

# IRTPRO™

## Detection of differential item functioning

The trace line model has many uses, but Lord (1977, 1980) observed that the trace line is ideally suited to defining differential item functioning (DIF), or lack thereof.

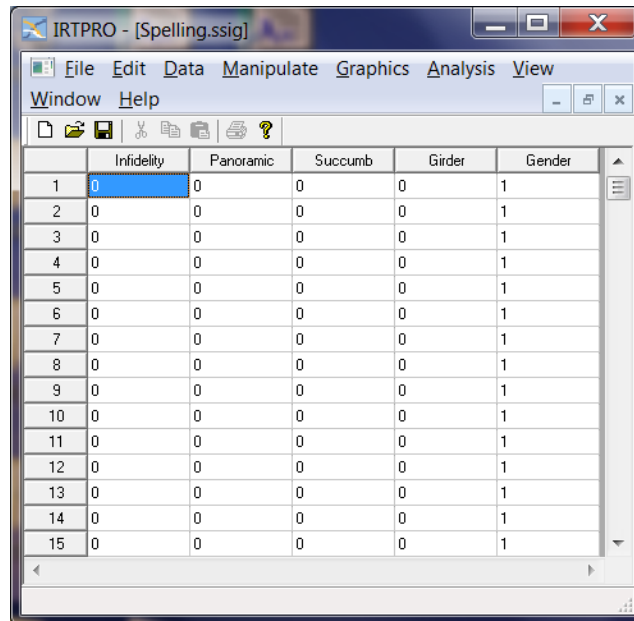
The value of the trace line at each level of  $\theta$  is the conditional probability of a correct response given that level of ability or proficiency. If we are considering the possibility that an item may function differently (exhibit DIF) for some focal group relative to some reference group, then in the context of IRT we are considering whether the trace lines differ for the two groups. If the trace lines are the same, there is no DIF. If the trace lines differ, there is DIF. Because the trace line for an item is determined by the item parameters, Lord (1977, 1980) noted that the question of DIF detection could be approached by comparing estimates of the item parameters between groups.

We illustrate DIF with data obtained from 659 undergraduates at the University of Kansas. The data obtained are based on a conventional orally administered spelling test. The reference group for this analysis includes the male students ( $N = 285$ ) and the focal group is made up of the female students ( $N = 374$ ). The original test comprised 100 words, but only 4 have been selected for use here. The four words to be spelled are Infidelity, Panoramic, Succumb, and Girder. These four items were selected because preliminary analysis suggested that they have very nearly equal discrimination parameters; this is convenient for purposes of illustration. The data were free-response, so there is no guessing to be considered. The words Infidelity, Panoramic and Succumb were selected to comprise an "anchor" (a set of items believed to involve no DIF) with information over a range of the  $\theta$ -continuum. The word Girder is the studied item; it was selected because it shows substantial differential difficulty for the two groups in these data. Thissen, Steinberg, and Wainer (1993) describe several DIF analyses of these data.

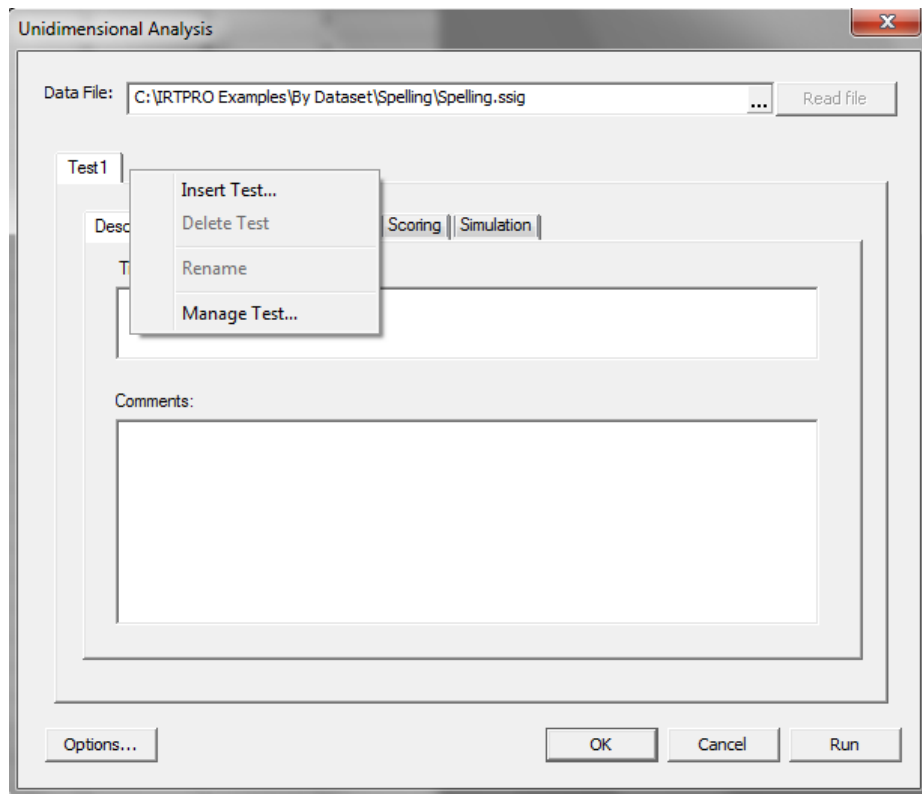
To see the data, use the **Open** file dialog under the **File** menu of IRTPRO, navigate to the **C:\IRTPRO Examples\By Dataset\Spelling** folder, select **Files of type: IRTPRO Data File (\*.ssig)** in the **Open File** dialog, and open the file **Spelling.ssig**. The first 15 cases of the data are shown below.

To initiate the unidimensional IRT analysis, select **Unidimensional IRT ...** under the **Analysis** menu. We start by adding two additional tests in order to fit a total of three models to the same data. By right-clicking on the right side of the **Test1** tab, a pop-up menu is obtained that enable the user to enter a new test, delete a test or rename an existing test, the default test names being

**Test1, Test2, ...** Once the **Test2** tab is displayed, the procedure is repeated to obtain **Test1, Test2, and Test3.**



	Infidelity	Panoramic	Succumb	Girder	Gender
1	0	0	0	0	1
2	0	0	0	0	1
3	0	0	0	0	1
4	0	0	0	0	1
5	0	0	0	0	1
6	0	0	0	0	1
7	0	0	0	0	1
8	0	0	0	0	1
9	0	0	0	0	1
10	0	0	0	0	1
11	0	0	0	0	1
12	0	0	0	0	1
13	0	0	0	0	1
14	0	0	0	0	1
15	0	0	0	0	1



Unidimensional Analysis

Data File: C:\IRTPRO Examples\By Dataset\Spelling\Spelling.ssig Read file

Test1

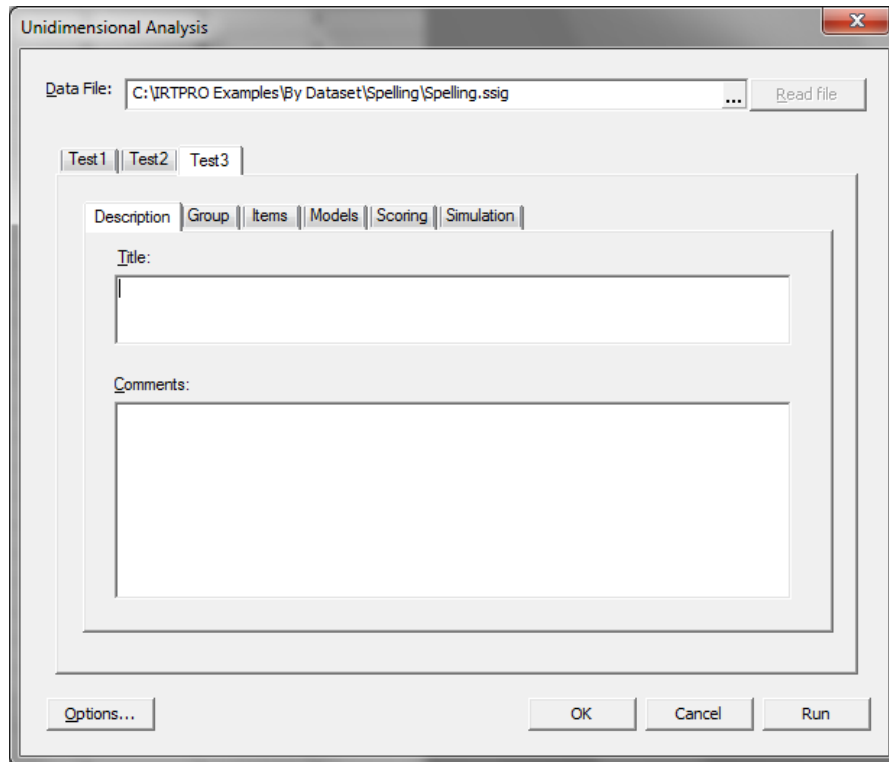
- Insert Test...
- Delete Test
- Rename
- Manage Test...

Scoring | Simulation

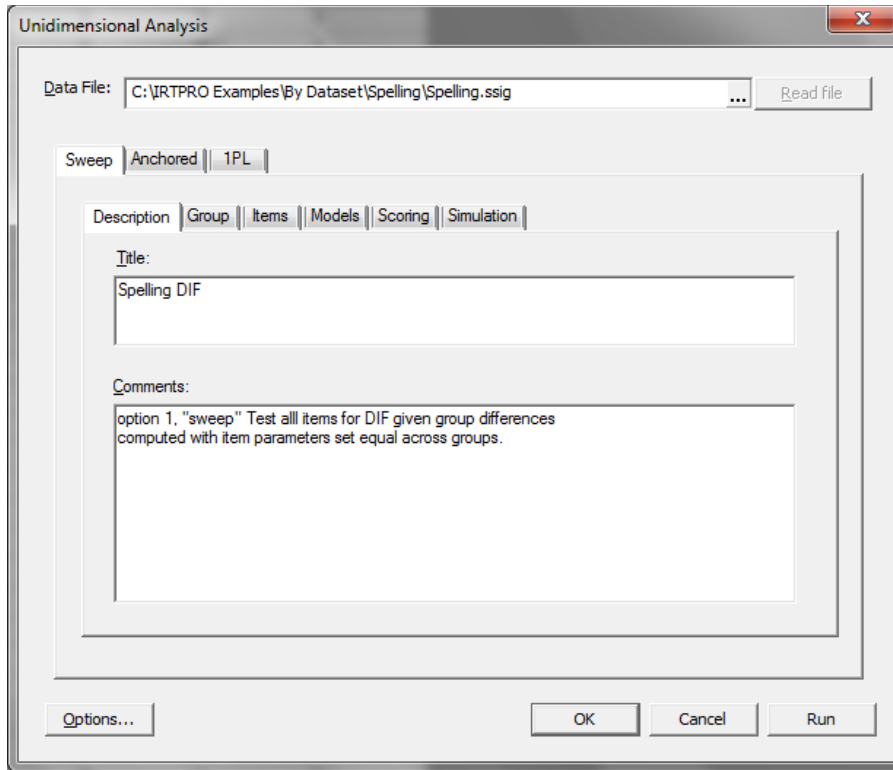
Comments:

Options... OK Cancel Run

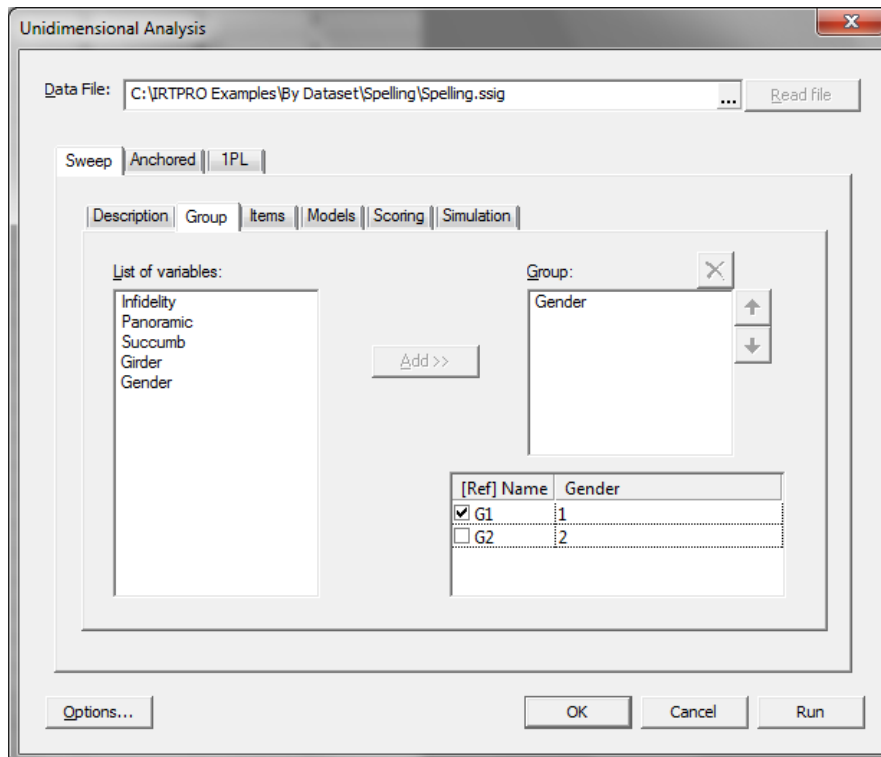
Right-click on each of the test tabs to invoke the **Insert Test...**, **Delete Test**, etc. drop-down menu. Click the **Rename** button and replace Test1 with Sweep, Test2 with Anchored, and Test3 with 1PL, the latter being more descriptive of the type of models to be fitted to the item response data.



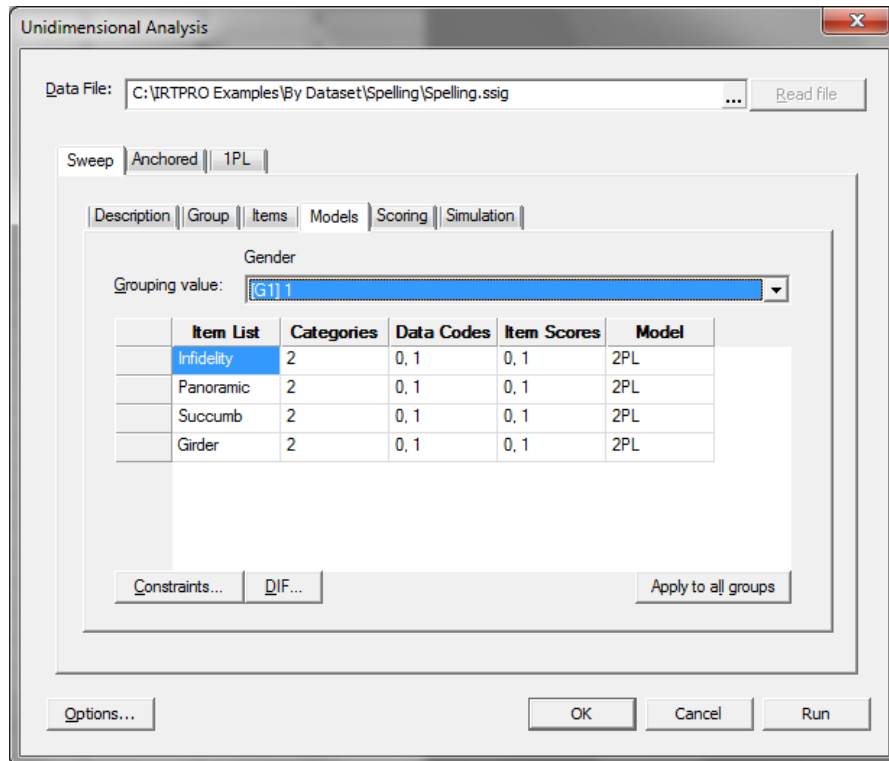
Start by clicking the **Sweep** tab to create the first test. Click the **Description** tab (the default tab for a new analysis) to add a title and comments.



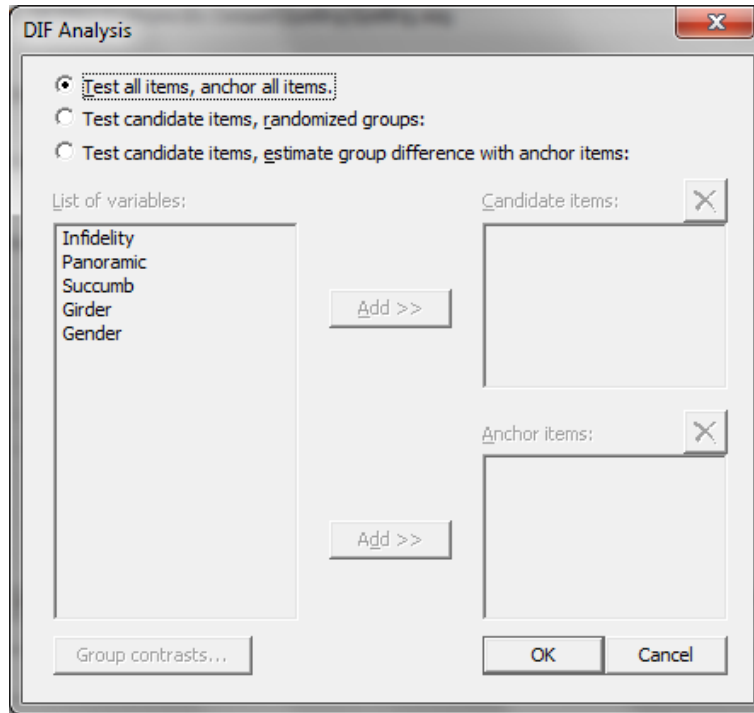
Next, click the **Groups** tab and select Gender as the grouping variable. The default reference group is the first group as shown below.



Use the **Items** tab in the **Sweep** tab in the **Unidimensional Analysis** dialog to select the four items to be analyzed, and click on the **Apply to all groups** button to select those items for both groups. Proceeding to the **Models** tab, it is seen that all four items have two categories and that the 2PL model is displayed as the default for each item.



In the **Models** window (see above), click the **DIF...** button. This brings up the DIF dialog shown next.



In this analysis we accept the default **Test all items, anchor all items**, so we click the **OK** button to return to the **Models** window and click the **Run** button to initiate the analysis. Portions of the output are as follows.

**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	Infidelity	<sup>2</sup> 0.88	0.26	<sup>1</sup> 1.30	0.18	-1.48	0.37
2	Panoramic	<sup>4</sup> 1.48	0.35	<sup>3</sup> 0.78	0.19	-0.52	0.14
3	Succumb	<sup>6</sup> 1.86	0.54	<sup>5</sup> -1.19	0.27	0.64	0.13
4	Girder	<sup>8</sup> 1.44	0.36	<sup>7</sup> 0.72	0.19	-0.50	0.14

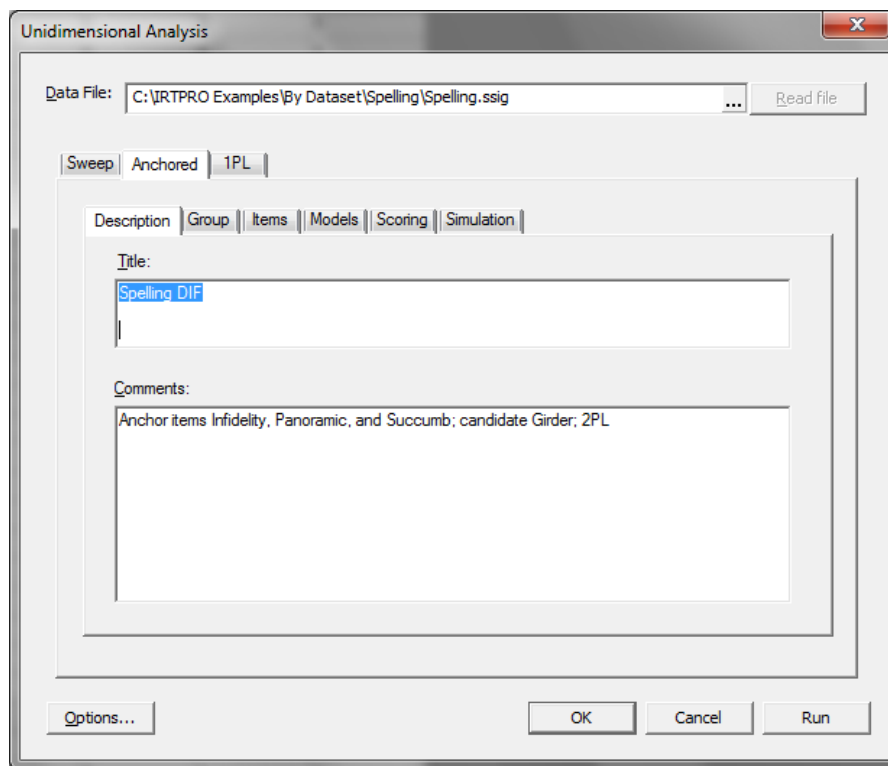
**2PL Model Item Parameter Estimates for Group 2, logit:  $a\theta + c$  or  $a(\theta - b)$**  ([Back to TOC](#))

Item	Label	a	s.e.	c	s.e.	b	s.e.
1	Infidelity	<sup>10</sup> 1.62	0.43	<sup>9</sup> 2.29	0.34	-1.41	0.22
2	Panoramic	<sup>12</sup> 1.51	0.37	<sup>11</sup> 0.94	0.19	-0.62	0.11
3	Succumb	<sup>14</sup> 1.14	0.29	<sup>13</sup> -0.78	0.13	0.69	0.21
4	Girder	<sup>16</sup> 1.64	0.43	<sup>15</sup> 0.14	0.16	-0.09	0.09

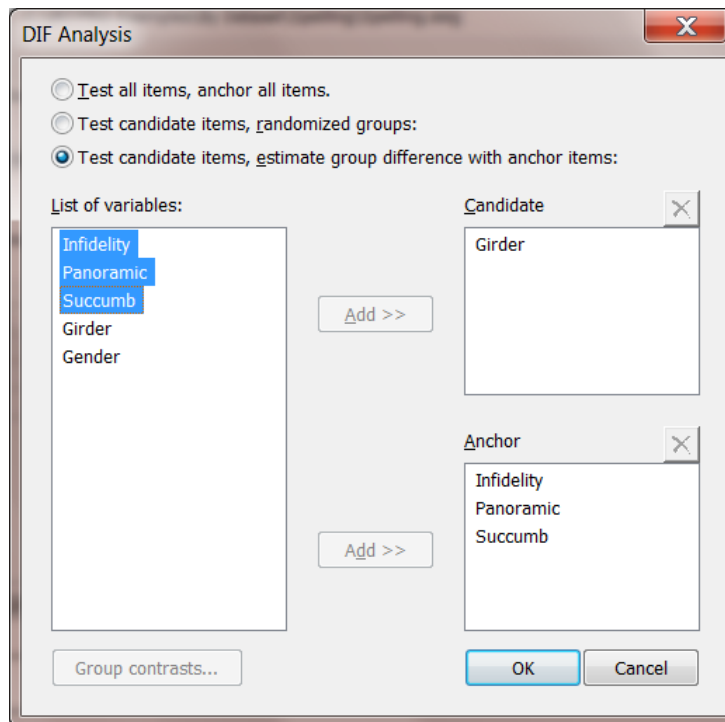
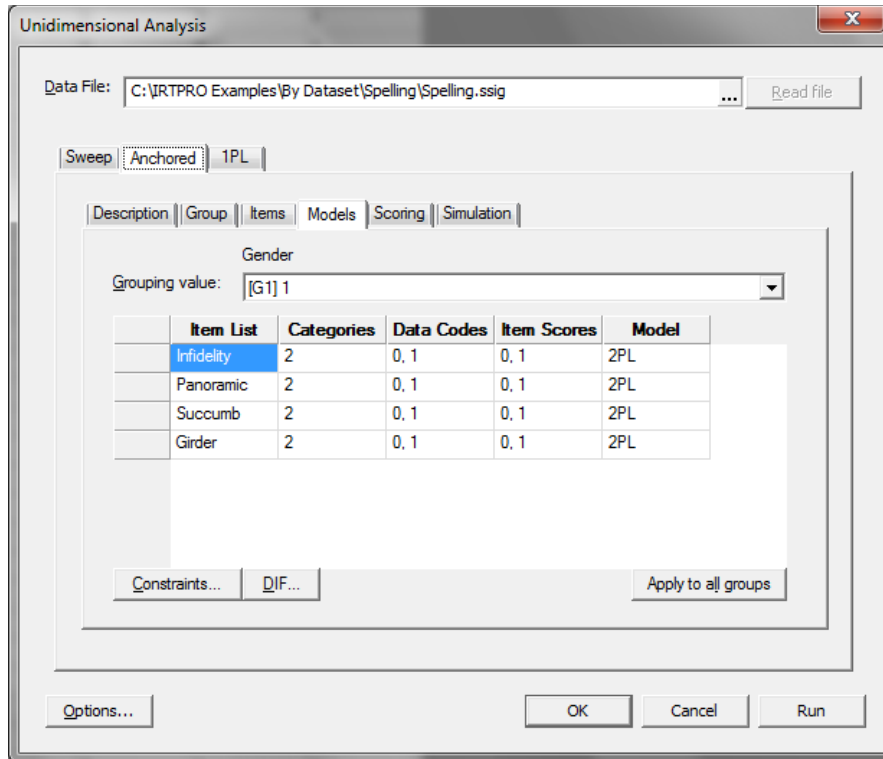
**DIF Statistics for Graded Items** ([Back to TOC](#))

Item numbers in:											
Group 1	Group 2	Total $X^2$	d.f.	p	$X^2_a$	d.f.	p	$X^2_{cja}$	d.f.	p	
1	1	7.5	2	0.0235	2.2	1	0.1348	5.3	1	0.0219	
2	2	0.5	2	0.7765	0.0	1	0.9516	0.5	1	0.4787	
3	3	2.1	2	0.3544	1.4	1	0.2393	0.7	1	0.4069	
4	4	8.2	2	0.0167	0.1	1	0.7164	8.1	1	0.0045	

To return to the **Unidimensional Analysis** environment, close the output window and select the **Unidimensional Analysis** option from the **Analysis** option on the main menu bar. To proceed to setting up the second test, click the **Anchored** tab and select **Description** to provide a title and comments.



From the **Group** tab, Gender is again selected as the grouping variable and, using the **Items** tab, all four items are selected. Proceeding to the **Models** tab, it is seen that each item is associated with the 2PL model.



Next, we click the **DIF** button to display the **DIF Analysis** window. Once this window is displayed, the **Test candidate items, estimate group difference with anchor items:** option is selected. Next, Girder is selected as candidate item and the remaining three items as anchor



items.

Click the **OK** button to return to the **Models** window and click the **Run** button to initiate the analysis. Portions of the output are as follows.

**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	Infidelity	<sup>2</sup> 1.18	0.24	<sup>1</sup> 1.59	0.16	-1.34	0.24
2	Panoramic	<sup>4</sup> 1.47	0.28	<sup>3</sup> 0.67	0.16	-0.46	0.11
3	Succumb	<sup>6</sup> 1.51	0.30	<sup>5</sup> -1.09	0.17	0.72	0.13
4	Girder	<sup>8</sup> 1.47	0.37	<sup>7</sup> 0.73	0.19	-0.50	0.13

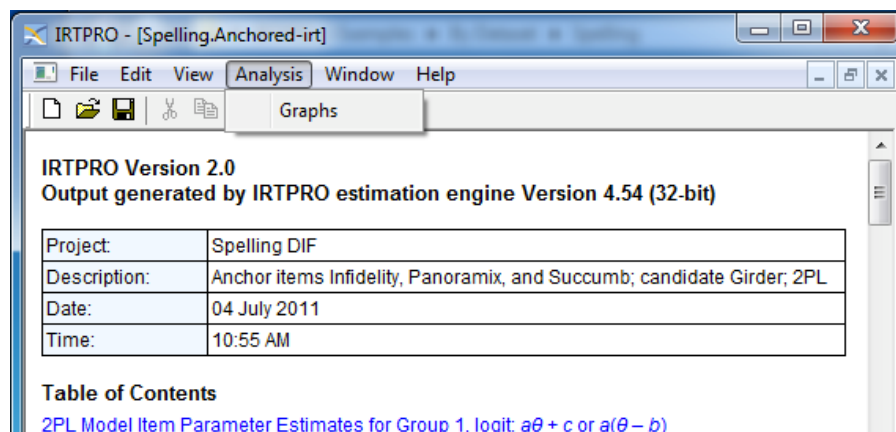
**2PL Model Item Parameter Estimates for Group 2, logit:  $a\theta + c$  or  $a(\theta - b)$**  ([Back to TOC](#))

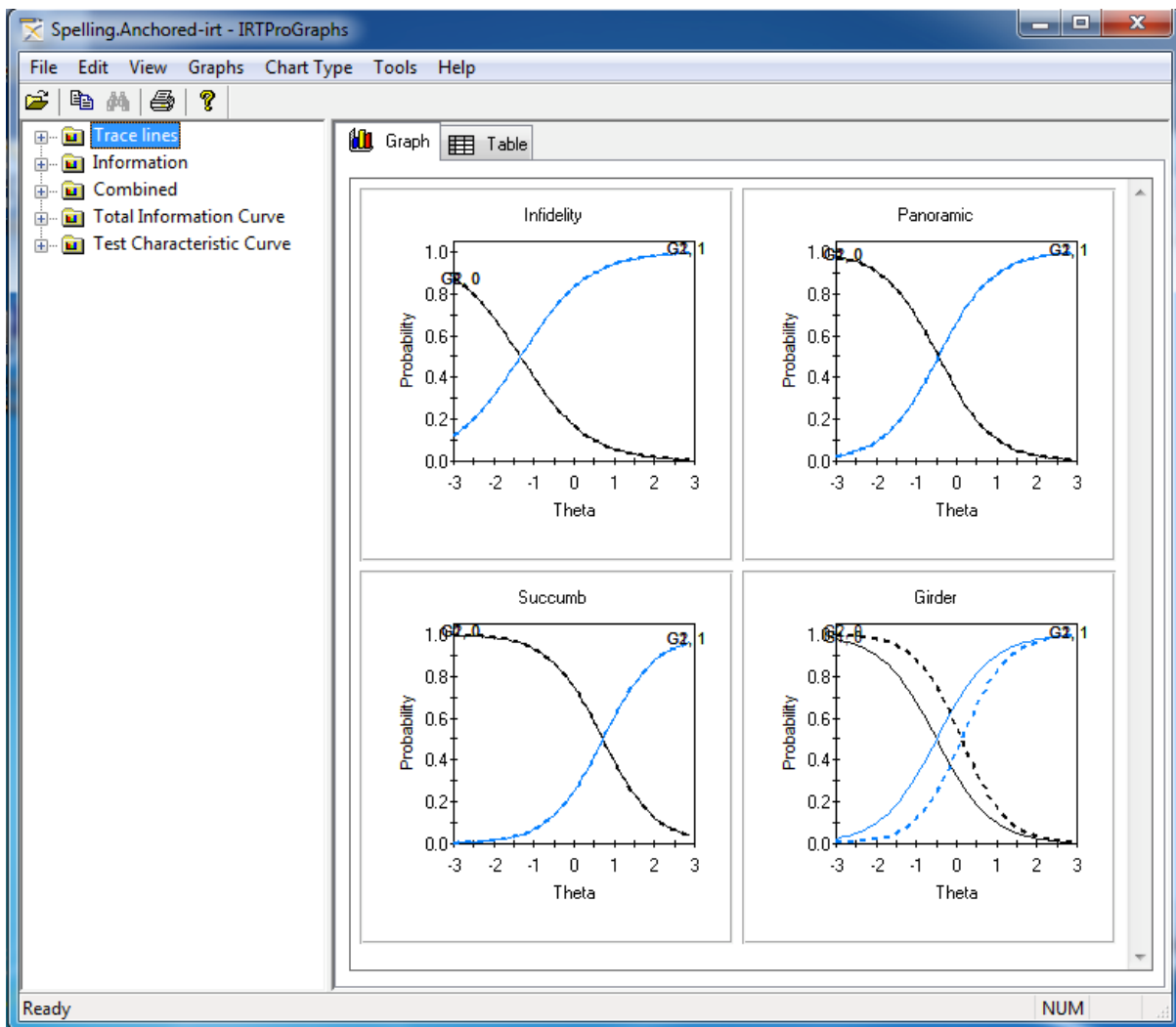
Item	Label	a	s.e.	c	s.e.	b	s.e.
1	Infidelity	<sup>2</sup> 1.18	0.24	<sup>1</sup> 1.59	0.16	-1.34	0.24
2	Panoramic	<sup>4</sup> 1.47	0.28	<sup>3</sup> 0.67	0.16	-0.46	0.11
3	Succumb	<sup>6</sup> 1.51	0.30	<sup>5</sup> -1.09	0.17	0.72	0.13
4	Girder	<sup>10</sup> 1.77	0.54	<sup>9</sup> -0.21	0.22	0.12	0.12

**DIF Statistics for Graded Items** ([Back to TOC](#))

Item numbers in:											
Group 1	Group 2	Total $X^2$	d.f.	p	$X^2_a$	d.f.	p	$X^2_{c a}$	d.f.	p	
4	4	11.2	2	0.0036	0.2	1	0.6413	11.0	1	0.0009	

While the output is displayed, the **Graphs** option is available under the **Analysis** item on the main menu bar.



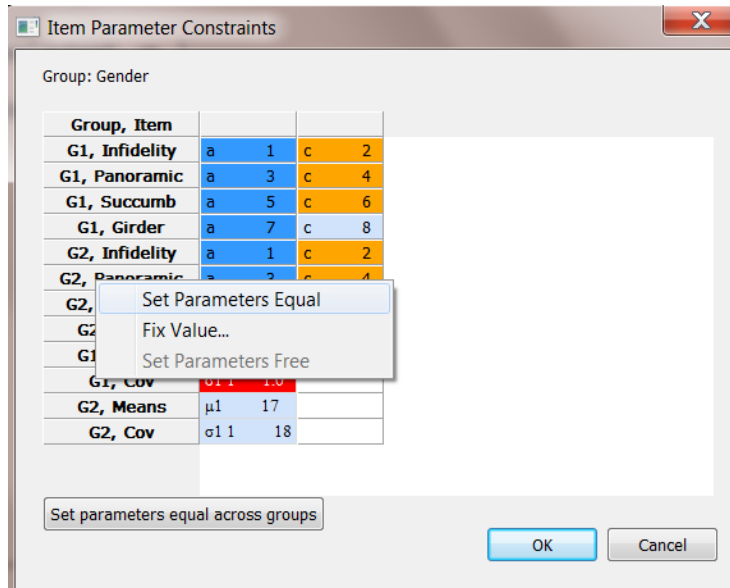


To return to the **Unidimensional Analysis** environment, close the graphics window (if it is open) and the output window, and select this option from the **Analysis** option on the main menu bar. To proceed to setting up the third test, click the **1PL** tab.

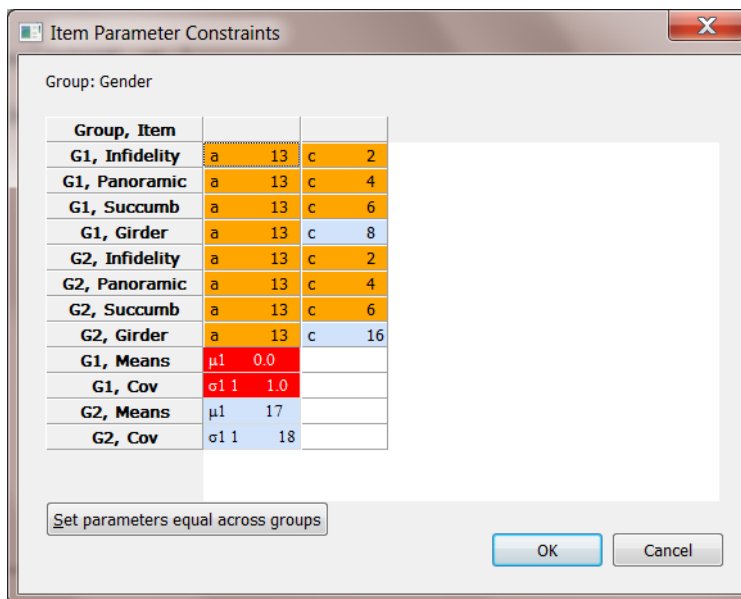
Use the **Description** tab to provide a title and comments. From the **Group** tab, Gender is again selected as the grouping variable and using the **Items** tab all four items are selected.

Next proceed to the **Models** tab, and click the **DIF** button to display the **DIF Analysis** window. Once this window is displayed, just as in the anchored example, the **Test candidate items, estimate group difference with anchor items:** option is selected. Next Girder is selected as candidate item and the remaining three items as anchor items. Click **OK** to return to the **Models** tab.

To use the 1PL model for the analysis, emulating Thissen, Steinberg, and Wainer (1993), we next click the **Constraints** button in the **Models** tab to bring up the **Constraints** dialog, and set all the slope ( $a$ ) parameters for both groups equal:



Note that it is important that such constraints are added after the DIF dialog, so that the constraints are in addition to a set of constraints automatically imposed by the anchored- DIF selection:



Once done, click the **OK** button to return to the **Models** window and click the **Run** button to

initiate the analysis. Portions of the output are as follows.

**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	Infidelity	<sup>6</sup> 1.41	0.15	<sup>1</sup> 1.70	0.16	-1.20	0.14
2	Panoramic	<sup>6</sup> 1.41	0.15	<sup>2</sup> 0.66	0.15	-0.47	0.10
3	Succumb	<sup>6</sup> 1.41	0.15	<sup>3</sup> -1.08	0.14	0.76	0.13
4	Girder	<sup>6</sup> 1.41	0.15	<sup>4</sup> 0.72	0.17	-0.51	0.12

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

Item	Label	a	s.e.	c	s.e.	b	s.e.
1	Infidelity	<sup>6</sup> 1.41	0.15	<sup>1</sup> 1.70	0.16	-1.20	0.14
2	Panoramic	<sup>6</sup> 1.41	0.15	<sup>2</sup> 0.66	0.15	-0.47	0.10
3	Succumb	<sup>6</sup> 1.41	0.15	<sup>3</sup> -1.08	0.14	0.76	0.13
4	Girder	<sup>6</sup> 1.41	0.15	<sup>5</sup> -0.21	0.18	0.15	0.13

**DIF Statistics for Graded Items** ([Back to TOC](#))

Item numbers in:											
Group 1	Group 2	Total $X^2$	d.f.	p	$X^2_a$	d.f.	p	$X^2_{c a}$	d.f.	p	
4	4	14.4	2	0.0007	0.0	1	1.0000	14.4	1	0.0001	

These DIF statistics are slightly deceptive; they list a two degrees of freedom total  $\chi^2$ , and a  $\chi^2$  test between the  $a$  parameters even though the  $a$  parameters are constrained equal (so the latter test must be 0.0), and there is only one degree of freedom for the test of the  $b$  parameter, which is the same as the overall test in this case. The software is not smart enough to count arbitrary constraints when doing DIF tests, and the user has to correct these things.