

# THE ANALYSIS OF CONTINUOUS DATA FROM MULTIPLE GROUPS

## 1. Introduction

In practice, many multivariate data sets are observations from several groups. Examples of these groups are genders, languages, political parties, countries, faculties, colleges, schools, etc. For these data sets, it is often of interest to determine whether or not the grouping variable has any influence on the structural equation model for the observed variables. The statistical methods for multiple group structural equation modeling may be used to determine whether or not the grouping variable has an influence on the model.

LISREL 8.7 for Windows (Jöreskog & Sörbom 2004) may be used to fit multiple group structural equation models to multiple group data. Traditional statistical methods such as Maximum Likelihood (ML), Robust Maximum Likelihood (RML), Weighted Least Squares (WLS), Diagonally Weighted Least Squares (DWLS), Generalized Least Squares (GLS) and Un-weighted Least Squares (ULS) are available for complete multiple group data while the Full Information Maximum Likelihood (FIML) method is available for multiple group data with missing values. The ML, RML, WLS, DWLS, GLS and ULS methods for multiple group structural equation modeling are described in Jöreskog & Sörbom (1999) while the FIML method is described in Du Toit & Du Toit (2001).

In this note, the FIML estimation method for data with missing values of LISREL 8.7 for Windows is used to fit a measurement model to two multivariate data sets consisting of the simulated scores of a sample of boys and girls on six psychological tests. These data sets are described in the next section. The measurement model is described in Section 3. Thereafter, the measurement model is fitted to separately to the two data sets. The multiple-group modeling feature of LISREL 8.7 for Windows is used in Section 5 to cross-validate the measurement model across the two gender groups. In Section 6, the invariance of the factor loadings across the two gender groups is tested.

## 2. The Data

The data are the simulated scores of 1250 girls and 1250 boys on six psychological tests. The raw data are listed in the PRELIS System Files (PSF) **MGGIRLS.PSF** and **MGBOYS.PSF** in the **MISSINGEX** subfolder of the LISREL 8.7 for Windows installation folder respectively. The first couple of lines of these two files are shown in the following PSF windows respectively.

	visperc	cubes	lozenges	paragraf	sentenc	wordmean	
1	-0.713	-1.957	-6.869	-9.000	-0.238	7.601	
2	-6.419	-8.642	-4.371	0.591	1.530	9.626	
3	8.788	-2.154	10.208	2.290	2.809	8.916	
4	9.723	-2.379	14.827	0.476	4.165	6.803	
5	-9.000	0.649	-8.022	-9.000	-7.507	-12.608	
6	-10.575	-9.000	-2.178	0.478	6.702	-9.000	
7	-9.000	7.800	-9.000	-9.000	-1.216	5.382	
8	-9.000	-2.138	5.238	-1.007	2.722	3.732	
9	2.595	5.720	-0.643	-0.039	-1.315	-9.000	
10	0.727	-7.937	-6.892	-9.000	-9.000	-27.537	
11	-3.663	-4.204	3.509	6.261	9.254	12.349	
12	-6.260	0.883	-14.828	4.856	0.880	12.270	
13	6.534	-9.000	-8.796	-9.000	-9.000	-1.789	
14	8.769	-3.621	-9.000	-3.469	-0.565	-9.000	
15	-2.607	-6.937	1.408	-0.649	-5.814	-9.000	
16	0.120	0.611	0.500	5.110	0.000	0.000	

	visperc	cubes	lozenges	paragraf	sentenc	wordmean	
1	-0.717	-1.950	-7.691	-9.000	-9.000	6.707	
2	-6.451	-8.069	-2.525	0.541	1.246	9.167	
3	8.831	-3.733	11.511	2.462	2.894	9.307	
4	-9.000	-4.125	16.955	1.225	4.645	8.286	
5	-9.000	1.083	-9.410	-9.000	-6.488	-11.788	
6	-10.627	-9.000	0.316	0.578	5.879	-9.000	
7	-9.000	6.405	-9.000	-9.000	-0.040	-9.000	
8	-9.000	-1.645	7.201	-0.431	2.983	4.680	
9	2.608	5.615	-2.337	-0.255	-1.387	-9.000	
10	0.730	-8.509	-6.620	-9.000	-9.000	-24.836	
11	-3.681	-3.835	5.573	5.550	7.617	10.738	
12	-6.291	1.971	-17.086	3.199	-0.966	8.663	
13	6.566	-9.000	-11.938	-9.000	-9.000	-2.802	
14	8.812	-5.281	-9.000	-2.741	0.142	-9.000	
15	-2.620	-9.000	3.515	-0.236	-4.566	-9.000	
16	0.140	1.171	1.700	1.500	0.000	0.000	

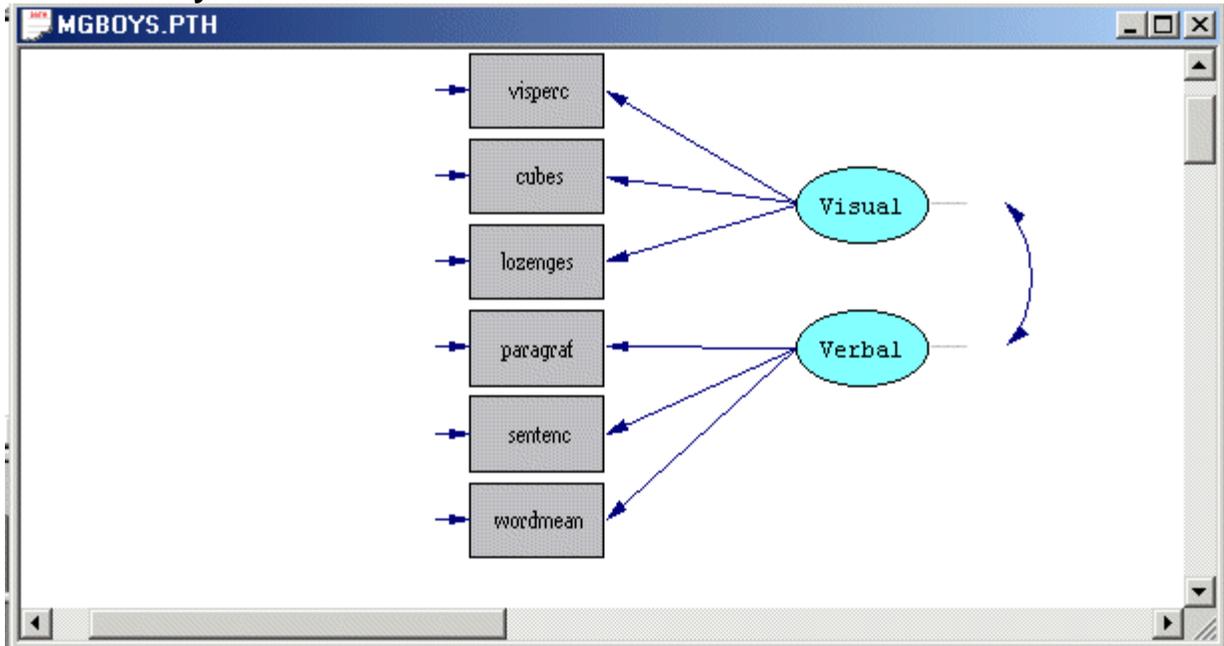
Note that the -9.000 entries represent missing values.

### 3. The Measurement Model

We consider six psychological tests for Junior High students. These six psychological tests are theoretically constructed to measure the visual perception and verbal ability of Junior

High students. A path diagram for the corresponding measurement model for visual perception and verbal ability is shown in Figure 1.

**Figure 1: A path diagram for a measurement model of Visual Perception and Verbal Ability**

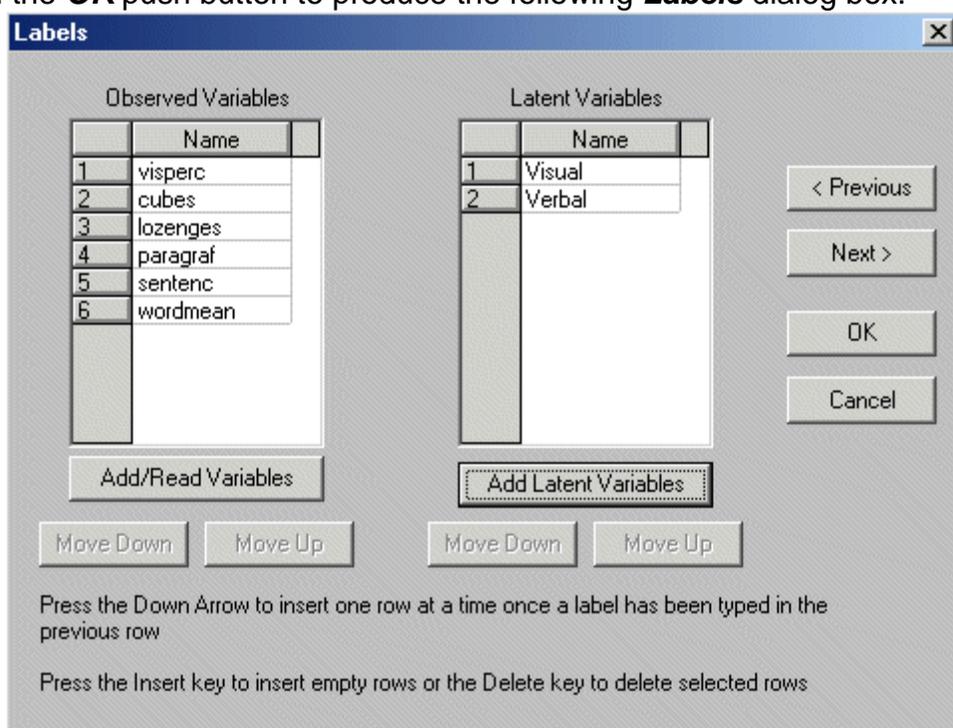


## 4. Separate Analyses

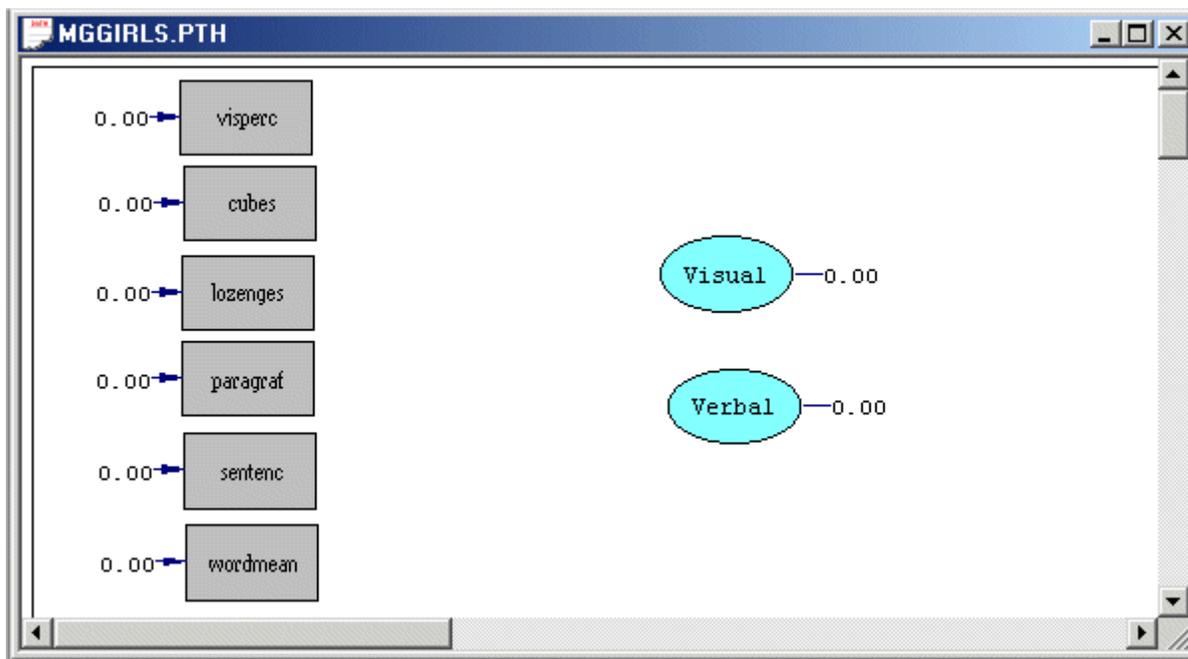
### Fitting the Model to the Girl Data

- Use the **New** option on the **File** menu of the root window to load the **New** dialog box.
- Select the **Path Diagram** option from the list box on the **New** dialog box to load the **Save As** dialog box.
- Enter **MGGIRLS.PTH** in the **File name** string field.
- Click on the **Save** push button to open an empty path diagram window.
- Select the **Title and Comments** option on the **Setup** menu to load the **Title and Comments** dialog box.
- Enter **A measurement model for visual perception and verbal ability of girls** in the **Title** string field.
- Click on the **Next** push button to load the **Group Names** dialog box.
- Click on the **Next** push button to load the **Labels** dialog box.
- Click on the **Add/Read Variables** push button to load the **Add/Read Variables** dialog box.
- Select the **PRELIS System File** option in the **Read from file:** dropdown list box.
- Click on the **Browse** push button to browse for the **MISSINGEX** folder of the LISREL 8.7 for Windows installation folder.
- Select the file **MGGIRLS.PSF**.
- Click on the **Open** push button to return to the **Add/Read Variables** dialog box.
- Click on the **OK** push button to return to the **Labels** dialog box.
- Click on the **Add Latent Variables** push button to load the **Add Variables** dialog box.
- Enter the label **Visual** for visual perception in the string field.

- Click on the **OK** push button to return to the **Labels** dialog box.
- Click on the **Add Latent Variables** push button to load the **Add Variables** dialog box.
- Enter the label **Verbal** for verbal ability in the string field.
- Click on the **OK** push button to produce the following **Labels** dialog box.

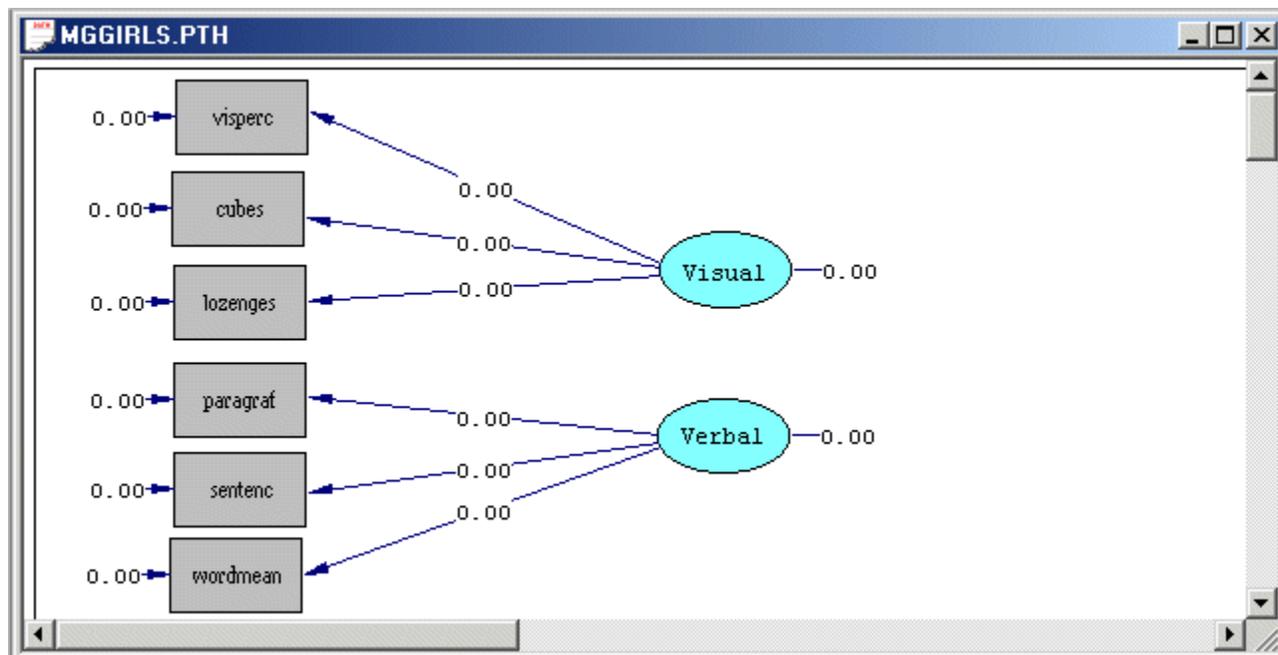


- Click on the **OK** push button to return to the PTH window for **MGGIRLS.PTH**.
- Click, drag and drop the labels of the six observed variables one at a time into the PTH window.
- Click, drag and drop the labels of the two latent variables one at a time into the PTH window to produce the following PTH window.



- Select the arrow icon on the drawing toolbar.
- Click and drag paths from **Visual** to **visperc**, **cubes** and **lozenges** respectively.

- Click and drag paths from **Verbal** to **paragraf**, **sentenc** and **wordmean** respectively to produce the following PTH window.



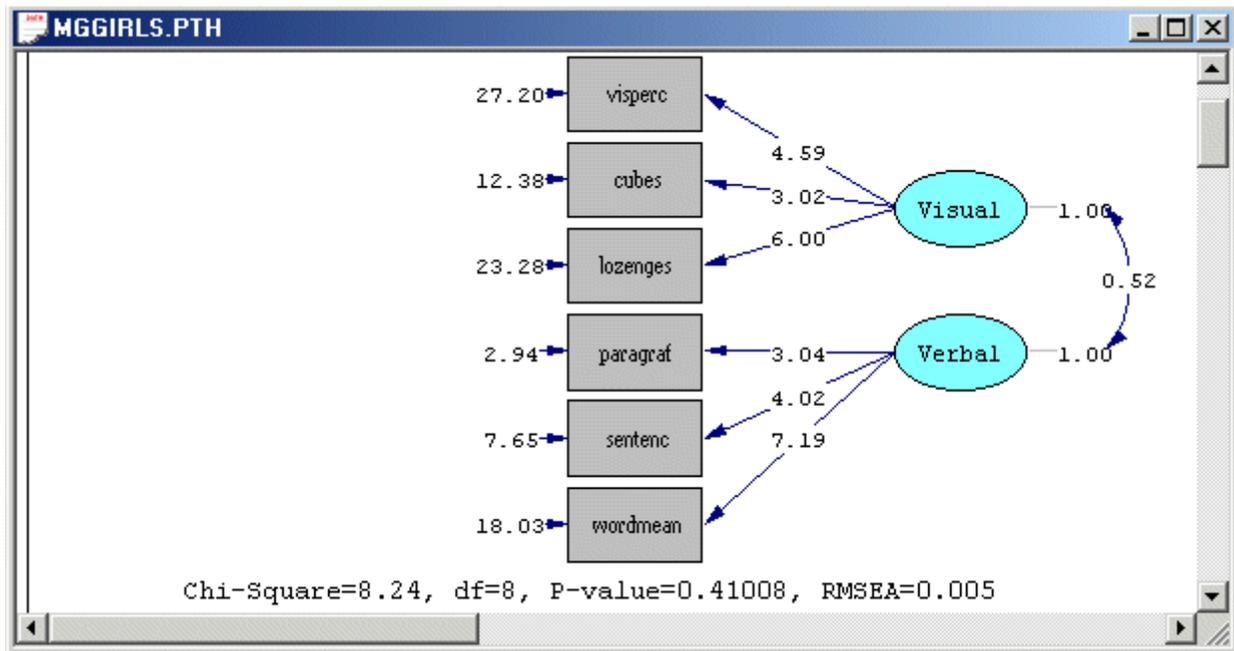
- Select the **Build SIMPLIS Syntax** option on the **Setup** menu to open the following SIMPLIS project window.

```

A measurement model for visual perception and verbal ability o
Raw Data from file 'C:\lisrel854_student\MISSINGEX\MGGIRLS.PSF
Latent Variables Visual Verbal
Relationships
visperc = Visual
cubes = Visual
lozenges = Visual
paragraf = Verbal
sentenc = Verbal
wordmean = Verbal
Path Diagram
End of Problem

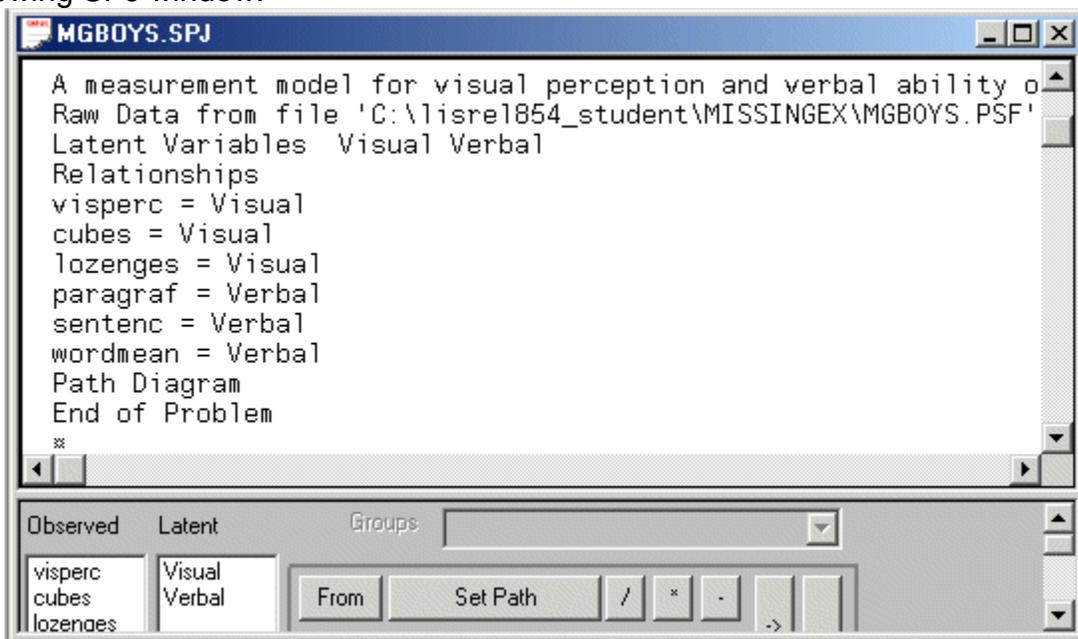
```

- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.

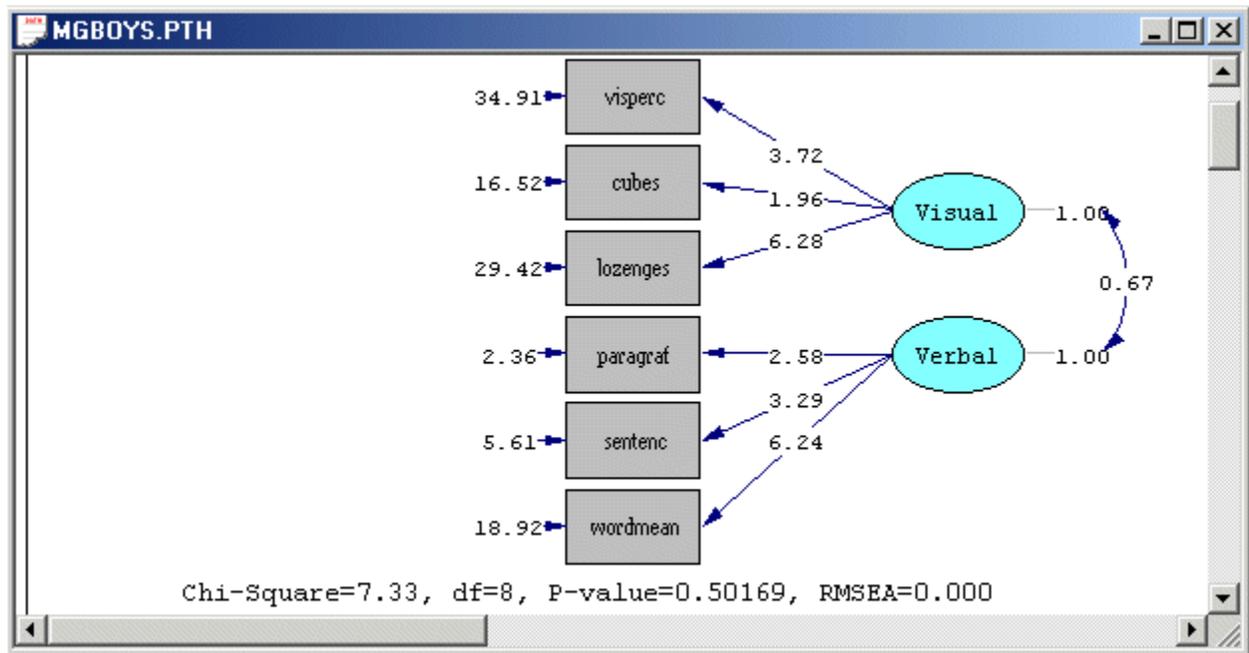


### Fitting the Measurement Model to the Boy Data

- Close the PTH window for **MGGIRLS.PTH**.
- Close the output window for **MGGIRLS.OUT** to return to the SPJ window for **MGGIRLS.SPJ**.
- Use the **Save As** option on the **File** menu of the root window to load the **Save As** dialog box.
- Enter the name **MGBOYS.SPJ** in the **File name** string field.
- Click on the **Save** push button to open the SPJ window for **MGBOYS.SPJ**.
- Change **girls** in the first line to **boys**.
- Change the file name **MGGIRLS.PSF** in the second line to **MBOYS.PSF** to create the following SPJ window.



- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.

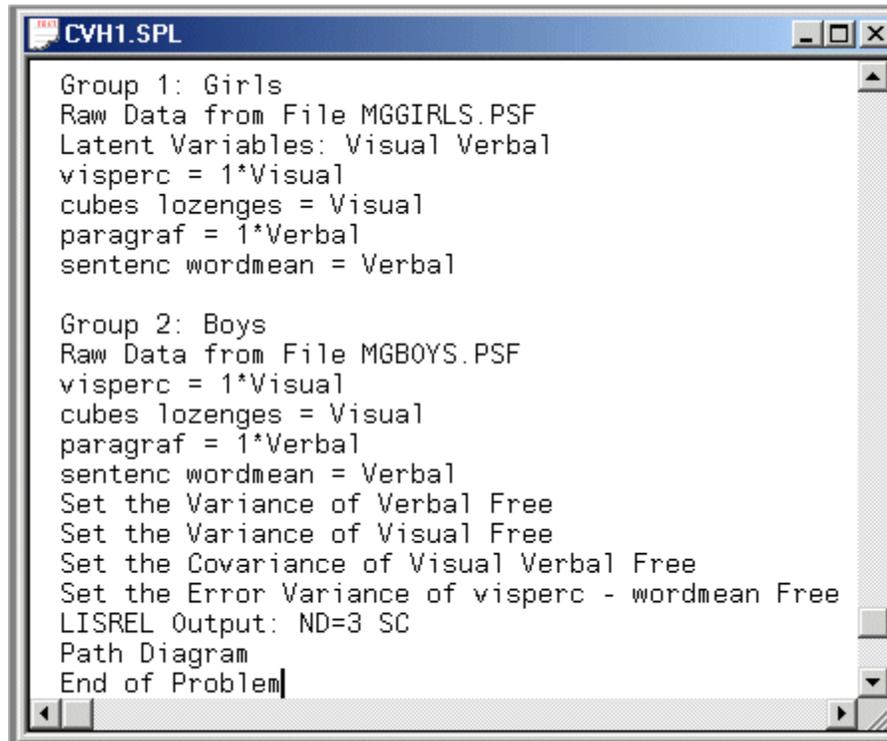


## 5. Cross Validation

The cross validation of a measurement model refers to the ability of the model to be invariant across two or more random samples from the same population. The measurement model in Figure 1 will be used to illustrate how the multiple group feature of LISREL 8.7 for Windows may be used to assess the cross validation of a measurement model. This assessment consists of testing the null hypothesis (H0) that the measurement model is identical across gender against the alternative hypothesis (H1) that the model is not identical across gender. A Chi-square difference test is used to test H0 and H1.

### Fitting the Multiple-group Model under H1

- Use the **New** option on the **File** menu of the root window to load the **New** dialog box.
- Select the **Syntax Only** option from the list box on the **New** dialog box to open the **SYNTAX1** text editor window.
- Enter the following commands into the **SYNTAX1** text editor window.

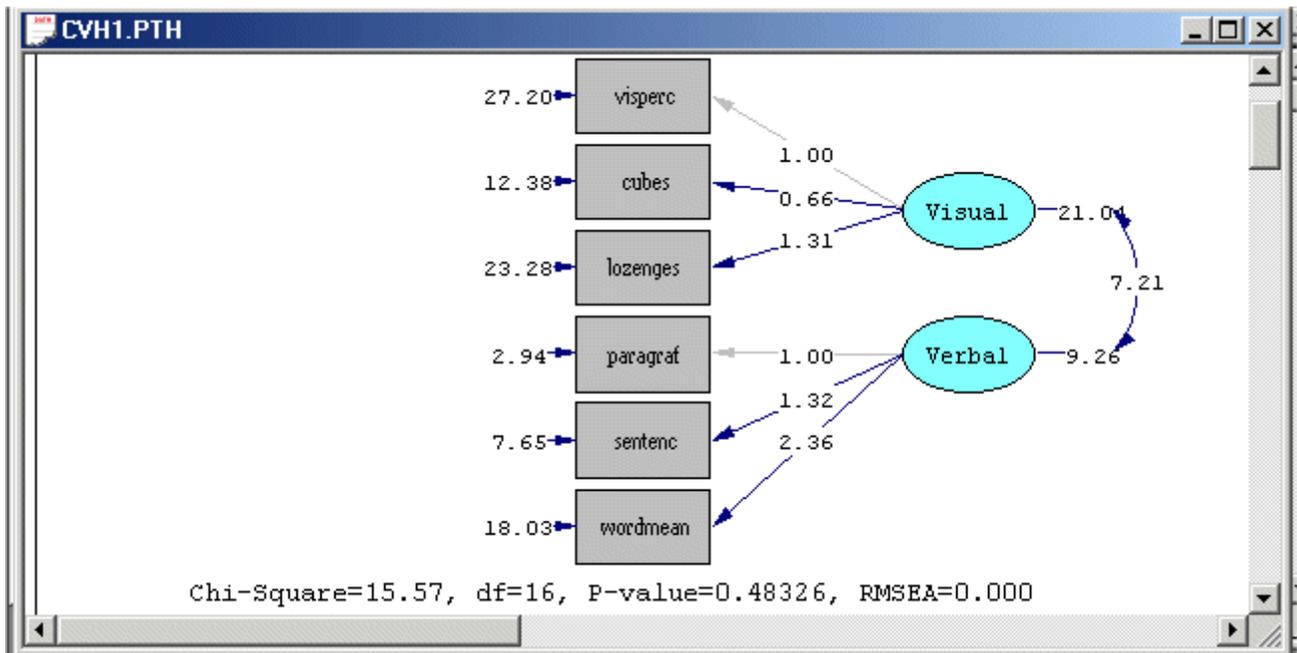


```
CVH1.SPL
Group 1: Girls
Raw Data from File MGGIRLS.PSF
Latent Variables: Visual Verbal
visperc = 1*Visual
cubes lozenges = Visual
paragraf = 1*Verbal
sentenc wordmean = Verbal

Group 2: Boys
Raw Data from File MGBBOYS.PSF
visperc = 1*Visual
cubes lozenges = Visual
paragraf = 1*Verbal
sentenc wordmean = Verbal
Set the Variance of Verbal Free
Set the Variance of Visual Free
Set the Covariance of Visual Verbal Free
Set the Error Variance of visperc - wordmean Free
LISREL Output: ND=3 SC
Path Diagram
End of Problem
```

- Line 1 specifies the first group as girls.
- Line 2 specifies the location and name of the raw data file for girls.
- Line 3 specifies descriptive labels for the two latent variables.
- Lines 4-7 specify the measurement model for girls.
- Line 9 specifies the second group as boys.
- Line 10 specifies the location and name of the raw data file for boys.
- Lines 11-18 specify the measurement model for boys. Lines 15-18 specify that the variance and covariance parameters of the model are different across the two groups.
- Line 19 requests the results in terms of the parameter matrices of the LISREL model for the measurement model in Figure 1. In addition, 3 decimal places (ND=3) and the completely standardized solutions (SC) are specified.
- Line 20 requests a path diagram (PTH) file.
- Line 21 indicates the end of the SIMPLIS commands to be processed.

- Use the **Save As** option on the **File** menu to load the **Save As** dialog box.
- Enter the name **CVH1.SPL** in the **File name** string field.
- Click on the **Save** push button to create the text editor window for **CVH1.SPL**.
- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.



### Fitting the Multiple-group Model under H0

- Use the **Open** option on the **File** menu to open the SIMPLIS syntax file **CVH1.SPL** in a text editor window.
- Delete lines 11-18.
- Use the **Save As** option on the **File** menu to load the **Save As** dialog box.
- Enter the name **CVH0.SPL** in the **File name** string field.
- Click on the **Save** push button to produce the following text editor window.

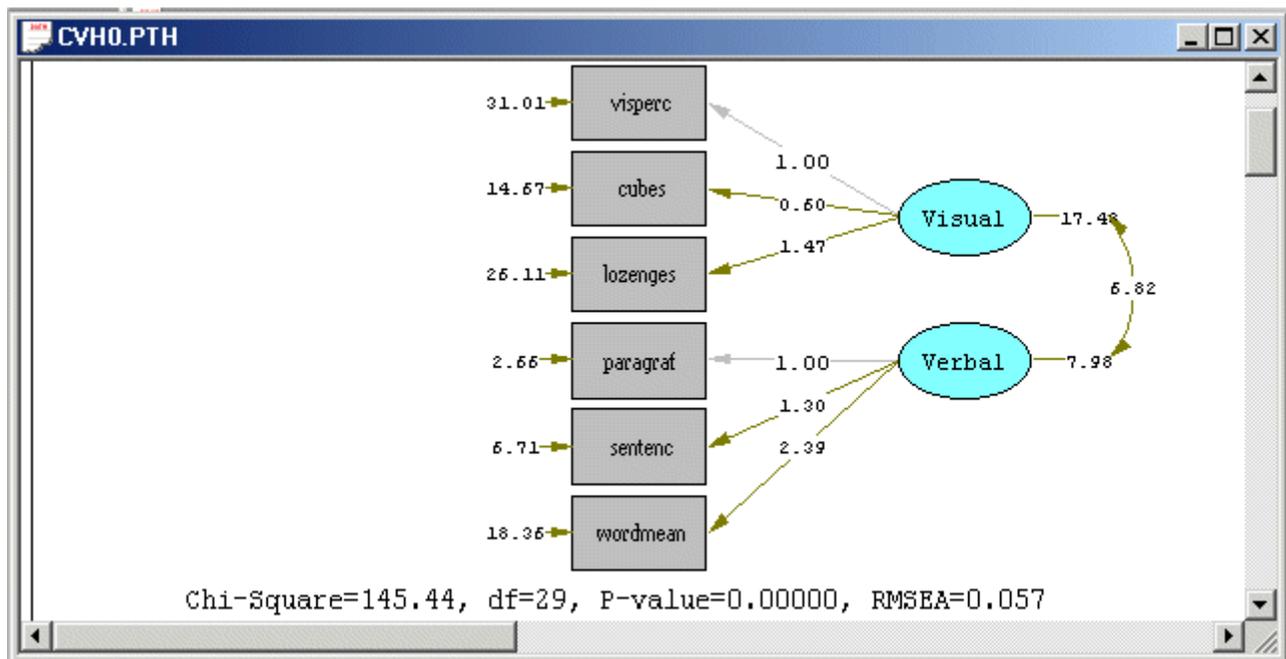
```

CVH0.SPL
Group 1: Girls
Raw Data from File MGGIRLS.PSF
Latent Variables
Visual Verbal
Relationships
visperc = 1*Visual
cubes lozenges = Visual
paragrap = 1*Verbal
sentenc wordmean = Verbal

Group 2: Boys
Raw Data from File MGBBOYS.PSF
LISREL Output: ND=3 SC
Path Diagram
End of Problem

```

- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.



### Performing the Chi-square difference test

A Chi-square difference test is used to assess the cross validation of the measurement model. In other words, a Chi-square difference test is used to test the null and alternative hypotheses. The test statistic value for the Chi-square difference test is merely the difference between the goodness-of-fit Chi-square test statistic values of the multiple group measurement models under the null and the alternative hypotheses. The associated degrees of freedom are merely the difference between the degrees of freedom of the multiple group measurement models under the null and the alternative hypotheses. The Chi-square difference test results for the measurement model for visual perception and verbal ability are summarized in the MS-Excel workbook **CV.XLS**. The contents of this file are shown below.

	A	B	C	D
1	HYPOTHESIS	CHISQ	DF	P-VALUE
2	EQUAL (H0)	145.442	29	1.88015E-17
3	UNEQUAL (H1)	15.571	16	0.483281241
4	DIFFERENCE	129.871	13	2.21864E-21

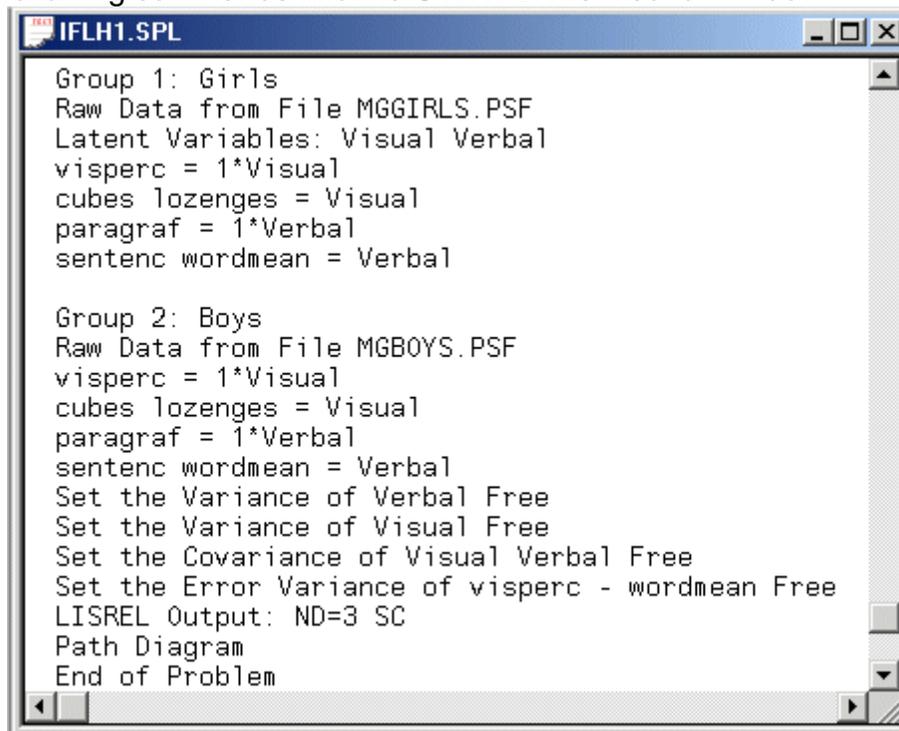
The small P-value suggests that there is sufficient evidence that the null hypothesis should be rejected. In other words, the cross validation of the measurement model for the six psychological tests across gender is not supported by the data of the two samples.

## 6. Invariance of Factor Loadings

The measurement model in Figure 1 will now be used to illustrate how the multiple group feature of LISREL 8.7 for Windows may be used to assess whether or not the factor loadings of the measurement model are invariant across the 2 gender groups. This assessment consists of testing the null hypothesis (H0) that the factor loadings are identical across gender against the alternative hypothesis (H1) that the factor loadings are not identical across gender. A Chi-square difference test is used to test H0 and H1.

### Fitting the Multiple-group Model under H1

- Use the **New** option on the **File** menu of the root window to load the **New** dialog box.
- Select the **Syntax Only** option from the list box on the **New** dialog box to open the **SYNTAX1** text editor window.
- Enter the following commands into the **SYNTAX1** text editor window.



```
IFLH1.SPL
Group 1: Girls
Raw Data from File MGGIRLS.PSF
Latent Variables: Visual Verbal
visperc = 1*Visual
cubes lozenges = Visual
paragraf = 1*Verbal
sentenc wordmean = Verbal

Group 2: Boys
Raw Data from File MGBBOYS.PSF
visperc = 1*Visual
cubes lozenges = Visual
paragraf = 1*Verbal
sentenc wordmean = Verbal
Set the Variance of Verbal Free
Set the Variance of Visual Free
Set the Covariance of Visual Verbal Free
Set the Error Variance of visperc - wordmean Free
LISREL Output: ND=3 SC
Path Diagram
End of Problem
```

Line 1 specifies the first group as girls.

Line 2 specifies the location and name of the raw data file for girls.

Line 3 specifies descriptive labels for the two latent variables.

Lines 4-7 specify the measurement model for girls.

Line 9 specifies the second group as boys.

Line 10 specifies the location and name of the raw data file for boys.

Lines 11-18 specify the measurement model for boys. Lines 15-18 specify that the variance and covariance parameters of the model are different across the two groups.

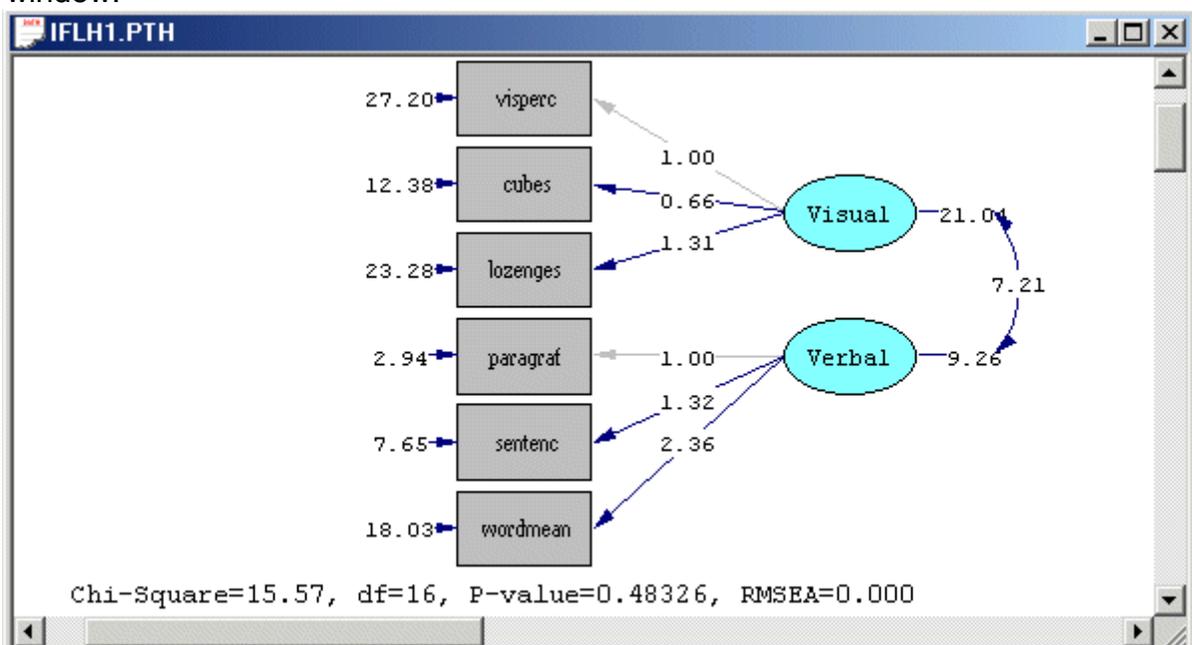
Line 19 requests the results in terms of the parameter matrices of the LISREL model for the measurement model in Figure 1. In addition, 3 decimal places (ND=3) and the completely standardized solutions (SC) are specified.

Line 20 requests a path diagram (PTH) file.

Line 21 indicates the end of the SIMPLIS commands to be processed.

- Use the **Save As** option on the **File** menu to load the **Save As** dialog box.
- Enter the name **IFLH1.SPL** in the **File name** string field.
- Click on the **Save** push button to create the text editor window for **CVH1.SPL**.

- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.



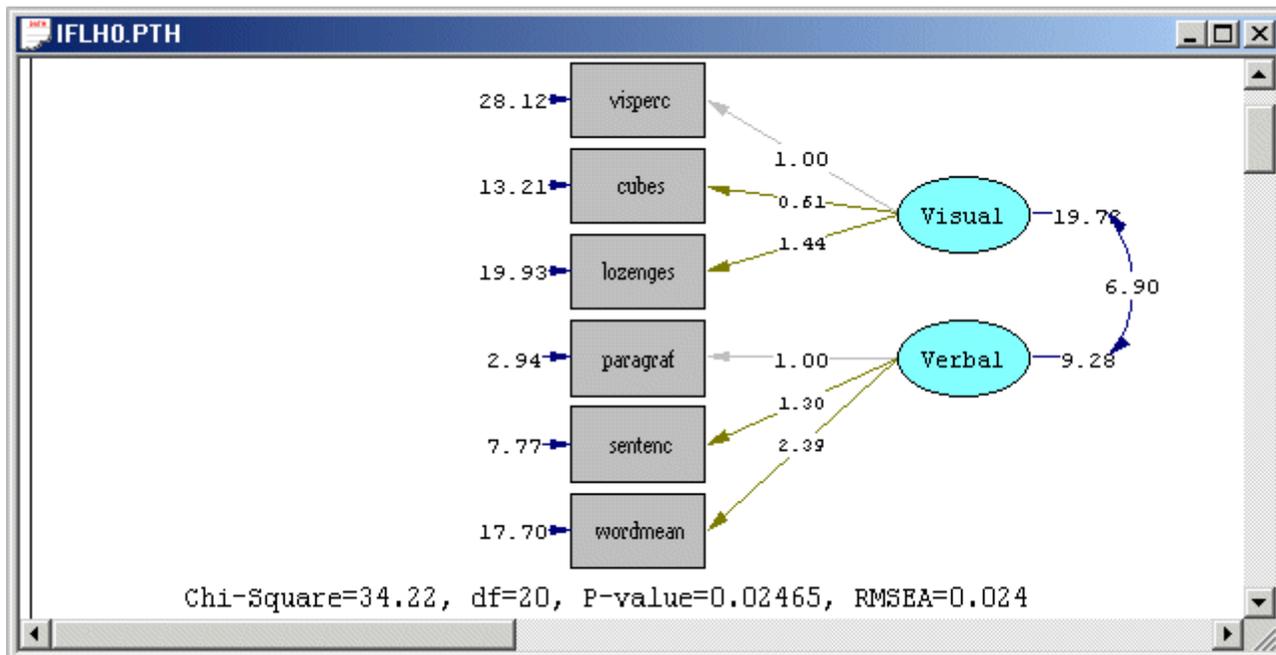
### Fitting the Multiple-group Model under H0

- Use the **Open** option on the **File** menu to open the SIMPLIS syntax file **IFLH1.SPL** in a text editor window.
- Delete lines 11-14.
- Use the **Save As** option on the **File** menu to load the **Save As** dialog box.
- Enter the name **IFLH0.SPL** in the **File name** string field.
- Click on the **Save** push button to produce the following text editor window.

```
IFLH0.SPL
Group 1: Girls
Raw Data from File MGGIRLS.PSF
Latent Variables: Visual Verbal
visperc = 1*Visual
cubes lozenges = Visual
paragraf = 1*Verbal
sentenc wordmean = Verbal

Group 2: Boys
Raw Data from File MGBBOYS.PSF
Set the Variance of Verbal Free
Set the Variance of Visual Free
Set the Covariance of Visual Verbal Free
Set the Error Variance of visperc - wordmean Free
LISREL Output: ND=3 SC
Path Diagram
End of Problem
```

- Click on the **Run LISREL** icon on the main toolbar to produce the following PTH window.



### Performing the Chi-square difference test

A Chi-square difference test is used to assess whether or not the factor loadings of the measurement model are invariant across the two gender groups. In other words, a Chi-square difference test is used to test the null and alternative hypotheses. The test statistic value for the Chi-square difference test is merely the difference between the goodness-of-fit Chi-square test statistic values of the multiple group measurement models under the null and the alternative hypotheses. The associated degrees of freedom are merely the difference between the degrees of freedom of the multiple group measurement models under the null and the alternative hypotheses. The Chi-square difference test results for the measurement model for visual perception and verbal ability are summarized in the MS-Excel workbook *IFL.XLS*. The contents of this file are shown below.

	A	B	C	D
1	HYPOTHESIS	CHISQ	DF	P-VALUE
2	EQUAL (H0)	34.223	20	0.024655171
3	UNEQUAL (H1)	15.571	16	0.483281241
4	DIFFERENCE	18.652	4	0.000919818

The small P-value suggests that there is sufficient evidence that the null hypothesis should be rejected. In other words, there is sufficient evidence that the factor loadings for boys and girls are different.

### References

- Du Toit, M. & Du Toit, S.H.C. (2001).  
*Interactive LISREL: User's Guide*. Lincolnwood, IL: Scientific Software International, Inc.
- Jöreskog, K.G. & Sörbom, D. (1999).  
*LISREL 8: User's Reference Guide*.  
 Lincolnwood, IL: Scientific Software International, Inc.
- Jöreskog, K. & Sörbom, D. (2004).  
*LISREL 8.7 for Windows [Computer Software]*.  
 Lincolnwood, IL: Scientific Software International, Inc.