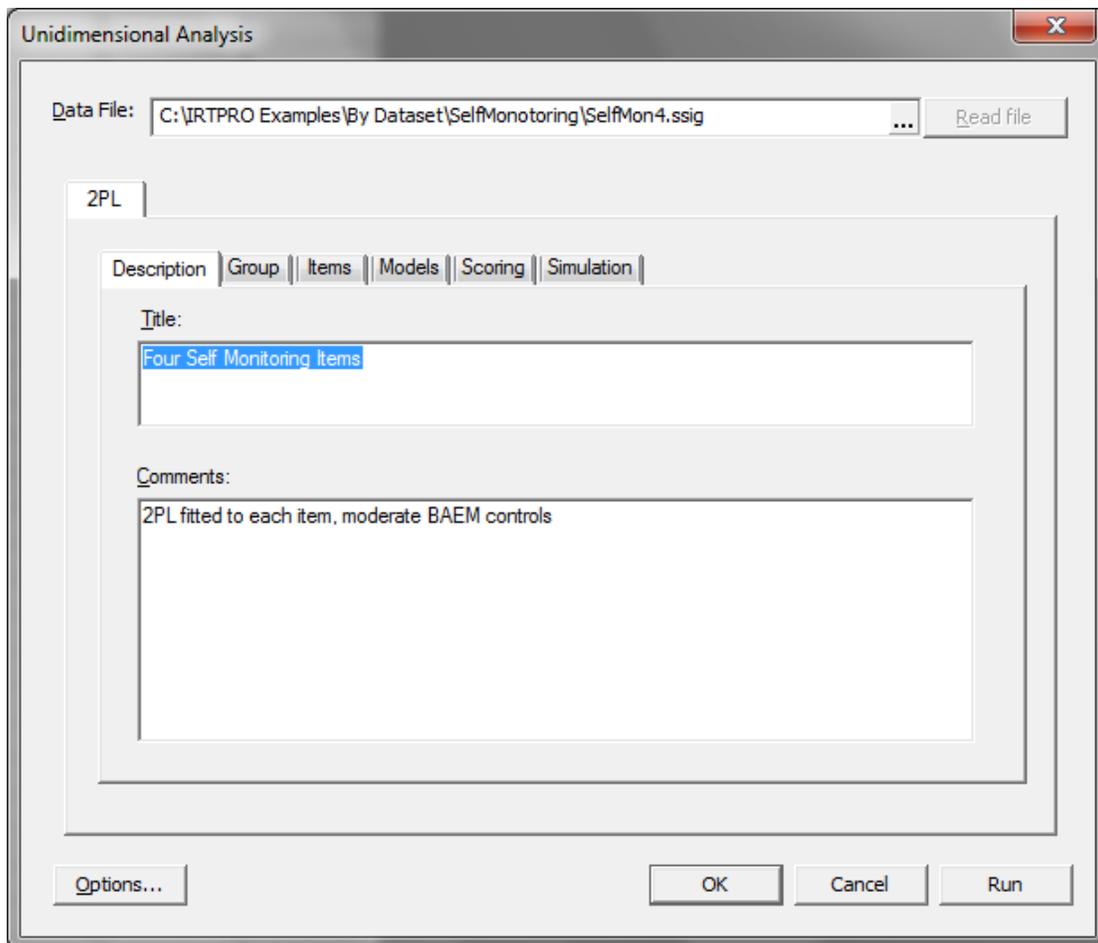
The screenshot shows the IRTPRO software interface. The title bar reads 'IRTPRO - [SelfMon4.ssig]'. The menu bar includes 'File', 'Edit', 'Data', 'Manipulate', 'Graphics', and 'Analysis'. The toolbar contains various icons for file operations. The main window displays a data table with 15 rows and 5 columns. The columns are labeled 'SelfMon8', 'SelfMon13', 'SelfMon18', and 'SelfMon20'. The first row is highlighted in blue.

	SelfMon8	SelfMon13	SelfMon18	SelfMon20
1	0	1	0	1
2	0	0	1	1
3	1	1	1	1
4	1	1	1	1
5	0	0	0	1
6	1	0	0	0
7	0	1	0	1
8	0	0	0	0
9	0	0	0	1
10	0	1	1	1
11	1	1	0	1
12	0	1	0	0
13	0	1	0	1
14	1	0	0	1
15	0	0	0	1

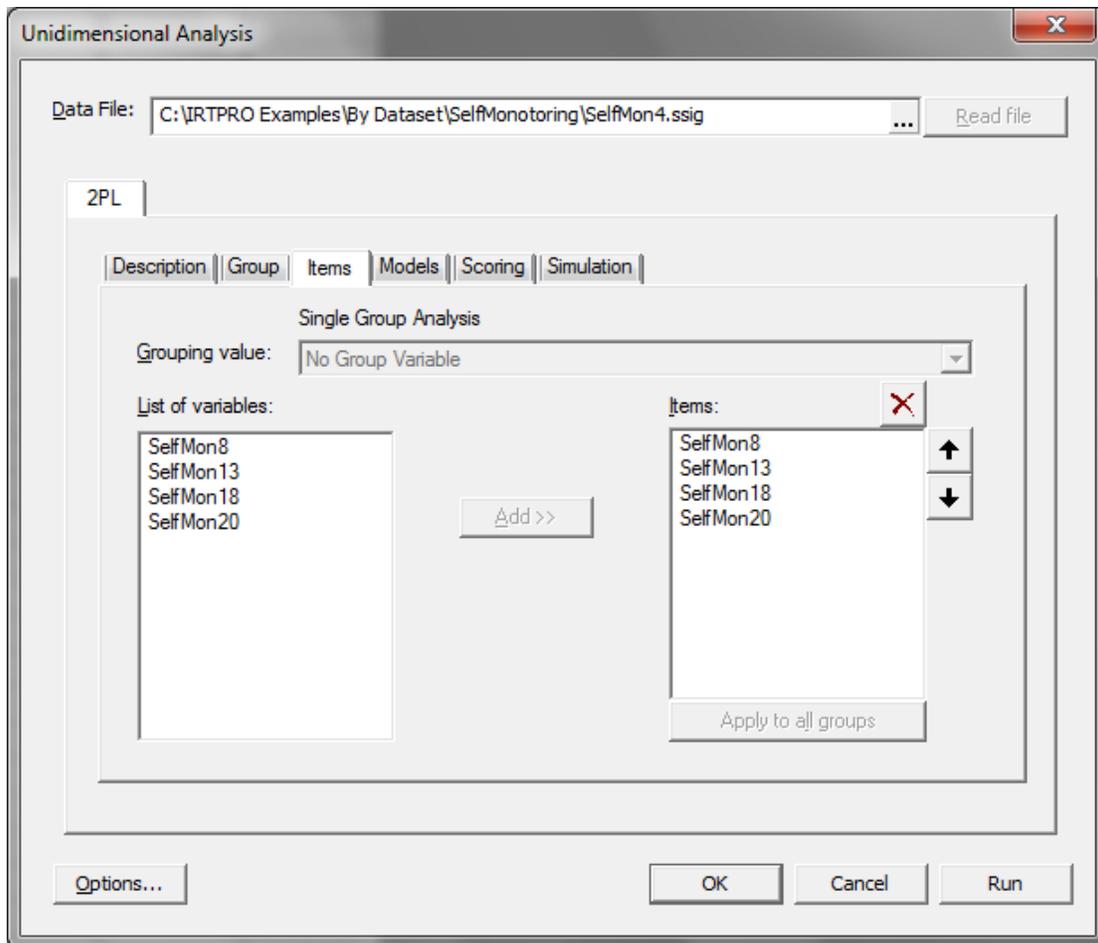
To set up the analyses, select the **Unidimensional IRT...** option from the **Analysis** menu to invoke the unidimensional analysis window.



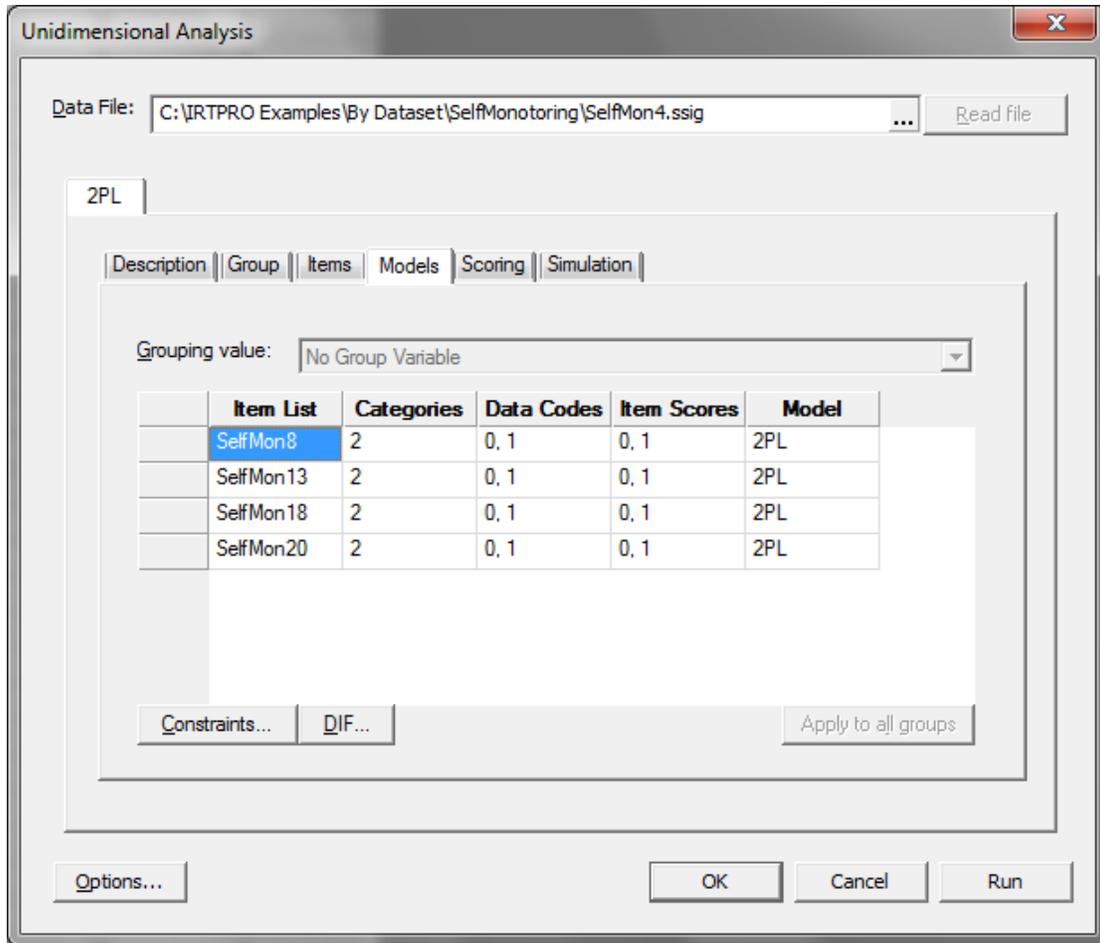
This window has five tabs called **Description**, **Group**, **Items**, **Models** and **Scoring**. Start with the default tab **Description** and provide a title and comments in the appropriate text boxes as shown below. Note that the default name for the current analysis is Test1. As will be shown later, more tests based on the same dataset may be inserted and each of these can be renamed to something that may be more suitable. In the present case, right-click on the Test1 tab and rename it to 2PL.



Since this data contains no grouping variable, the **Group** tab is skipped and we proceed to the **Items** tab, where all four items from the **List of variables** are selected. Then use the **Add** button to list these items under **Items**.



Because the 2PL model is the default for dichotomous items, the entry of information for the analysis is now complete, and clicking on the **Run** button in the lower right of the **Unidimensional Analysis** dialog will produce the results. However, to see more details of how the data will be modeled, selection of the **Models** tab shows the list of items, their data codes, the translation of those codes into response categories, and the model selected:



Click the **Run** button to run the 2PL analysis. Portions of the output file **SelfMon4.2PL-irt.htm** are shown below. We find that the slope parameter for item 2 is estimated to be much lower than for the other three items:

**2PL Model Item Parameter Estimates for Group 1, logit:  $a\theta + c$  or  $a(\theta - b)$**  [\(Back to TOC\)](#)

Item	Label	a	s.e.	c	s.e.	b	s.e.		
1	SelfMon8	2	4.12	1.37	1	-0.09	0.28	0.02	0.07
2	SelfMon13	4	0.12	0.13	3	0.49	0.10	-4.23	4.67
3	SelfMon18	6	2.41	0.46	5	-0.91	0.22	0.38	0.08
4	SelfMon20	8	2.02	0.34	7	0.89	0.19	-0.44	0.09

Note that the values in the table above are not exactly the same as those reported by Thissen & Steinberg (2009) for two reasons: (1) Thissen & Steinberg (2009) tabulate the slopes divided by 1.7, or in the so-called "normal metric", to compare to normal ogive slopes, and (2) even if that is corrected, there are slight numerical differences between the results. This is due to different numerical quadrature used for the Newton-Raphson estimation in R that

Thissen & Steinberg (2009) used vs. the EM estimation used here in IRTPRO.

If we click on the entry **Likelihood-based Values and Goodness of Fit Statistics** in the table of contents for this 2PL fit to these items, we find that the model appears to fit very well:

Statistics based on the full item x item x ... classification			
$\chi^2$	Degrees of freedom	Probability	RMSEA
5.03	7	0.6570	0.00

Statistics based on one- and two-way marginal tables			
M <sub>2</sub>	Degrees of freedom	Probability	RMSEA
3.17	2	0.2063	0.04
Note: M <sub>2</sub> is based on full marginal tables.			
Note: Model-based weight matrix is used.			

The statistics based on the full item classification can rarely be computed in IRT applications, because they require that the sample is sufficiently large to "fill" all of the cells of cross-classification created by listing all response patterns. For four dichotomous items, that is a manageable 16 cells, and so IRTPRO tabulates the observed and expected frequencies (and some other values) for each response pattern as follows:

Response Pattern Observed and Expected Frequencies, Standardized Residuals, EAPs and SDs for Group 1 ([Back to TOC](#))

Item:				Frequencies		Standard		
1	2	3	4	Observed	Expected	Residual	EAP[ $\theta u$ ]	SD[ $\theta u$ ]
0	0	0	0	44	44.72	-0.11	-1.07	0.65
0	0	0	1	25	25.07	-0.01	-0.44	0.49
0	0	1	0	4	3.55	0.24	-0.35	0.47
0	0	1	1	4	6.35	-0.94	0.04	0.43
0	1	0	0	65	64.87	0.02	-1.02	0.63
0	1	0	1	40	39.11	0.15	-0.41	0.48
0	1	1	0	6	5.60	0.17	-0.33	0.46
0	1	1	1	12	10.48	0.48	0.06	0.43
1	0	0	0	7	5.90	0.45	-0.02	0.43
1	0	0	1	25	20.10	1.12	0.36	0.46
1	0	1	0	2	3.88	-0.96	0.44	0.47
1	0	1	1	38	39.42	-0.24	1.02	0.62
1	1	0	0	9	9.67	-0.22	0.00	0.43
1	1	0	1	29	34.44	-0.97	0.39	0.46
1	1	1	0	8	6.72	0.50	0.47	0.48
1	1	1	1	75	73.10	0.25	1.07	0.63

The  $\chi^2$  value (5.03, on 7 d.f.) reported above is obtained by using the likelihood-ratio chi-square to compare the observed and expected frequencies in that table. In this case, those values do not differ by more than is expected by sampling error.

For scales that involve more items, or more response categories, the complete cross-classification is often too large to "fill"; for example, for 6 five-alternative items the full cross-classification would have  $5^6 = 15,625$  cells, which can only be sparsely populated by a few hundred respondents. So in that case, as well as this one, an updated version of the  $M_2$  statistic proposed by Maydeu-Olivares & Joe (2005, 2006) may be used as a proxy for  $\chi^2$ . The  $M_2$  statistic is based on the one- and two-way marginal tables of the complete cross-classification; those subtables are easier to "fill" with reasonable sample sizes.

IRTPRO can also compute a trace line diagnostic statistic for each item, which is a generalization for polytomous responses of the  $S$ - $\chi^2$  item-fit statistic suggested by Orlando & Thissen (2000, 2003). For the 2PL fit to these four items, these statistics are tabulated as shown in the following table.

S-X2 Item Level Diagnostic Statistics

Item	Label	$\chi^2$	d.f.	Probability
1	SelfMon8	0.51	2	0.7748
2	SelfMon13	0.23	2	0.8909
3	SelfMon18	1.86	2	0.3950
4	SelfMon20	0.23	2	0.8902

The summary table (above) indicates that for all four of these items, the trace lines have been fitted sufficiently well that the model-expected proportions responding 0 and 1 match the observed data well.

For each item, the complete tables printed below have one row for each summed score for the "other items" (for each of these four items, the summed score on the "other items" ranges 0 – 3), and within those scores the observed and expected frequencies are tabulated. The entries are printed in blue if the observed frequency exceeds the expected frequency, and in red if too few responses are observed.

**Item 1 S-X2(2) = 0.5 , p = 0.7748 [\(Back\)](#)**

Score	Category 0		Category 1	
	Observed	Expected	Observed	Expected
0	44	45.1	7	5.9
1	94	95.6	36	34.4
2	50	48.5	75	76.5
3	12	10.9	75	76.1

**Item 2 S-X2(2) = 0.2 , p = 0.8909 [\(Back\)](#)**

Score	Category 0		Category 1	
	Observed	Expected	Observed	Expected
0	44	44.5	65	64.5
1	36	35.3	55	55.7
2	31	29.6	49	50.4
3	38	39.6	75	73.4

Item 3  $S-X^2(2) = 1.9$  ,  $p = 0.3950$  ([Back](#))

Score	Category 0		Category 1	
	Observed	Expected	Observed	Expected
0	44	44.5	4	3.5
1	97	93.5	12	15.5
2	74	72.4	58	59.6
3	29	33.3	75	70.7

Item 4  $S-X^2(2) = 0.2$  ,  $p = 0.8902$  ([Back](#))

Score	Category 0		Category 1	
	Observed	Expected	Observed	Expected
0	44	44.2	25	24.8
1	76	77.0	69	68.0
2	17	17.8	79	78.2
3	8	7.0	75	76.0

The tables shown above illustrate good fit. In cases in which the model fits poorly, deviations between observed and expected may be large, and there may be long "runs" of red (or blue) in columns, as the trace line is "over" or "under" due to some model misspecification. (The latter can really only be observed in longer tests.)

Because the "full tables" for  $S-\chi^2$  can be very large, printing them is optional. Printing of optional results is controlled by clicking the **Options** button (bottom left-hand corner of the **Unidimensional Analysis** window) and then selecting the **Advanced Options, Miscellaneous** dialog:

Advanced Options X

Test: 2PL Apply to all tests

Estimation | Starting Values | Priors | Miscellaneous | Save | Simulate

Number of decimal places in tabular 2

Number of processors: 4

Print table of standardized residuals

Compute Chen-Thissen LD and item fit statistics

Compute limited-information overall model fit statistics

Print each item's goodness of fit frequency table  
Minimum expected 1

Print factor loadings

Print parameter numbers

Print diagnostic information

Print d\_ump file  Print full d\_ump file

OK Cancel Apply