NIELS J. BLUNCH

Introduction to STRUCTURAL EQUATION MODELING USING IBM SPSS STATISTICS AND EQS

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

Los Angeles | London | New Delhi
Singapore | Washington DC

SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Editor: Jai Seaman
Assistant editor: James Piper
Production editor: Victoria Nicholas
Copyeditor: Neville Hankins
Proofreader: Kate Campbell
Indexer: David Rudeforth
Marketing manager: Sally Ransom
Cover design: Shaun Mercier
Typeset by: C&M Digitals (P) Ltd, Chennai, India
Printed and bound by CPI Group (UK) Ltd,
Croydon, CR0 4YY

MIX
Paper from responsible sources
FSC® C013604

Library of Congress Control Number: 2015934011

British Library Cataloguing in Publication data

A catalogue record for this book is available from the British Library

ISBN 978-1-47391-621-0
ISBN 978-1-47391-622-7 (pbk)

7

The Measurement Model in SEM: Confirmatory Factor Analysis

We start by examining the differences between the three models, which are usually put under the common designation of factor analysis: namely, principal components analysis, exploratory factor analysis and confirmatory factor analysis – the latter being in fact the measurement model of SEM.

Next, you will learn two rules for identification in confirmatory factor models and you will explore estimation of confirmatory factor models through a classic example. You will also learn how EQS can help you to obtain parsimony, i.e. to find as simple and uncomplicated (but still well-fitting) model as possible.

Then you will learn how confirmatory factor analysis can be used to select items for inclusion in a measurement model or for use in a summated scale. You will also learn how to use SEM to measure reliability and validity in ways that are more in accordance with the theoretical definition of these concepts than those presented in Chapter 2.

The chapter ends with a short discussion of reflective and formative indicators, and points to a problem in item selection that has often led researchers astray.

1 The Three Factor Models

In component analysis the components are linear functions of the original variables, whereas in factor analysis – whether exploratory or confirmatory – the roles are reversed: the variables are considered functions of latent variables called factors.

However, there are a number of important differences between exploratory factor analysis and confirmatory factor analysis.

In exploratory factor analysis:

1. Every manifest variable is connected with every latent variable (as in component analysis).
2. Error terms are uncorrelated.
3. All parameters are estimated from the data.

In confirmatory factor analysis some or all of the above rules are violated:

1. Manifest variables are only connected with some pre-specified latent variables, the ideal being that every manifest variable is an indicator for one and only one factor.

2. Some error terms may be allowed to correlate.
3. Some of the parameters may be restricted to certain values or to have the same values as other parameters, or they may be restricted to fulfill other conditions.

The differences among the three factor models are depicted in Figure 1 (it deserves mentioning that the component model and the exploratory factor model are shown prior to a possible rotation that could introduce correlations among the components or factors).

Comparing Figure 1 with Figure 1.5, it is obvious that the measurement model in SEM is a confirmatory factor model, and that it is this model that was the basis for the arguments against the classical methods of measuring reliability and validity in Chapter 2.

Remember that an indicator X_i can be:

- a single item in a many-item scale;
- a simple sum of several items, i.e. a summated scale; or
- a weighted sum of several items, e.g. a principal component.

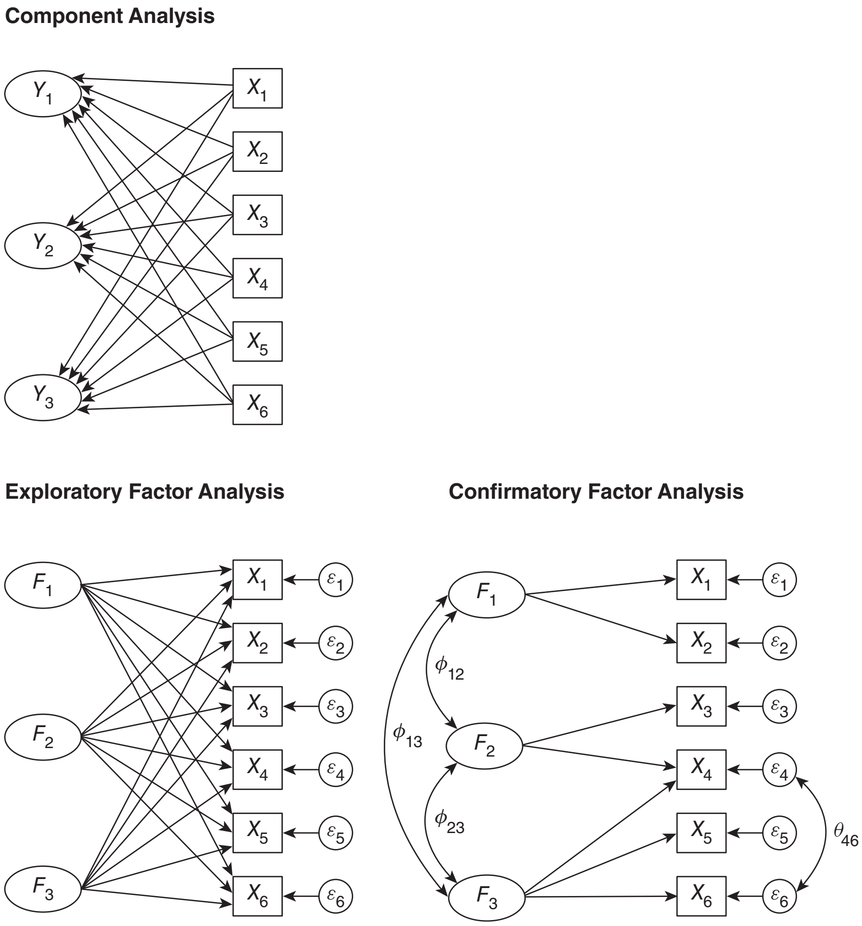


Figure 1 The three factor models

2 Identification and Estimation of Confirmatory Factor Models

In order to obtain identification every factor must be assigned a scale, either by fixing its variance or by fixing one of its regression coefficients; the same goes for the error terms. Further, the *t*-rule must be fulfilled but, as you have learned, this rule is only necessary, not sufficient.

Identification in confirmatory factor models

Two rules – both of which are sufficient, but not necessary – are worth mentioning (cf. Figure 4.2):

1. *The three-indicator rule*: A confirmatory factor model is identified if:
   (a) Every factor has at least three indicators.
   (b) No manifest variable is an indicator for more than one factor.
   (c) The error terms are not correlated.
2. *The two-indicator rule*: A confirmatory factor model with at least two factors is identified if:
   (a) Every factor has at least two indicators.
   (b) No manifest variable is an indicator for more than one factor.
   (c) The error terms are not correlated.
   (d) The covariance matrix for the latent variables does not contain zeros.

The main advantage of confirmatory models is that prior knowledge can be taken into account when formulating the model. Further, confirmatory models open up for various methods of testing the models.

Example 1
Democracy in developing countries

Bollen has published several studies on the determinants of democratic development (e.g. Bollen 1979, 1980). In this example I use data from Bollen (1989). The data given in Table 1 contain eight variables. The first four are indications of degree of democracy in 75 developing countries in 1960, the four variables being:

V1 = freedom of the press

V2 = freedom of group opposition

V3 = fairness of election

V4 = elective nature of the legislative body

The next four variables are the same variables measured the same way in 1965.

Table 1 Example 1: covariance and means matrix

	V1	V2	V3	V4	V5	V6	V7	V8
V1	6.89							
V2	6.25	15.58						
V3	5.84	5.84	10.76					
V4	6.09	9.51	6.69	11.22				
V5	5.06	5.60	4.94	5.70	6.83			
V6	5.75	9.39	4.73	7.44	4.98	11.38		
V7	5.81	7.54	7.01	7.49	5.82	6.75	10.80	
V8	5.67	7.76	5.64	8.01	5.34	8.25	7.59	10.53
Means	5.46	4.26	6.56	4.45	5.14	2.98	6.20	4.04

A component analysis of the covariance matrix gives the first three eigenvalues as

$$\begin{matrix} 57.17 \\ 8.47 \\ 5.26 \end{matrix} \tag{1}$$

and presents a strong case for a one-component solution, although we know that in fact there are two sets of measurements, one for 1960 and one for 1965.

If we insist on a two-component solution, the component loadings (covariances) are (after a varimax rotation)

$$\begin{matrix} & F1 & F2 \\ V1 & 1.28 & 1.86 \\ V2 & 3.53 & 0.96 \\ V3 & 0.53 & 2.94 \\ V4 & 2.23 & 1.94 \\ V5 & 1.14 & 1.78 \\ V6 & 2.79 & 1.09 \\ V7 & 1.49 & 2.47 \\ V8 & 2.12 & 1.80 \end{matrix} \tag{2}$$

where we see that the two components in no way represent the two time periods.

Of course you can try an oblique rotation (which will give a correlation of 0.63 between the two components), or use exploratory factor analysis instead of component analysis, but none of these techniques will separate the two time periods.

This shows the danger of letting the automatic use of exploratory methods lead you astray without theoretical considerations.

Now, let us see how such considerations can guide us:

1. We must maintain two factors, one expressing the degree of democracy in 1960 and the other the degree of democracy in 1965.
2. It is reasonable to assume that error terms in the same (manifest) measurement in the two years will correlate.
3. In the same way we will assume that the error terms for V2 and V4 (and V6 and V8) correlate, as these measurements are based on the same written source.

A model along these lines is shown in Figure 2.

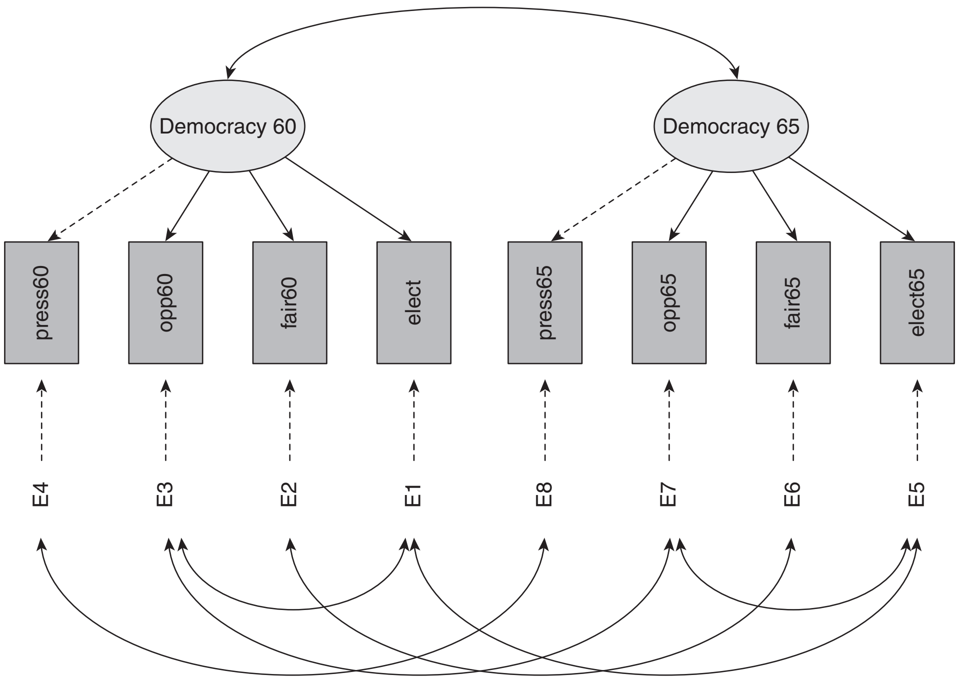


Figure 2 Example 1: the model

In order to estimate the model, two conditions must be satisfied:

1. We must create scales for the latent variables.
2. The model must be identified.

As mentioned earlier, the first condition can be met in two ways. The simpler way is to fix one of the regression coefficients for each factor to 1.00. This will transfer the scale of the indicator in question to its latent variable. Another possibility is to standardize the factors by fixing their variances to 1.00. On the screen the arrows connecting press60 and press65 to their respective factors are shown in red (as are the arrows connecting the E-variables to their manifest variables), signaling that I have chosen the first possibility.

Regarding point 2 above, we have 23 parameters to estimate:

6 coefficients

7 covariances

10 variances

As input we have 8 variances and 28 covariances. We thus have 13 'pieces of information' or degrees of freedom left over for testing. The *t*-rule is satisfied. However, the three-indicator rule cannot be used because of the correlated error terms. So the model is *possibly* identified.

A program for estimation of the model is given in Table 2.

Table 2 Example 1: EQS program

```
/TITLE
  Example 7.1. First run. Data from Bollen 1989
/SPECIFICATIONS
  DATA='C:\Users\Sony\Desktop\EQS bog\EQS Chap 7\Bollen.ess';
  VARIABLES=8; CASES=75;
  METHOD=ML; ANALYSIS=COVARIANCE; MATRIX=COVARIANCE;
/LABELS
  V1=press60; V2=opp60; V3=fair60; V4=elect60; V5=press65;
  V6=opp65; V7=fair65; V8=elect65;
  F1=dem60; F2=dem65
/EQUATIONS
  V1 =   1F1 + E1;
  V2 =   *F1 + E2;
  V3 =   *F1 + E3;
  V4 =   *F1 + E4;
  V5 =   1F2 + E5;
  V6 =   *F2 + E6;
  V7 =   *F2 + E7;
  V8 =   *F2 + E8;
/VARIANCES
   F1 = *;
   F2 = *;
   E1 TO E8 = *;
/COVARIANCES
  F2,F1 = *;
  E4,E2 = *;
  E5,E1 = *;
  E6,E2 = *;
  E7,E3 = *;
  E8,E4 = *;
  E8,E6 = *;
/PRINT
  FIT=ALL;
  TABLE=COMPACT;
/END
```

The program format should be familiar to you by now. In the /SPECIFICATIONS paragraph I state that the data should be chosen from the EQS data file 'Bollen.ess' and ask for GLS estimation. No new commands are involved, so I do not expect you to have any trouble with the contents.

Table 3 shows selected output.

Table 3 Example 1: output from first run (extract)

```
STANDARDIZED RESIDUAL MATRIX:                                              (1)


                        PRESS60    OPP60     FAIR60    ELECT60   PRESS65
                          V1        V2        V3        V4        V5
     PRESS60   V1        .001
     OPP60     V2       -.001      .014
     FAIR60    V3        .064     -.066     -.002
     ELECT60   V4       -.024      .018     -.004      .000
     PRESS65   V5        .002      .001      .026      .009     -.001
     OPP65     V6        .038      .034     -.096      .048     -.025
     FAIR65    V7       -.014      .002     -.014     -.006      .014
     ELECT65   V8       -.012      .035     -.050      .017     -.025


                        OPP65     FAIR65    ELECT65
                          V6        V7        V8
     OPP65     V6        .010
     FAIR65    V7       -.022     -.003
     ELECT65   V8        .015      .013      .004

              AVERAGE ABSOLUTE STANDARDIZED RESIDUAL =             .0203
  AVERAGE OFF-DIAGONAL ABSOLUTE STANDARDIZED RESIDUAL =             .0248


GOODNESS OF FIT SUMMARY FOR METHOD = ML

 INDEPENDENCE MODEL CHI-SQUARE    =   454.661 ON   28 DEGREES OF FREEDOM

 INDEPENDENCE AIC =      398.661    INDEPENDENCE CAIC =      305.771
        MODEL AIC =      -13.501           MODEL CAIC =      -56.628

 CHI-SQUARE =       12.499 BASED ON      13 DEGREES OF FREEDOM         (2)
 PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS         .48718

 THE NORMAL THEORY RLS CHI-SQUARE FOR THIS ML SOLUTION IS       11.679.

 FIT INDICES                                                           (3)
 -----------
 BENTLER-BONETT     NORMED FIT INDEX =      .973
 BENTLER-BONETT NON-NORMED FIT INDEX =     1.003
 COMPARATIVE FIT INDEX (CFI)         =    1.000
 BOLLEN'S            (IFI) FIT INDEX  =    1.001
 MCDONALD'S          (MFI) FIT INDEX  =    1.003
 JORESKOG-SORBOM'S   GFI  FIT INDEX   =     .962
 JORESKOG-SORBOM'S AGFI   FIT INDEX   =     .895
```

```
ROOT MEAN-SQUARE RESIDUAL (RMR)      =      .324
STANDARDIZED RMR                     =    .030
ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =       .000
90% CONFIDENCE INTERVAL OF RMSEA  (        .000,           .111)


MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)          (4)
PARAMETER ESTIMATES (B) WITH STANDARD ERRORS AND TEST STATISTICS (Z)
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                                                                R-
DEP.VAR.       PREDICTOR    B       BETA   S.E.   Z      SQUARED
----------------------------------------------------------------
V1  (PRESS60 )                                               .718
          F1  ( DEM60   ) 1.000    .848
          E1  (PRESS60 ) 1.000    .531
V2  (OPP60    )                                              .516
          F1  (  DEM60 ) 1.267*   .719  .185  6.831@
          E2  (OPP60    ) 1.000   .695
V3  (FAIR60  )                                               .524
          F1  ( DEM60   ) 1.069*   .724  .154  6.925@
          E3  (FAIR60  ) 1.000    .690
V4  (ELECT60 )                                               .715
          F1  ( DEM60   ) 1.274*   .846  .148  8.626@
          E4  (ELECT60 ) 1.000    .534
V5  (PRESS65 )                                               .620
          F2  ( DEM65   ) 1.000    .788
          E5  (PRESS65 ) 1.000    .616
V6  (OPP65    )                                              .567
          F2  ( DEM65   ) 1.227*   .753  .182  6.755@
          E6  (OPP65    ) 1.000    .658
V7  (FAIR65  )                                               .706
          F2  ( DEM65   ) 1.343*   .840  .174  7.736@
          E7  (FAIR65  ) 1.000    .542
V8  (ELECT65 )                                               .692
          F2  ( DEM65   ) 1.309*   .832  .172  7.597@
          E8  (ELECT65 ) 1.000    .555


 VARIANCES OF INDEPENDENT VARIABLES
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.


                      VARIANCE    S.E.       Z
------------------------------------------------
FACTOR
           F1  ( DEM60  )4.943*  1.132  4.368@
           F2  ( DEM65  )4.240*  1.086  3.905@
ERROR
           E1  (PRESS60 )1.937*   .454  4.269@
           E2  (OPP60   )7.427*  1.400  5.306@
           E3  (FAIR60  )5.130*   .975  5.264@
```

(Continued)

Table 3 (Continued)

```
          E4   (ELECT60 )3.195*    .754   4.240@
          E5   (PRESS65 )2.595*    .530   4.892@
          E6   (OPP65   )4.879*    .938   5.203@
          E7   (FAIR65  )3.183*    .723   4.405@
          E8   (ELECT65 )3.231*    .728   4.438@


COVARIANCES AMONG INDEPENDENT VARIABLES
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.
                                COVA.   S.E.   Z         CORR.
-----------------------------------------------------------
F1,F2    (DEM60   ,DEM65   ) 4.412*    .979  4.505@     .964
E1,E5    (PRESS60 ,PRESS65 )  .635*    .374  1.697      .283
E2,E4    (OPP60   ,ELECT60 ) 1.292*    .714  1.810      .265
E2,E6    (OPP60   ,OPP65   ) 2.073*    .744  2.786@     .344
E3,E7    (FAIR60  ,FAIR65  )  .828*    .619  1.336      .205
E4,E8    (ELECT60 ,ELECT65 )  .472*    .460  1.024      .147
E6,E8    (OPP65   ,ELECT65 ) 1.274*    .588  2.165@     .321
```

A few comments should clarify the contents

1. The standardized residual covariances (i.e. the residual correlations) are all extremely small, so there is not much to gain by introducing more parameters. On the contrary, it is perhaps possible to simplify the model by placing restrictions on parameter values.
2. We see that a χ^2-test of the model has a *P*-value of 0.487. This means that there is a probability of 0.487 of getting this result or one that is more against our model, if the model is correct. Because of the reversed testing we are interested in the test *not* being significant. Therefore a *P*-value near 50% does not seem unsatisfying. We are, however, not interested in very large *P*-values, as this could be a sign of over-fitting, and therefore suggest that the model could be simplified.
3. All the fit measures point in the same direction: We should indeed look for simplifications of the model.
4. Next, there are the parameter estimates, their standard errors and test statistics. The only parameters that are not significant by traditional standards are (`FAIR60,FAIR65`) and (`ELECT60,ELECT65`) (one-sided tests, $\alpha = 0.05$. Remember that the '`@`' are for two-sided tests, but here one-sided tests seem more appropriate.

As the fit is so good, it should perhaps be possible to simplify the model. Generally, we want a parsimonious model with as few free parameters as possible. The danger with models that are 'too good' is that the good fit perhaps is obtained by profiting from peculiarities in the sample at hand, and that the results therefore would not show up in other samples from the same population. So, the simpler the model, the smaller the dangers of generalization – and also the simpler the model, the easier it is to interpret.

You could for instance argue that the connection between a latent variable and its indicator is the same in the two years, i.e. that the measuring instrument itself does not change between the two measurements. If so, you put in the following paragraph in the program in Table 2:

```
/CONSTRAINTS
(V2,F1) = (V6,F2);
(V3,F1) = (V7,F2);
(V4,F1) = (V8,F2);
```
(3)

You will observe that these restrictions are not inconsistent with the data. The main results of this second run are shown in the model in Table 4, and the fit measures for the two models are:

First run	Second run
χ^2=12.232	χ^2=15.074
df=13 P=0.487	df= 16 P=0.519

(4)

	χ^2	f
New model	15.074	16
Old model	12.232	13
Difference	2.842	3

(5)

We compare the two models as follows:

The difference in χ^2 is asymptotically distributed as χ^2 with three degrees of freedom. As a value of 2.842 with three degrees of freedom is not statistically significant according to traditional criteria, we prefer the new and simpler model.

This so-called χ^2-difference test is restricted to cases where one of the two models is nested under the other – that is, one model can be obtained by placing restrictions on the other.

Table 4 Example 1: selected output, second run

```
 MAXIMUM LIKELIHOOD SOLUTION (NORMAL DISTRIBUTION THEORY)

 PARAMETER ESTIMATES (B) WITH STANDARD ERRORS AND TEST STATISTICS (Z)
 STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                                                                R-
DEP.VAR.       PREDICTOR    B       BETA    S.E.     Z       SQUARED
--------------------------------------------------------------------
V1   (PRESS60 )                                                 .713
          F1   ( DEM60  ) 1.000     .845
          E1   (PRESS60 ) 1.000     .535
V2   (OPP60    )                                                .479
          F1   ( DEM60  ) 1.213*    .692   .144    8.423@
          E2   (OPP60   ) 1.000     .722
```

(Continued)

Table 4 (Continued)

V3 (FAIR60)						.582
	F1 (DEM60)	1.210*	.763	.126	9.619@	
	E3 (FAIR60)	1.000	.647			
V4 (ELECT60)						.704
	F1 (DEM60)	1.273*	.839	.123	10.378@	
	E4 (ELECT60)	1.000	.544			
V5 (PRESS65)						.644
	F2 (DEM65)	1.000	.802			
	E5 (PRESS65)	1.000	.597			
V6 (OPP65)						.581
	F2 (DEM65)	1.213*	.762	.144	8.423@	
	E6 (OPP65)	1.000	.647			
V7 (FAIR65)						.667
	F2 (DEM65)	1.210*	.817	.126	9.619@	
	E7 (FAIR65)	1.000	.577			
V8 (ELECT65)						.695
	F2 (DEM65)	1.273*	.833	.123	10.378@	
	E8 (ELECT65)	1.000	.553			

VARIANCES OF INDEPENDENT VARIABLES
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

			VARIANCE	S.E.	Z
FACTOR					
	F1	(DEM60)	4.770*	1.050	4.545@
	F2	(DEM65)	4.588*	1.043	4.398@
ERROR					
	E1	(PRESS60)	1.916*	.442	4.334@
	E2	(OPP60)	7.633*	1.392	5.485@
	E3	(FAIR60)	5.027*	.985	5.104@
	E4	(ELECT60)	3.256*	.737	4.416@
	E5	(PRESS65)	2.538*	.529	4.793@
	E6	(OPP65)	4.874*	.943	5.166@
	E7	(FAIR65)	3.348*	.714	4.690@
	E8	(ELECT65)	3.267*	.735	4.448@

COVARIANCES AMONG INDEPENDENT VARIABLES
STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

		COVA.	S.E.	Z	CORR.
F1,F2	(DEM60 ,DEM65)	4.520*	.987	4.578@	.966
E1,E5	(PRESS60 ,PRESS65)	.582*	.372	1.564	.264
E2,E4	(OPP60 ,ELECT60)	1.411*	.699	2.017@	.283
E2,E6	(OPP60 ,OPP65)	2.098*	.748	2.806@	.344
E3,E7	(FAIR60 ,FAIR65)	.741*	.624	1.188	.181
E4,E8	(ELECT60 ,ELECT65)	.480*	.462	1.039	.147
E6,E8	(OPP65 ,ELECT65)	1.275*	.595	2.143@	.319

In introducing these restrictions I had an eye on the output in Table 3, but my choice of restrictions was based on substantive reasoning. I was not just fishing in the output.

The revision of the original model was theory driven, but if model revision is more or less driven by empirical evidence the 'testing' is not meaningful, because you test on the same data that have formed the hypotheses. For example, if you drop non-significant parameters from a model you can be almost sure that the χ^2-difference test will also be non-significant.

A digression: if you cannot substantiate that the measuring instrument functions in the same way on the two occasions, you cannot compare the degree of democracy at the two points in time, if you should want to do so (more about this in Chapter 9).

Sample or population?

There is a problem with the analysis in Example 1: should we consider the 75 developing countries as a random sample drawn from a larger population (and, if so, how is that population defined?), or do they constitute the complete population? This is a fundamental question.

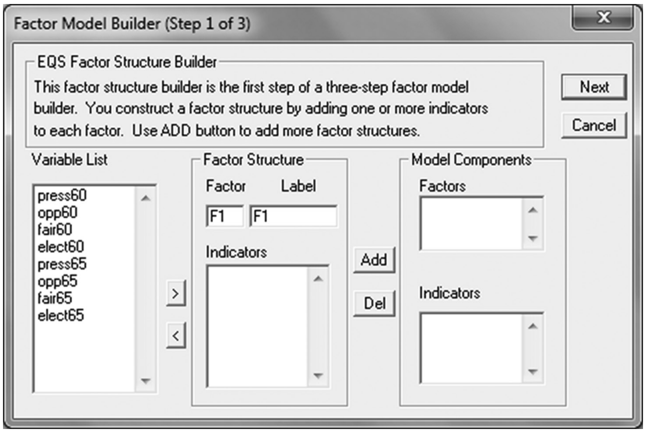
The basic idea in statistical estimation and testing is that you use a sample to draw conclusions about the unknown parameters of the population, taking the sampling error into consideration. But if your 'sample' is in reality the population, your 'parameter estimates' *are* the population parameters, and consequently there is no need for testing!

Even if the sample could be considered as drawn at random from some population, it is in all probability so large a fraction of that population that special formulas should be used to calculate the standard errors. Remember this if your analysis in based on macro data.

Programming using the 'Diagrammer'/'Factor Model'

After opening the data file (Bollen.ess), click the 'Diagrammer' icon and in the window in Figure 5.12 select 'Factor Model'. The window in Figure 3a will appear.

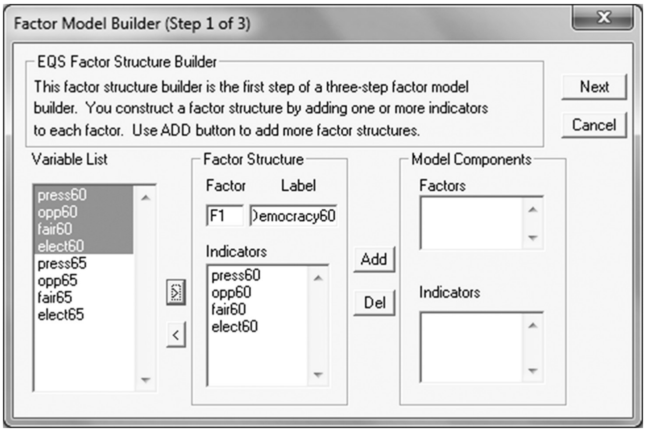
(a)



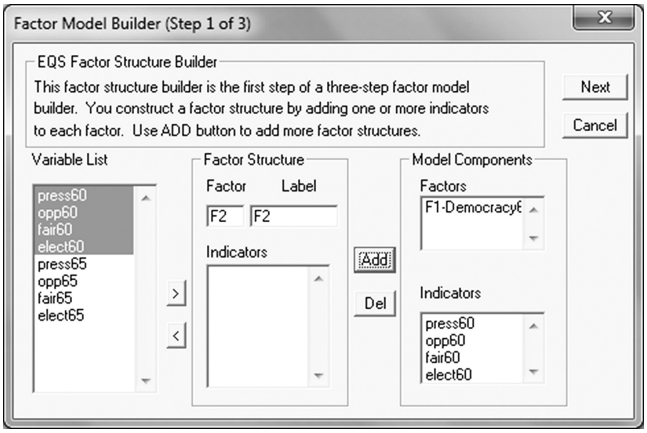
(Continued)

Figure 3 (Continued)

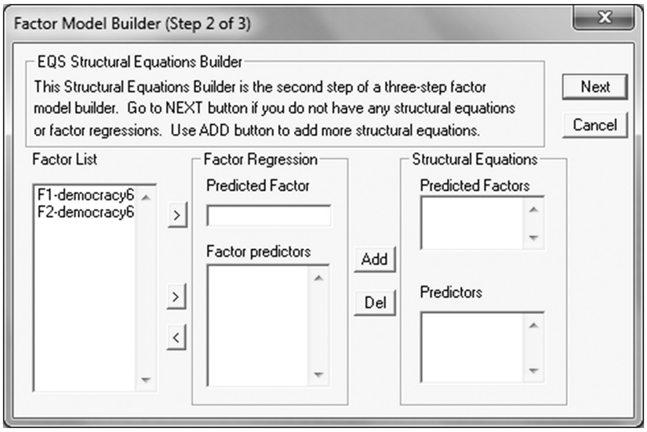
(b)



(c)



(d)



(e)

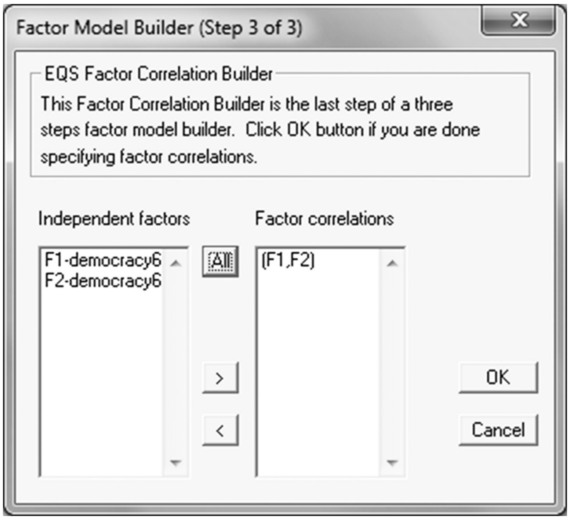


Figure 3 Example 1: the three steps in the 'Diagrammer/Factor Model

Label Factor 1 'democracy60'. Then select 'press60', 'opp60', 'fair60' and 'elect60' and by clicking the little arrow copy them to 'Indicators' as shown in panel (b). Then select 'Add', and the variables are moved to the 'Model Components' as shown in panel (c).

At the same time 'F2' appears in the 'Factor Structure' section of the window and you repeat the procedure, labeling F2 'democracy65' and using the last four indicators.

Pressing 'OK' moves you to the next step in the process – see panel (d). This gives you the possibility to include causal paths between factors, i.e. to build a general structural equation model like the ones treated in the next chapter. As such a connection is irrelevant in this case, just press 'OK' to go the third and last step.

The window in panel (e) appears. Select 'All' and when you click 'OK' the window in Figure 4 will appear – and you have finished your programming.

As you can see, the drawing in Figure 4 differs from the one in Figure 2, but feel free to experiment on your own with modifications to the drawing.

On the screen the arrows connecting the first indicator to its factor are red, signaling that the regression coefficient is fixed at 1.00 by default.

There is, however, a hidden problem here. You will remember that EQS allows only eight characters in labels, whereas 'Diagrammer' does not have such a restriction. This means that when you try to run the program shown in Figure 4, EQS will cut off all letters but the first eight, and you will receive an error message that F1 and F2 have the same label, so you have to rename the two factors before you run the program (e.g. call them dem60 and dem65 as in Table 2).

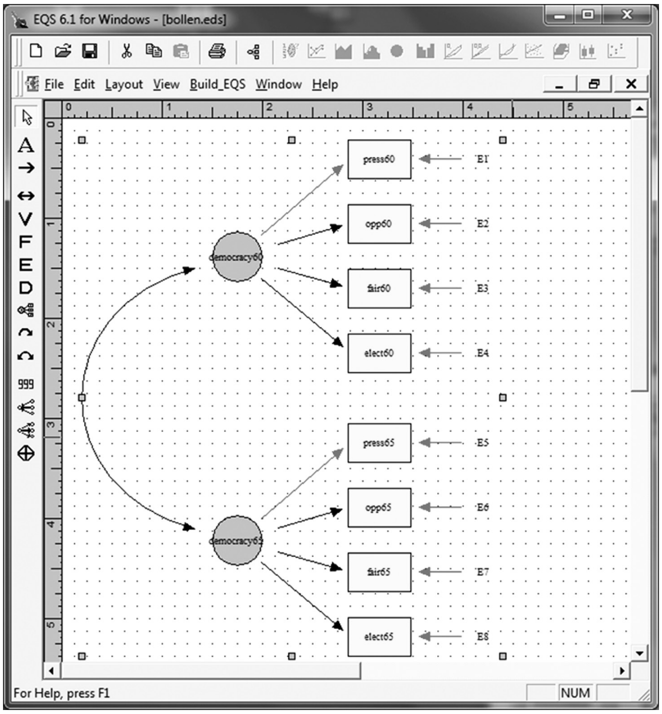


Figure 4 Example 1: the final model

3 Item Selection and Scale Construction

Very often confirmatory factor analysis will give a more differentiated picture of a scale's characteristics than traditional item analysis or exploratory factor analysis.

In Example 2.1 an item analysis in SPSS reduced the scale by three items and increased Cronbach's α from 0.831 to 0.866. In Example 3.3 we reached the same conclusion using principal components analysis.

These are the two classical techniques for item analysis, but they are not without drawbacks:

1. A large Cronbach's α and/or large item–rest correlations are no guarantee of unidimensionality.
2. In component analysis (and exploratory factor analysis) every manifest variable is connected to every latent variable, whereas in the ideal scale, every manifest variable is connected to only one factor. This blurs the picture (cf. Figure 1).
3. In both cases the possibilities for statistical testing are limited.

Example 2
Constructing a scale to measure 'style of processing

Childers, Houston, and Heckler (1985) constructed a summated scale for measuring 'style of processing' (SOP), i.e. a person's preference to engage in a verbal and/or visual modality of processing information about his or her environment.

You met this scale in Example 1.3.

The SOP scale is a 22-item, four-point scale, of which 11 items are assumed to reflect a verbal processing style and 11 a visual processing style. Childers et al. proposed the two sub-scales to be used either separately or in combination as one scale – although they preferred using the combined scale.

The items are shown in Table 5.

The scale was used by Sørensen (2001) in a project on consumer behavior, the respondents being 88 randomly selected Danish housewives. The scale used was a seven-point scale and not a four-point scale like the original one. You can find Sørensen's data on the companion website.

Table 5 Example 2: SOP scale

1. I enjoy work that requires the use of words.
2. There are some special times in my life that I like to relive by mentally 'picturing just how everything looked.*
3. I can never seem to find the right word when I need it.*
4. I do a lot of reading.
5. When I'm trying to learn something new, I'd rather watch a demonstration than read how to do it.*
6. I think I often use words in the wrong way.*
7. I enjoy learning new words.
8. I like to picture how I could fix up my apartment or a room if I could buy anything I wanted.*
9. I often make written notes to myself.
10. I like to daydream.*
11. I generally prefer to use a diagram rather than a written set of instructions.*
12. I like to 'doodle'.*
13. I find it helps to think in terms of mental pictures when doing many things.*
14. After I meet someone for the first time, I can usually remember what they look like, but not much about them.*
15. I like to think in synonyms of words.
16. When I have forgotten something I frequently try to form a mental 'picture' to remember it.*
17. I like learning new words.
18. I prefer to read instructions about how to do something rather than have someone show me.
19. I prefer activities that don't require a lot of reading.*
20. I seldom daydream.
21. I spend very little time trying to increase my vocabulary.*
22. My thinking often consists of mental 'pictures' or images.*

Notes: *Denotes items that are reverse scored. Items 1, 3, 4, 6, 7, 9, 15, 17, 18, 19, and 21 compose the verbal component. Items 2, 5, 8, 10 through 14, 16, 20, and 22 compose the visual component.

Cronbach's α was 0.76 for the verbal sub-scale and 0.71 for the visual sub-scale, while for the combined scale it was 0.74. In their original paper Childers et al. stated the same α to be 0.81, 0.86 and 0.88 respectively. In most circumstances, values of α of this size will lead to acceptance of the scales as reliable *and unidimensional.*

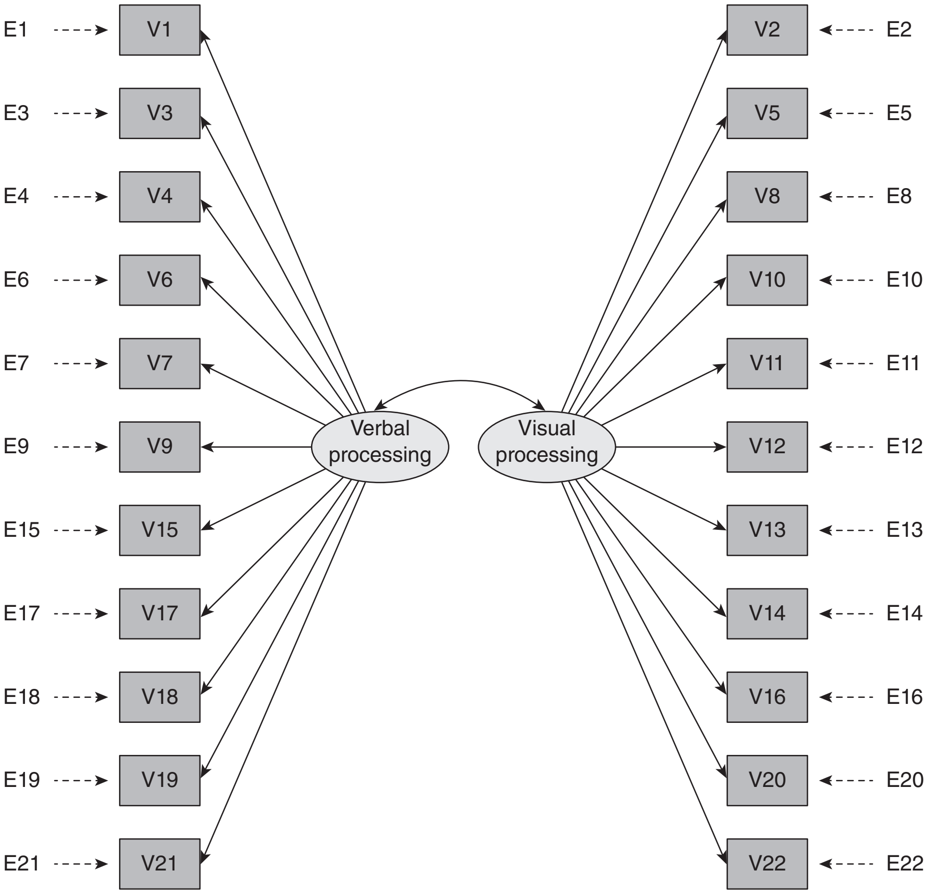


Figure 5 Example 2: model of Childers et al.'s SOP scale

Here I will use Sørensen's data in order to evaluate the scale; the measurement model is shown in Figure 5.

In the program shown in Table 6, you will meet one new paragraph:

```
/LMTEST
PROCESS=SIMULTANEOUS;
SET=PEE,GVF;
```

This deserves a few remarks, but you will have to wait till I comment on the output in Table 8.

But, first, let us first take a look at the fit measures in Table 7.

Table 6 Example 2: program for first run

```
/TITLE
  Example 7.2. First run. Data from Sørensen (2001)
/SPECIFICATIONS
  DATA='C:\Users\Sony\Desktop\EQS bog\EQS Chap 7\SOPX.ESS';
  VARIABLES=22; CASES=88;
  METHOD=GLS; ANALYSIS=COVARIANCE; MATRIX=COVARIANCE;
/LABELS
  F1=VERBAL PROCESSING; F2=VISUAL PROCESSING;
/EQUATIONS
  V1 = *F1 + E1;   V3 = *F1 + E3;   V4 = *F1 + E4;   V6 = *F1 + E6;
  V7 = *F1 + E7;   V9 = *F1 + E9;  V15 = *F1 + E15; V17 = *F1 + E17;
 V18 = *F1 + E18; V19 = *F1 + E19; V21 = *F1 + E21;

  V2 = *F2 + E2;   V5 = *F2 + E5;   V8 = *F2 + E8;  V10 = *F2 + E10;
 V11 = *F2 + E11; V12 = *F2 + E12; V13 = *F2 + E13; V14 = *F2 + E14;
 V16 = *F2 + E16; V20 = *F2 + E20;
 V22 = *F2 + E22;
/VARIANCES
  F1 = 1;
  F2 = 1;
  E1 TO E22 = *;
/COVARIANCES
  F2,F1   = *;
/LM TEST
  PROCESS=SIMULTANEOUS;
  SET=PEE, GVF;
/PRINT
  FIT=ALL;
  TABLE=COMPACT;
/END
```

Table 7 Example 2: fit indices for first run

```
CHI-SQUARE =       242.239 BASED ON      208 DEGREES OF FREEDOM
  PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       0.05189

  FIT INDICES
  -----------
  BENTLER-BONETT     NORMED FIT INDEX  =      0.203
  BENTLER-BONETT NON-NORMED FIT INDEX  =      0.479
  COMPARATIVE FIT INDEX (CFI)          =      0.531
  BOLLEN'S          (IFI) FIT INDEX    =      0.643
  MCDONALD'S        (MFI) FIT INDEX    =      0.823
  JORESKOG-SORBOM'S  GFI  FIT INDEX    =      0.747
  JORESKOG-SORBOM'S AGFI  FIT INDEX    =      0.692
  ROOT MEAN-SQUARE RESIDUAL (RMR)      =      0.899
  STANDARDIZED RMR                     =      0.230
  ROOT MEAN-SQUARE ERROR OF APPROXIMATION (RMSEA)     =      0.043
  90% CONFIDENCE INTERVAL OF RMSEA  (       0.000,           0.065)
```

A chi-square of 242.239 with 208 degrees of freedom giving a *P*-value of 0.052 is a little on the low side. What is more serious is that the fit indices are not quite satisfactory: the IFI and CFI indices below 0.80 should be taken seriously, so it is time to look at freeing parameters.

In the previous example you met the χ^2-difference test which was used to test if two models, one of which was nested under the other, were statistically different as far as fit was concerned.

You could of course use that same test here. For every fixed parameter (in this case parameters fixed at zero) – or groups of parameters – you would consider for 'unfixing', you could add them to the model and test whether the new model was significantly better than the old one.

That, however, would be a very tedious job, with a lot of model building and a lot of testing.

Fortunately EQS offers a way to do all the tests in one go without having to construct the various models you want to consider.

If you consider adding free parameters, use the Lagrange multiplier (LM) test, and if you think of fixing free parameters (not relevant in this case) use the Wald (W) test. The χ^2-difference, LM and W tests are asymptotically equivalent, i.e. in large samples they will give the same results.

The command

```
/LMTEST
PROCESS=SIMULTANEOUS;
```

orders EQS to do an LM test roughly equivalent to a series of χ^2-difference tests, one for each fixed parameter, starting with the one that is expected to contribute most to a better fit, and then proceeding with the second most significant etc.

If I had not inserted the command

```
SET=PEE,GVF;
```

the output would have printed out a list of every fixed parameter in the model – and that would have been a very long list indeed. Remember that all parameters that are not in the model are considered fixed (at zero) and, for most of them, including them in the model would make no sense, e.g. introducing causal paths between manifest variables. The `SET` command restricts the number of parameters that are considered for inclusion in the model.

To understand the meaning of this command, you must understand that EQS groups parameters in three groups (or matrices):

1. A group (matrix) called PHI and abbreviated P consists of covariances between independent variables.
2. A group called GAMMA and abbreviated G consists of regression coefficients involving both dependent and independent variables.
3. A group called BETA and abbreviated B that contains regression coefficients involving only dependent variables.

Dependent and independent are used here in the EQS sense. In the present case, only the modifications in the connections between V and F variables (i.e. VF parameters) and correlations between the Es (i.e. EE parameters) are worth considering.

The command `SET = PEE, GVF` tells EQS that the parameters to be considered are EE variables (found in the P matrix) and VF parameters (found in the G matrix).

If you use interactive programming, you will at one point choose ‘Build EQS/ LMTEST’ (see Figure 6a) to activate the window in panel (b), where you can see the three parameter matrices and their sub-matrices. When this window opens, several of the check boxes are checked by default, but just uncheck them and instead place your check marks as shown in the panel (c).

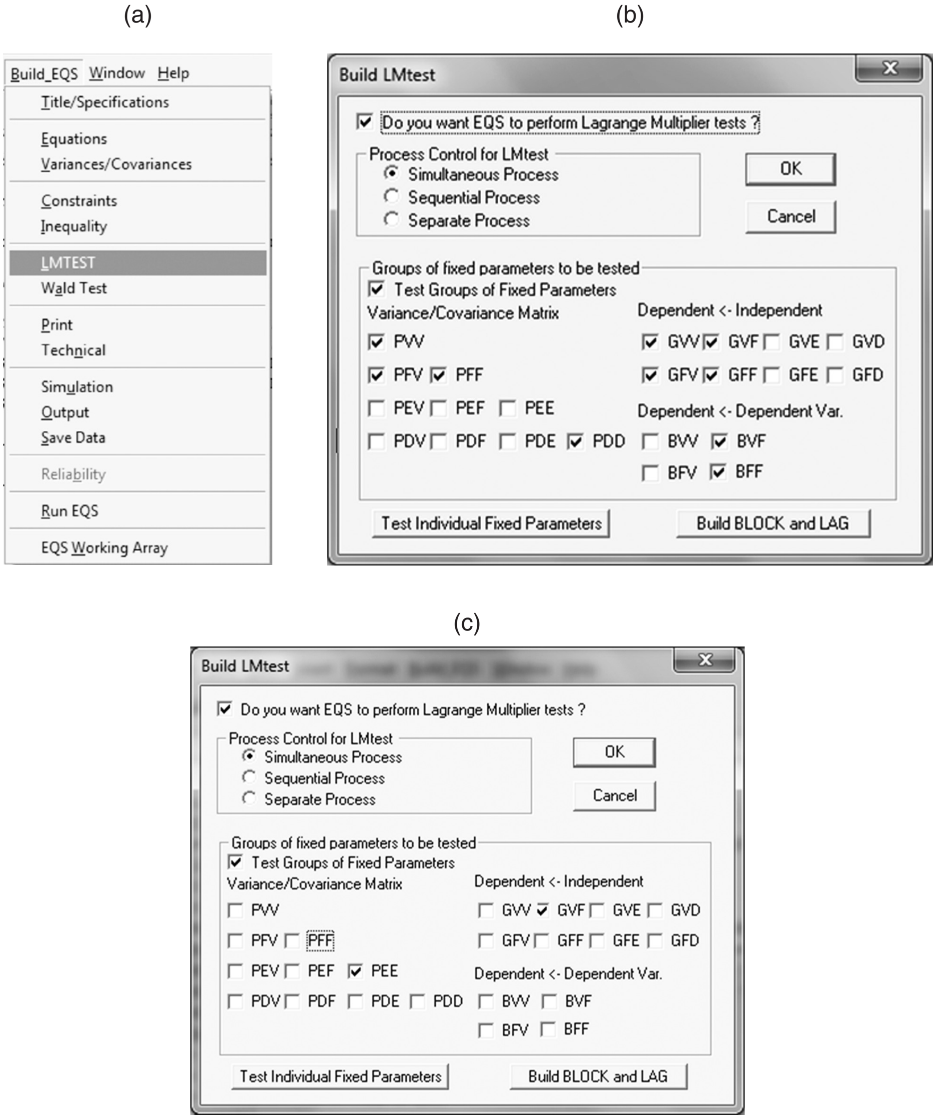


Figure 6 Programming ‘LMTEST’

From panel (b), you will guess that

```
PROCESS=SIMULTANEOUS;
```

is the default, so I could just as well have omitted it. The two alternatives

```
PROCESS=SEQUENTIAL;
```

and

```
PROCESS=SEPARATE;
```

give you the possibility to specify the sequence of the tests – for example, to specify a sequence in accordance with an a priori theory. You will have to consult the manual (Bentler, 2006) for further information.

Table 8 Example 2: output first run – Lagrange multiplier test

LAGRANGE MULTIPLIER TEST (FOR ADDING PARAMETERS)

ORDERED UNIVARIATE TEST STATISTICS:

NO	CODE		PARAMETER	CHI-SQUARE	PROB.	HANCOCK 208 DF PROB.	PARAM. CHANGE	STANDARDIZED CHANGE	PREDICTED RMSEA	CFI
1	2	6	E13,E11	9.146	0.002	1.000	0.820	0.580	0.038	0.642
2	2	6	E8,E6	7.163	0.007	1.000	-0.371	-0.521	0.039	0.615
3	2	6	E8,E3	6.908	0.009	1.000	0.487	0.584	0.040	0.612
4	2	6	E8,E7	6.884	0.009	1.000	-0.255	-0.572	0.040	0.611
5	2	6	E3,E2	6.884	0.009	1.000	-0.658	-0.578	0.040	0.611
6	2	6	E15,E5	5.972	0.015	1.000	-0.688	-0.393	0.040	0.599
7	2	6	E21,E6	5.673	0.017	1.000	0.454	0.434	0.041	0.595
8	2	6	E6,E3	5.554	0.018	1.000	0.517	0.635	0.041	0.593
9	2	6	E9,E4	5.530	0.019	1.000	-0.885	-0.370	0.041	0.593
10	2	6	E18,E5	5.382	0.020	1.000	0.670	0.565	0.041	0.591
-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-	-	-	-
253	2	6	E16,E2	0.000	0.993	1.000	-0.002	-0.001	0.044	0.517
254	2	0	F2,F2	0.000	1.000	1.000	0.000	0.000	0.044	0.517
255	2	0	F1,F1	0.000	1.000	1.000	0.000	0.000	0.044	0.517

MULTIVARIATE LAGRANGE MULTIPLIER TEST BY SIMULTANEOUS PROCESS IN STAGE 1

PARAMETER SETS (SUBMATRICES) ACTIVE AT THIS STAGE ARE:
PEE GVF

		CUMULATIVE MULTIVARIATE STATISTICS			UNIVARIATE INCREMENT					
							HANCOCK'S SEQUENTIAL			
STEP	PARAMETER	CHI-SQUARE	D.F.	PROB.	CHI-SQUARE	PROB.	D.F.	PROB.	PREDICTED RMSEA	CFI
1	E13,E11	9.146	1	0.002	9.146	0.002	208	1.000	0.038	0.642
2	E8,E6	16.389	2	0.000	7.242	0.007	207	1.000	0.033	0.728

3	E8,E7	23.177	3	0.000	6.788	0.009	206	1.000	0.028	0.807
4	E12,E8	28.917	4	0.000	5.740	0.017	205	1.000	0.023	0.872
5	E19,E10	34.706	5	0.000	5.789	0.016	204	1.000	0.016	0.938
6	E9,E4	40.215	6	0.000	5.509	0.019	203	1.000	0.001	1.000
7	E3,E2	45.211	7	0.000	4.996	0.025	202	1.000	99.999	1.000
8	E21,E6	50.094	8	0.000	4.883	0.027	201	1.000	99.999	1.000
9	E15,E5	55.044	9	0.000	4.950	0.026	200	1.000	99.999	1.000
10	E2,E1	59.898	10	0.000	4.854	0.028	199	1.000	99.999	1.000
11	E6,E3	64.340	11	0.000	4.443	0.035	198	1.000	99.999	1.000
12	E19,E6	68.959	12	0.000	4.619	0.032	197	1.000	99.999	1.000

```
*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.
```

You do not need to bother with the first two columns in Table 8 as they are only technical codes. In the third column you will find a list of fixed parameters (in this case they are all fixed at zero) sorted according to the size of increase in the model fit they would cause if the were freely estimated. The fourth and fifth columns show the chi-square and *P*-value of a test of the null hypothesis that the model fit would *not* increase if the parameter in question were freely estimated.

As all the *P*-values in the first ten cases are less than 0.05 you would reject this null hypothesis in (at least) the first ten cases based on traditional criteria, but the problem with such a decision is that the various tests are not independent. They do not tell you what would happen if more than one parameter were set free simultaneously.

The next column shows the Hancock *P*-values, which are some sort of Bonferroni probabilities analogous to what you have probably met in your introductory statistics course. As you will guess from the *P*-values of 1.000, this criterion is much too conservative, and I will not recommend basing any conclusions on it.

The next two columns show the expected change in parameter values if the parameter in case were set free. At present these parameters are not in the model, i.e. they are fixed at zero. Consequently the expected parameter changes are the expected parameter values.

Because of the interdependencies among the marginal LM tests, it is safer to base decisions on the multivariate tests in the second panel in the table.

The first line shows the consequences of freeing the parameter (`E13,E11`), the next the consequences of simultaneously freeing (`E13,E11`) and (`E8,E6`), the third the consequences of freeing (`E13,E11`), (`E8,E6`) and (`E8,E7`), etc.

In the third, fourth and fifth columns you will see the chi-square, df and *P* for simultaneously freeing the parameters, mentioned in the line and the lines above. The next two columns show marginal tests of freeing a parameter if all parameters in the lines above have also been freed.

It is tempting to free parameters, starting from the top until the marginal test becomes insignificant, but do not yield to the temptation.

As I advocated earlier, it is usually a good idea to base suggestions for modifications to a model on substantive rather than purely empirical evidence, so let us take a look at the items in Table 5.

A few of them catch the eye:

1. Items 10 and 20 say nearly the same thing and serve as mutual controls on the consistency of the answers, so they should correlate more than by their common cause: ‘visual’.

2. The same can be said of items 7 and 17.
3. Items 3 and 6 are also very close. If you cannot find a special word when you need it and therefore use another word, I think you will find that other word less suitable in the situation.

At least you should consider introducing these correlations in the model.

4. You should also expect correlations among items 5, 11 and 18. Whereas most of the other items only mention one of the processing styles, these three explicitly mention both, and compare them.

   According to their wording these three items could be placed in both subscales. You should therefore consider letting them load on both factors.

A program along these lines is given in Table 9 and the output in Table 10.

If you compare Table 10 with Table 7, you will observe that I have omitted the lines with E variables to save space. This will be done in all output tables from now on.

Table 9 Example 2: program for the second run

```
/TITLE
  Example 7.2. Second run. Data from Sørensen (2001)
/SPECIFICATIONS
  DATA='C:\Users\Sony\Desktop\EQS bog\EQS Chap 7\SOPX.ESS';
  VARIABLES=22; CASES=88;
  METHOD=GLS; ANALYSIS=COVARIANCE; MATRIX=COVARIANCE;
/LABELS
  F1=VERBAL PROCESSING; F2=VISUAL PROCESSING;
/EQUATIONS
  V1  =  *F1 + E1;   V3 = *F1 + E3;   V4 = *F1 + E4;   V6 = *F1 + E6;
  V7  =  *F1 + E7;   V9 = *F1 + E9;  V15 = *F1 + E15; V17 = *F1 + E17;
  V18 =  *F1 + *F2 + E18;            V19 = *F1 + E19; V21 = *F1 + E21;

  V2  =  *F2 + E2;   V5 = *F1 + *F2 + E5;              V8 = *F2 + E8;
  V10 =  *F2 + E10; V11 =  *F1 + *F2 + E11;           V12 = *F2 + E12;
  V13 =  *F2 + E13; V14 = *F2 + E14;  V16 = *F2 + E16; V20 = *F2 + E20;
  V22 =  *F2 + E22;
/VARIANCES
  F1 = 1;
  F2 = 1;
  E1 TO E22 = *;
/COVARIANCES
  F2,F1   = *;
  E10,E20 = *;
  E7,E17  = *;
  E3,E6   = *;
/LM TEST;
  PROCESS=SIMULTANEOUS;
  SET=PEE, GVF;
/PRINT
  FIT=ALL;
  TABLE=COMPACT;
/END
```

Table 10 Example 2: selected output (second run) (E lines removed to save space)

```
CHI-SQUARE =       216.290 BASED ON      202 DEGREES OF FREEDOM
  PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS        .23235

  PARAMETER ESTIMATES (B) WITH STANDARD ERRORS AND TEST STATISTICS (Z)
  STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                                                             R-
DEP.VAR.       PREDICTOR  B      BETA   S.E.   Z     SQUARED
------------------------------------------------------------
V1  (   V1     )                                        .276
            F1 (VERBALPR) .543*    .526  .177  3.075@
V2  (   V2     )                                        .367
            F2 (VISUALPR)  .807*   .606  .198  4.067@
V3  (   V3     )                                        .286
            F1 (VERBALPR)  .760*   .535  .204  3.735@
V4  (   V4     )                                        .045
            F1 (VERBALPR)  .294*   .212  .219  1.345
V5  (   V5     )                                        .598          (1)
            F1 (VERBALPR) 1.731*   .901  .349  4.956@
            F2 (VISUALPR) -.417*  -.217  .373 -1.117
V6  (   V6     )                                        .245
            F1 (VERBALPR)  .559*   .495  .169  3.299@
V7  (   V7     )                                        .097
            F1 (VERBALPR)  .256*   .311  .137  1.863
V8  (   V8     )                                        .551
            F2 (VISUALPR)  .874*   .742  .161  5.436@
V9  (   V9     )                                        .173
            F1 (VERBALPR)  .779*   .416  .267  2.920@
V10 (   V10    )                                        .135
            F2 (VISUALPR)  .561*   .367  .211  2.656@
V11 (   V11    )                                        .386          (1)
            F1 (VERBALPR) 1.243*   .691  .317  3.915@
            F2 (VISUALPR) -.204*  -.113  .338  -.602
V12 (   V12    )                                        .036
            F2 (VISUALPR)  .323*   .189  .244  1.324
V13 (   V13    )                                        .630
            F2 (VISUALPR) 1.280*   .794  .218  5.884@
V14 (   V14    )                                        .137
            F2 (VISUALPR)  .698*   .370  .233  3.000@
V15 (   V15    )                                        .055          (2)
            F1 (VERBALPR)  .341*   .235  .236  1.442
V16 (   V16    )                                        .204
            F2 (VISUALPR)  .680*   .452  .213  3.185@
V17 (   V17    )                                        .079          (2)
            F1 (VERBALPR)  .227*   .282  .136  1.665
V18 (   V18    )                                        .722          (1)
            F1 (VERBALPR) 1.833*  1.124  .377  4.861@
            F2 (VISUALPR) -.977*  -.599  .413 -2.363@
V19 (   V19    )                                        .536
            F1 (VERBALPR) 1.264*   .732  .203  6.218@
V20 (   V20    )                                        .173
            F2 (VISUALPR)  .647*   .416  .229  2.827@
```

(Continued)

Table 10 (Continued)

```
V21 (  V21    )                                                .010          (2)
          F1  (VERBALPR)  .125*   .099   .210    .596
V22 (  V22    )                                                .591
          F2  (VISUALPR) 1.284*   .769   .198   6.486@

  COVARIANCES AMONG INDEPENDENT VARIABLES
  STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.


                                 COVA.  S.E.     Z       CORR.
 ---------------------------------------------------------------
   F1,F2    (VERBALPR,VISUALPR)  .669*   .126   5.308@    .669        (3)
   E3,E6    (  V3    ,  V6    )  .460*   .213   2.158@    .391
   E7,E17   (  V7    ,  V17   )  .353*   .106   3.330@    .585
  E10,E20   (  V10   ,  V20   ) 1.425*   .359   3.965@    .708
```

With a chi-square of 216.290 with 202 degrees of freedom giving a *P*-value of 0.233, the model seems satisfactory, and a look at the fit indices confirms this:

CFI	SRMR	RMSEA	LO 90	HI 90
0.804	0.197	0.029	0.000	0.055

Although CFI is not quite up to standard (> 0.90–0.95), it is considerably larger than in the first run, and the other indices are all fine.

By studying the parameters of the model, you can observe the following:

1. Item 18 has significant loadings on both factors whereas items 5 and 11, which were supposed to measure visual processing style, actually both have significant loadings on 'verbal' and non-significant loadings on 'visual'. However, all coefficients have the expected signs.
2. The following regression coefficients are non-significant (one-sided test, $\alpha = 0.05$): 15, 17, 21, 5, 11 and 12.
3. All the covariances are significant, and the correlation between the two factors is 0.669 (as the two factors are standardized, the correlation equals the covariance).

What, then, can we conclude?

1. The SOP scale has two dimensions as suggested by Childers et al. – in fact it has *more* than two dimensions showing up as highly significant correlations between items in the same main dimension.
2. Some of the items have very little loading on the factor they are supposed to measure. A reformulation or exclusion should be considered.
3. Although the two sub-scales correlate, there are in fact *two* dimensions, and treating them as one scale could – depending on the research question – cause problems, because it assumes that the two ways of processing are alternatives.
4. In fact it has long been known that verbal and visual processes take place in functionally separate cognitive systems and thus are not alternatives (Paivio, 1971).

From this last point you will learn that it is of the utmost importance to spend a good deal of time making clear the nature of the concepts you intend to measure. This preliminary step is all too often not given the necessary care.

Now, in light of point 4 above, you may wonder: 'Why do the two factors correlate at all?'

A possible explanation is that the correlation is caused by item 18 loading on both factors, and that the correlation would disappear if this item (together with items 5 and 11) were removed.

This line of reasoning immediately gives birth to another question: the SOP scale is intended to measure preference, but do all items really measure preference? It seems to me that items 3, 6 and 14 measure ability rather than preference. Of course preference and ability must be expected to correlate, because we generally prefer activities where we feel we have the largest potential – but is ability not exogenous, and should the arrows connecting items 3, 6 and 14 to their latent variables not point in the opposite direction?

These two last points have taught us an important lesson: that serious thinking about the subject area is much preferred to thoughtless dependence on computer output.

If we delete items 5, 11 and 18 together with the three 'ability questions' we get the output in Table 11.

Table 11 Example 2: output from third run

```
CHI-SQUARE =       108.183 BASED ON      101 DEGREES OF FREEDOM
  PROBABILITY VALUE FOR THE CHI-SQUARE STATISTIC IS       0.29440

PARAMETER ESTIMATES (B) WITH STANDARD ERRORS AND TEST STATISTICS (Z)
  STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.

                                                              R-
 DEP.VAR.        PREDICTOR  B      BETA     S.E.    Z      SQUARED
 ---------------------------------------------------------------
V1   (  V1    )                                               .552
            F1  (VERBALPR) 1.030*    .743  .177  5.812@
V2   (  V2    )                                               .171
            F2  (VISUALPR)  .610*    .413  .209  2.923@
V4   (  V4    )                                               .219
            F1  (VERBALPR)  .787*    .468  .232  3.397@
V7   (  V7    )                                               .595
            F1  (VERBALPR)  .851*    .771  .141  6.027@
V8   (  V8    )                                               .219
            F2  (VISUALPR)  .497*    .468  .164  3.028@
V9   (  V9    )                                               .039      (3)
            F1  (VERBALPR)  .392*    .198  .293  1.337
V10  (  V10   )                                               .079
            F2  (VISUALPR)  .441*    .281  .229  1.922
V12  (  V12   )                                               .019      (3)
            F2  (VISUALPR)  .244*    .137  .255   .958
V13  (  V13   )                                               .655
            F2  (VISUALPR) 1.501*    .809  .221  6.788@
V15  (  V15   )                                               .123
            F1  (VERBALPR)  .577*    .350  .242  2.383@
```

(Continued)

Table 11 (Continued)

```
V16 (  V16   )                                                   .225
           F2  (VISUALPR)  .800*   .474  .219  3.660@
V17 (  V17   )                                                   .493
           F1  (VERBALPR)  .770*   .702  .142  5.422@
V19 (  V19   )                                                   .132
           F1  (VERBALPR)  .564*   .364  .231  2.443@
V20 (  V20   )                                                   .224
           F2  (VISUALPR)  .926*   .473  .252  3.674@
V21 (  V21   )                                                   .223
           F1  (VERBALPR)  .744*   .473  .215  3.464@
V22 (  V22   )                                                   .559
           F2  (VISUALPR) 1.259*   .748  .198  6.367@

  COVARIANCES AMONG INDEPENDENT VARIABLES
  STATISTICS SIGNIFICANT AT THE 5% LEVEL ARE MARKED WITH @.


                                         COVA.  S.E.   Z       CORR.
 --------------------------------------------------------------------
   F1,F2   (VERBALPR,VISUALPR) -.064*    .163   -.394    -.064      (1)
   E7,E17  (  V7    ,  V17  )  .278*    .143   1.939     .506      (2)
  E10,E20  (  V10   ,  V20  ) 1.997*    .423   4.722@    .767      (2)
```

The fit of this model is not too bad:

CFI	SRMS	RMSEA	LO 90	HI 90
0.893	0.171	0.029	0.000	0.064

Of the three models, the third one is the best fitting. However, the most striking support for this model is found using the decision theoretic measures:

	Model 1	Model 2	Model 3
AIC	−174	−188	−94
CAIC	−897	−890	−445

Looking at the parameter estimates, we can see the following:

1. The correlation between the two factors disappeared, as it should according to theory.
2. The other two correlations are significant (remember: one-sided tests!).
3. Items 9 and 12 are not significant.

The purist would remove items 7 and 20 in order to get rid of the remaining correlations. However, as these correlations would probably be (nearly) constant and repeat themselves in future uses of the scale, they would not reduce its reliability. You could also consider removing or rephrasing the non-significant items.

If your research demands that you compare your study with others that used the SOP scale, it will perhaps be necessary for you to use the items as a summated scale (or scales). If that is not a demand, I would prefer to use the items un-summated. You could consider parceling if you need to keep down the number of parameters.

Example 1 (continued)
Democracy in developing countries

In the previous example you were introduced to the LM test. Now, I revert to Example 1 to show you how this test works in connection with constrained parameters.

If you add the paragraph

```
/LM TEST;
PROCESS=SIMULTANEOUS;
```

to the program for the second run, the output will include the lines in Table 12, which clearly shows that none of the restrictions are unwarranted. You should compare the content of this table with the result of the χ^2-difference test used earlier.

The last part of the section, testing the addition of parameters, is of course irrelevant in this case.

Table 12 Example 1: using the LM test to test constraints

```
LAGRANGE MULTIPLIER TEST (FOR RELEASING CONSTRAINTS)

  CONSTRAINTS TO BE RELEASED ARE:
          CONSTR:    1   (V2,F1)-(V6,F2)=0;
          CONSTR:    2   (V3,F1)-(V7,F2)=0;
          CONSTR:    3   (V4,F1)-(V8,F2)=0;

           UNIVARIATE TEST STATISTICS:
                                                     PARAM.
   NO     CONSTRAINT     CHI-SQUARE   PROBABILITY   CHANGE
   --     -----------    ----------   -----------   ------
    1     CONSTR:    1      0.361        0.548       0.095
    2     CONSTR:    2      2.402        0.121      -0.272
    3     CONSTR:    3      0.004        0.950       0.009


         CUMULATIVE MULTIVARIATE STATISTICS           UNIVARIATE INCREMENT
         ---------------------------------            -------------------

  STEP  PARAMETER    CHI-SQUARE  D.F.  PROBABILITY       CHI-SQUARE  PROBABILITY
  ---- -----------   ----------  ----  -----------       ----------  -----------
    1  CONSTR:    2     2.402     1       0.121              2.402       0.121
    2  CONSTR:    1     2.550     2       0.279              0.148       0.701
    3  CONSTR:    3     2.577     3       0.462              0.027       0.869


  LAGRANGE MULTIPLIER TEST (FOR ADDING PARAMETERS)

ORDERED UNIVARIATE TEST STATISTICS:
```

(Continued)

Table 12 (Continued)

PREDICTED NO	CODE	PARAMETER	CHI-SQUARE	PROB.	HANCOCK 16 DF PROB.	PARAM. CHANGE	STANDAR-DIZED CHANGE	RMSEA	CFI
--	----	---------	------	-----	-----	------	------	-----	-----
1	2 12	V3,F2	2.982	0.084	1.000	-0.305	-0.041	99.999	1.000
2	2 12	V7,F1	1.943	0.163	1.000	0.235	0.034	99.999	1.000
3	2 0	V1,F1	0.523	0.469	1.000	0.095	0.017	99.999	1.000
4	2 0	V5,F2	0.523	0.469	1.000	-0.095	-0.017	99.999	1.000
5	2 12	V2,F2	0.508	0.476	1.000	0.113	0.014	99.999	1.000
6	2 12	V1,F2	0.451	0.502	1.000	0.088	0.016	99.999	1.000
7	2 12	V5,F1	0.420	0.517	1.000	-0.083	-0.014	99.999	1.000
8	2 12	V6,F1	0.188	0.664	1.000	-0.066	-0.009	99.999	1.000
9	2 12	V8,F1	0.024	0.877	1.000	-0.021	-0.003	0.007	1.000
10	2 12	V4,F2	0.020	0.887	1.000	0.020	0.003	0.007	1.000

```
*** NOTE *** IF PREDICTED RMSEA COULD NOT BE CALCULATED, 99.999 IS PRINTED.
             IF PREDICTED  CFI  COULD NOT BE CALCULATED,  9.999 IS PRINTED.


***** NONE OF THE UNIVARIATE LAGRANGE MULTIPLIERS IS SIGNIFICANT,
***** THE MULTIVARIATE TEST PROCEDURE WILL NOT BE EXECUTED.
```

4 Reliability and Validity

The classical methods of judging the reliability and validity of a measuring instrument all have their shortcomings since none of them actually take the latent variables explicitly into account as part of the measurement model.

Therefore the theoretical definitions of reliability and validity as coefficients of determination when regressing the measurement on the theoretical constructs cannot be used as a basis for calculations. However, using SEM as the basis for judging reliability and validity should open up this possibility.

Reliability

In Equation (2.3a) the reliability coefficient was defined as the coefficient of determination when a measurement (an indicator, a manifest variable) is regressed on its latent variable(s). Using SEM, you can use this definition to calculate reliability coefficients.

In the various outputs in this chapter you will find these squared multiple correlations in the last column of the estimated regression coefficients.

It comes as no surprise that the two items with the smallest reliabilities are the ones that have non-significant regression coefficients. However, a few of the highly significant items have rather small reliabilities and should be discarded or reformulated.

Anyway, if apart from measuring SOP you also want to measure several other concepts in order to construct a ‘causal’ model, 16 (not to mention 22!) items are quite a lot. If the other latent variables in your model require a similar number of items your questionnaire will easily grow to a length that could cause messy data, because respondents might refuse to participate, not answer all the questions, or fill out the questionnaire more or less at random in order to get the job done as quickly as possible.

In their efforts to obtain perfection scale, constructors very often end up with scales that are too long for practical use. The SOP scale is perhaps an example of this.

This way of calculating a reliability coefficient has at least five advantages:

1. It is based on the very definition of reliability.
2. It is possible to calculate the reliability of every single item and not just a sum of them.
3. It can be used whether the measurements are parallel, tau-equivalent or congeneric.
4. It does not assume errors to be uncorrelated across items.
5. It can be used when an item is an indicator for more than just one latent variable.

If you have only cross-sectional data it is in general not possible to estimate the specific variance, which is then absorbed into the error variance. In this case you must either assume that the specific variance is zero or consider the estimated reliabilities to be lower bounds.

Reliability in EQS

You can order EQS to print a wide assortment of reliability coefficients, all building on different assumptions, but you must be careful to use only coefficients that make sense in the actual case – remember from Example 6.2 that EQS printed Cronbach's α although it had absolutely no meaning in the context.

Before we look at the complications involved in measuring reliability in a multi-factor model such as the SOP scale, let us revert to a simpler one-factor example.

Example 3
Reliability of a one-factor scale

As mentioned several times in Chapter 2, calculation of Cronbach's α presupposes that the scale in question is unidimensional. So let us open the data set 'Fish1' (used in Example 2.1) as an EQS data file and construct the program in Table 13.

The output is given in Table 14, on which I will give the following comments:

Table 13 Example 3: program for EQS calculation of reliability for a one-factor model

```
/TITLE
  Reliability Analysis using the dataset 'Fish1'
/SPECIFICATIONS
  DATA='c:\users\sony\documents\fish1.ess';
VARIABLES=17; CASES=89;
  METHOD=ML; ANALYSIS=COVARIANCE; MATRIX=RAW;
/LABELS
  V1=FORB35;  V2=FORB36;  V3=FORB37;  V4=FORB38;  V5=FORB39;
  V6=FORB40;  V7=FORB41;  V8=FORB42;  V9=TILB43;  V10=TILB44;
  V11=TILB45; V12=TILB46; V13=TILB47; V14=TILB48; V15=TILB49;
  V16=TILB50; V17=FORB36R;
/RELIABILITY
  SCALE=V1,V3,V4,V5,V6,V7,V8,V17;
/PRINT
  TABLE=EQUATION;
/END
```

Table 14 Example 3: EQS calculation of reliability for a one-factor model: output

```
RELIABILITY COEFFICIENTS
------------------------
CRONBACH'S ALPHA                        =      0.831                    (1)
RELIABILITY COEFFICIENT RHO             =      0.847                    (2)
MAXIMAL WEIGHTED INTERNAL CONSISTENCY RELIABILITY  =      0.886       (3)

MAXIMAL RELIABILITY CAN BE OBTAINED BY WEIGHTING THE VARIABLES AS FOLLOWS:
FORB35    FORB37    FORB38    FORB39    FORB40    FORB41    FORB42
 0.2146    0.0954    0.0793    0.3590    0.2055    0.0523    0.1812
FORB36R
 0.4440

STANDARDIZED FACTOR LOADINGS FOR THE FACTOR THAT GENERATES           (4)
MAXIMAL RELIABILITY FOR THE UNIT-WEIGHT COMPOSITE

BASED ON THE MODEL (RHO):
FORB35    FORB37    FORB38    FORB39    FORB40    FORB41    FORB42
 0.7344    0.4561    0.4001    0.7950    0.6833    0.2390    0.6936
FORB36R
 0.8490
```

1. First in the output is Cronbach's α (which of course has exactly the same value as in Table 2.2). As you will remember, classical test theory does not include latent variables, and as calculation of α uses only manifest variables, it does not build the theoretical definition of reliability as formulated in Equation (2.3a).

   Also remember that α builds on the restrictive assumption that the items are at least equivalent.
2. Next comes Raykov's ρ, which builds on the theoretical definition of reliability and is more general than α as it only demands the measurements to be congeneric; ρ is in most cases larger than α. Furthermore, ρ builds on the actual measurement model, which could have more than one factor. However, when you use the `/RELIABILITY` paragraph a one-factor model is assumed. ρ maintains the equal weighting of items.

   Raykov (1997) presents a fine discussion and a comparison of the two reliability coefficients and even shows an EQS program for calculating ρ in the one-factor case.
3. The next few lines show the reliability (ρ) that could be obtained if the items were optimally weighted and the optimal weights. As you will observe, the three items with the lowest weights are exactly the same as those that were discarded in our traditional item analyses in Examples 2.1 and 3.3.
4. Last in the output are the standardized factor loadings (regression coefficients) for a model with equal-weighted items.

Example 4 (continued from Example 2)
Reliability of a multi-factor scale

If you include the statement

```
RELIABILITY = YES
```

In the /PRINT paragraph in the program in Table 6, the output will include the reliability coefficients shown in Table 15.

Table 15 Example 2: Reliability coefficients in a many-factor model

```
RELIABILITY COEFFICIENTS
  ------------------------
  CRONBACH'S ALPHA                                                 =  0.745  (1)
  COEFFICIENT ALPHA FOR AN OPTIMAL SHORT SCALE                     =  0.862  (2)
  BASED ON THE FOLLOWING  2 VARIABLES
    V7         V17
  RELIABILITY COEFFICIENT RHO                                      =  0.540  (3)
  GREATEST LOWER BOUND RELIABILITY                                 =  0.921  (4)
  LI-BENTLER CORRECTED GREATEST LOWER BOUND RELIABILITY            =  0.901
  GLB RELIABILITY FOR AN OPTIMAL SHORT SCALE                       =  0.928  (5)
  BASED ON 19 VARIABLES, ALL EXCEPT:
    V9         V12        V14
  BENTLER'S DIMENSION-FREE LOWER BOUND RELIABILITY                 =  0.921  (6)
  LI-BENTLER CORRECTED DIMENSION-FREE LOWER BOUND RELIABILITY      =  0.901
  SHAPIRO'S LOWER BOUND RELIABILITY FOR A WEIGHTED COMPOSITE       =  0.947  (7)

  WEIGHTS THAT ACHIEVE SHAPIRO'S LOWER BOUND:
    V1         V2         V3         V4         V5         V6         V7
   0.2679     0.0401     0.3213     0.1406     0.3282     0.2637     0.2418
    V8         V9         V10        V11        V12        V13        V14
   0.2400     0.1928     0.0268     0.2368     0.1308     0.1459     0.1575
    V15        V16        V17        V18        V19        V20        V21
   0.2102     0.1348     0.2427     0.2744     0.3218     0.0553     0.1430
    V22
   0.1720

  STANDARDIZED FACTOR LOADINGS FOR THE FACTOR THAT GENERATES
  MAXIMAL RELIABILITY FOR THE UNIT-WEIGHT COMPOSITE

  BASED ON THE MODEL (RHO):
    V1         V2         V3         V4         V5         V6         V7
   0.2651     0.4266    -0.1454     0.1065     0.0909     0.0762     0.3562
    V8         V9         V10        V11        V12        V13        V14
   0.3425     0.0308     0.6107    -0.0265    -0.0897     0.2935     0.0940
    V15        V16        V17        V18        V19        V20        V21
   0.1512     0.1041     0.3651     0.0234     0.0969     0.6580     0.1724
    V22
   0.2782

  BASED ON THE GREATEST LOWER BOUND (GLB):
    V1         V2         V3         V4         V5         V6         V7
   0.4536     0.2801     0.5299     0.3045     0.5899     0.5632     0.4042
    V8         V9         V10        V11        V12        V13        V14
   0.4487     0.3019     0.2605     0.4932     0.2171     0.3797     0.2985
    V15        V16        V17        V18        V19        V20        V21
   0.2630     0.3624     0.3503     0.5012     0.6250     0.2677     0.2231
    V22
   0.4393
```

1. First in the list is *Cronbach's α*. In this case it is irrelevant, because it builds on a one-factor model and in addition demands that all regression coefficients and all error variances are equal and that all errors are uncorrelated. This is indeed a very unrealistic model in most cases. The two-factor structure in this case is reason enough to invalidate *α*.

2. The next two lines, which tell us that we could obtain a maximal α using only the two items V7 and V17, are of course irrelevant for the same reason.
3. Next is ρ, which takes into account the two-factor structure of the model. ρ also assumes unit weighting of the items and, in case of correlated error terms, these correlations are considered to be 'noise'.

   Perhaps this last remark deserves a few more words for clarification.

   Recall Equation (2.2b), which in a 'rougher' form could be formulated as follows:

   $$\text{variance of } X = \text{variance explained by model} + \text{unexplained variance} \tag{6}$$

   where X is a manifest variable, and we defined the reliability as

   $$\rho_{xx} = \frac{\text{variance explained by model}}{\text{variance of } X} \tag{7}$$

   In the general factor model (whether it has one or more factors), we have several Xs, i.e. we have a *vector* of Xs, and consequently a covariance *matrix* (the input covariance matrix), which is divided into an implied (or model) matrix and a residual (or error) matrix (cf. Chapter 1's Equation (20)):

   $$\widehat{\Sigma} = \widehat{\Sigma}_m + \widehat{\Sigma}_\varepsilon \tag{8}$$

   The reliability is calculated using a formula analogous to (7), but with covariance matrices substituted for variances. By considering error covariances as 'noise', error variances *and* error covariances are put into $\widehat{\Sigma}_\varepsilon$, and $\widehat{\Sigma}_m$ takes everything else.

   Whereas α assumes a one-factor structure and ρ builds on the actual model (in this case a two-factor structure), the following reliability measure builds on a many-factor structure with an unspecified number of factors and they consider all covariances to be 'true' – that is, all covariances are put into $\widehat{\Sigma}_m$.
4. *Greatest lower bound reliability* demands that all variances are non-negative, whereas *Bentler's dimension-free lower bound reliability* does not. However, they both have an upward bias that could be serious in small samples. The two Li-Bentler corrections allow for this.
5. Also shown in the output is the reliability of an optimal short GLB scale together with a description of which items are included in this scale.
6. *Shapiro's lower bound reliability for a weighted composite* (Shapiro, 1982) is based on weighting the items using the weights that maximize reliability, the weights also being shown in the output.
7. Last in the output are standardized factor loadings that maximize the reliability measures ρ and GLB (both of which presuppose unit weighted scores).

Validity

Calculation of a suitable measure for the validity is more complicated.

Let us assume that we have a manifest variable V, which is an indicator of more than one latent variable in our model. If as our starting point we take the validity of V to be the extent to which V is connected to each concept, F1, F2, … it is assumed to measure, then the most straightforward measures of validity are the regression coefficients (V1,F1), (V1,F2), … in raw or standardized form. Each of these measures has the same advantages and drawbacks that we know from traditional regression analysis. One advantage of the standardized coefficients is that they are independent of the measurement units. This is very important in the case of SEM, where the measurement scales of the variables and especially those of the latent variables are often more or less arbitrary. On the other hand, standardized coefficients depend on the variances in the populations, and if we wish to compare several groups from different populations this could be a problem.

Another way is to start with the reliability coefficient as the coefficient of determination when regressing the variable V on all variables that have a direct effect on it. This is of course the squared correlation coefficient as defined in the section above on reliability.

If we want the validity coefficient to measure that part of the variance in V attributable to, for example, F1, we must deduce from the squared multiple correlation coefficient the variation caused by all other factors influencing V.

We can then define the validity coefficient of V with regard to F1 as

$$U_{\mathrm{V.F1}} = R^2_{\mathrm{V}} - R^2_{\mathrm{V(F1)}} \tag{9}$$

where R^2_{v} is the coefficient of determination obtained by regressing V on all variables that have a direct influence on V, while $R^2_{\mathrm{V(F1)}}$ is the coefficient of determination obtained by regressing V on the same variables except F1. The symbols in (9) are taken from Bollen (1989), who proposed the measure, to designate *unique validity variance*.

Example 2 (continued)
Constructing a scale to measure 'style of processing'

Let us estimate the unique validity variance of *process18* with regard to *verbal* in the model for the second run.

We have

$$U_{\mathrm{V18}.verbal} = R^2_{\mathrm{V18}} - R^2_{\mathrm{V18}(verbal)} \tag{10}$$

R^2_{V18} can be read from Table 10, and $R^2_{\mathrm{V18}(verbal)}$ is obtained by regressing process18 only on 'visual':

$$U_{\mathrm{V18},\, verbal} = 0.722 - 0.064 = 0.658 \tag{11}$$

As pointed out several times, the concept of validity is much more problematic both to define and to measure than reliability. The reader is referred to Bollen (1989) and Raykov (2012) for further reading.

5 Reflexive and Formative Indicators

In a confirmatory factor model (and also in an exploratory factor model), the arrows point *from* the latent variable to its manifest indicators. As a consequence, indicators of the same latent variable must correlate (cf. Figure 9a).

It is a very common mistake among newcomers to SEM to overlook this simple fact, and to use indicators that are not *necessarily* correlated.

The classic example is that you want to measure a person's consumption of alcoholic beverages, and to that end you use a series of questions each measuring the consumption of one single beverage. That is, indicators such as:

Consumption of beer

Consumption of wine

Consumption of whisky

Consumption of cognac

etc.

While a (weighted) sum of these variables is a measure of total consumption of alcohol, there is no reason to believe that all these indicators should be correlated.

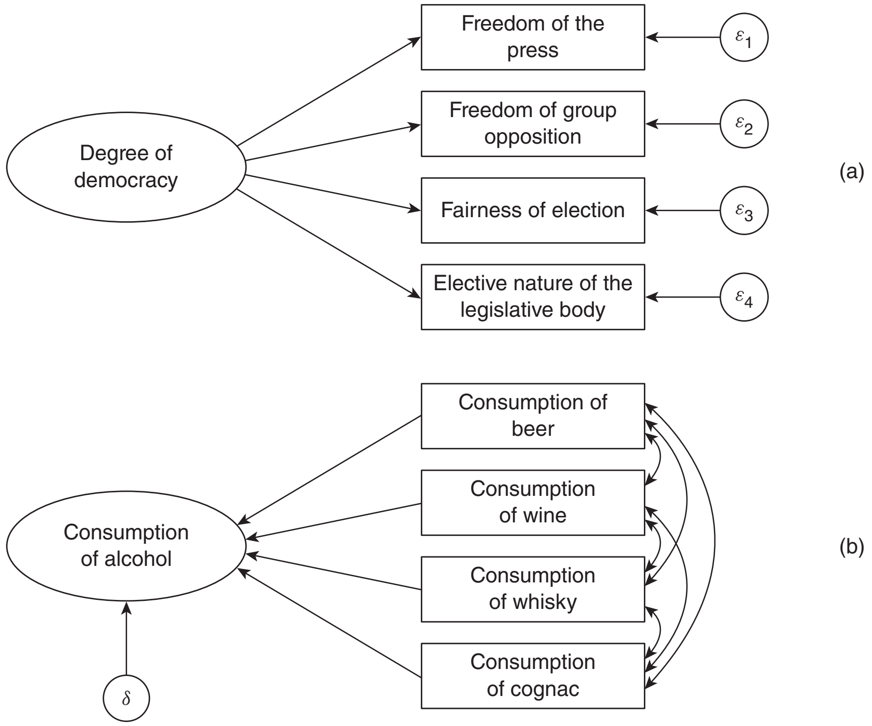


Figure 7 Reflexive (a) and formative (b) indicators

In a graphic illustration of this situation, the arrows should point *from* the indicators and *to* the latent variable (cf. Figure 9b).

The indicators in panel (a) of the figure are called *reflexive*: they *reflect* the underlying latent variable. In contrast the indicators in panel (b) are called *formative*: they *form* or define the latent variable, which in this case is not latent at all, as it is a function of manifest variables.

While a certain amount of correlation should exist among reflexive indicators for the same latent variable, correlations among formative indicators are not a necessity.

It is important to be aware that classical test theory and concepts like reliability and validity assume indicators to be reflexive. Using formative indicators in SEM programs like EQS is rather complicated. Apart from identification problems (that can be very tricky to solve) you can see from the figure that making the indicators exogenous means that they lose their error terms, and that measurement errors will instead be absorbed in the disturbance of the latent variable together with other sources of unexplained variance if the latent variable is affected by other variables in a larger model.

In fact, you have already met examples of indicators that could perhaps be considered formative, namely the three problematic items forb38, forb39 and forb41 in the first fish example (Example 2.1).

In many applications of SEM, socio-economic variables are used to characterize a person. The various indicators used are most realistically seen as formative, but it is not uncommon to see them treated as reflexive.

The many problems with using formative indicators in traditional SEM are given an excellent treatment by Kline (2006).

If your model includes several latent variables with formative indicators, you could try using another SEM technique, namely *partial least squares* (PLS) invented by Herman Wold (Wold, 1975). An introduction to this technique can be found in Fornell and Cha (1994). A deeper treatment with many examples is given by Vinzi, Chin, Henseler, and Wang (2010). The newest textbook is by Hair, Hult, and Ringle (2014).

You could say that PLS relates to covariance-based SEM like component analysis relates to factor analysis (Compare Figure 7 with Figure 1).

In this chapter you met the following concepts:

- confirmatory factor analysis
- three-indicator rule
- two-indicator rule
- R^2 as a measure of reliability
- unique variance as a measure of validity
- χ^2-difference test
- Lagrange multiplier test

You have also been introduced to the EQS paragraphs:

```
/CONSTRAINTS
```

and

```
/RELIABILITY
```

and to the various reliability measures in EQS.

Questions

1. State the differences among the three (main) factor models, and discuss their virtues and vices.
2. Can you suggest further modifications to the model in Figure 3?
3. Reflecting on your own studies or research, comment on the various instruments EQS offers for helping you with model modifications. Discuss their virtues and vices.
4. Comment on the differences between reflective and formative indicators. Why is this distinction important?

References

Bentler, P. M. (2006). *EQS 6 Structural equations program manual*. Encino, CA: Multivariate Software, Inc.

Bollen, K. A. (1989). *Structural equation modeling with latent variables*. New York: Wiley.

Bollen, K. A. (1979). Political democracy and the timing of development. *American Sociological Review, 44*, 572–587.

Bollen, K. A. (1980). Issues in the comparative measurement of political democracy. *American Sociological Review, 45*, 370–390.

Bollen, K. A. (1989). *Structural equation modeling with latent variables*. New York: Wiley.

Childers, T. L., Houston, M. J., & Heckler, S. (1985). Measurement of individual differences in visual versus verbal information processing. *Journal of Consumer Research, 12*, 124–134.

Fornell, C., & Cha, J. (1994). Partial least squares. In P. Bagozzi (Ed.), *Advanced methods of marketing research*. Oxford: Blackwell.

Hair, J. F., Hult, G. T. M., & Ringle, C. (2014). *A primer on least squares structural equation modeling*. Thousand Oaks, CA: Sage.

Kline, R. B. (2006). Reverse arrow dynamics: Formative measurement and feedback loops. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 43–68). Charlotte, NC: Information Age Publishing.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*, 173.

Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 472–492). New York: Guilford Press.

Shapiro, A. (1982). Weighted minimum trace factor analysis. *Psychometrika, 47*, 243–264.

Sørensen, E. (2001). *Means-end chains og laddering i et kognitivt perspektiv.* PhD thesis, Aarhus School of Business, Aarhus, Denmark.

Vinzi, V. E., Chin, W. W., Henseler, J., & Wang, H. (2010). *Handbook of partial least squares: Concepts, methods and applications in marketing and related fields*. New York: Springer.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock (Ed.), *Quantitative sociology: International perspectives on mathematical and statistical modeling*. New York: Academic Press.