

Paper 1651-2014
Confirmatory Factor Analysis and Structural Equation Modeling of
Noncognitive Assessments using PROC CALIS

Steven Holtzman, Educational Testing Service, Princeton, NJ

ABSTRACT

Noncognitive assessments, which measure constructs such as time management, goal-setting and personality, are becoming more prevalent today in research within the domains of academic performance and workforce readiness. Many instruments that are used for this purpose contain a large number of items that can each be assigned to specific facets of the larger construct. The factor structure of each instrument emerges from a mixture of psychological theory and empirical research, often by performing exploratory factor analysis (EFA) using the SAS[®] procedure PROC FACTOR. Once an initial model is established, it is important to perform confirmatory factor analysis (CFA) to confirm that the hypothesized model provides a good fit to the data. If outcome data are collected, such as grades, structural equation modeling (SEM) should also be employed to investigate how well the assessment predicts these measures. This paper will demonstrate how the SAS[®] procedure PROC CALIS is useful for performing CFA and SEM. Examples of these methods will be demonstrated and proper interpretation of the fit statistics and resulting output will be illustrated.

INTRODUCTION

Assessments that measure cognitive ability, such as mathematics, science and reading have been investigated for decades. Many researchers have recently been investigating ways of measuring nonacademic domains, utilizing numerous instruments which measure various personal qualities, such as time management, personality, teamwork, or social support. It has often been found that these measures can have strong relationships with a student's academic performance and workforce readiness.

Many noncognitive assessments require individuals to provide responses to a set of items. These items are typically self-report in nature, and tend to utilize Likert-type agreement response scales with three to seven choices ranging from Strongly Disagree to Strongly Agree, as an example. Although the items constituting a particular noncognitive assessment can be grouped under an overall construct (e.g., time management), it is common for subsets of items to describe specific facets within the larger construct (e.g., meeting deadlines, effective organization). The dimensionality of an assessment often emerges from previous research in psychological theory, where researchers determine the assignment of each item to its corresponding facet. In other cases where the categorization may not be so clear, the dimensionality is determined by empirical research, often by performing exploratory factor analysis (EFA). This process explores the possible underlying structure of a set of interrelated variables without imposing any preconceived structure on the data (Child, 1990). As a result, the number of latent variables and the underlying factor structure can be identified. This method can be performed using the SAS[®] procedure PROC FACTOR and has been the focus of many papers and presentations over the past years (Steinberg, 2010).

CONFIRMATORY FACTOR ANALYSIS

Whether the factor structure of a noncognitive instrument is determined using psychological theory or empirical research, it is important to perform confirmatory factor analysis (CFA), a special case of what is known as structural equation modeling (SEM). SEM typically refers to models where causal relationships are hypothesized to exist between latent variables. The CFA process determines whether the hypothesized structure provides a good fit to the data, or in other words, that a relationship between the observed variables and their underlying latent, or unobserved, constructs exist (Child, 1990). The CFA would also provide evidence that all items are properly aligned with the correct latent variables within the general construct being measured. A good example of a full SEM model would be a model that investigates whether a CFA model could successfully predict a particular outcome variable, typically also a latent variable.

There are numerous steps in performing a CFA. The first step is to determine the model that will be tested. Data should then be collected to test the model. If the model was proposed using EFA, either a completely different dataset should be used for the CFA, or the initial dataset should first be split randomly with different subsamples being used for each procedure (EFA and CFA, respectively).

Several things should be checked to ensure that the data are appropriate for testing the hypothesized model. In order to guarantee each factor is clearly represented by a sufficient number of items, known as overdetermination, it is important to look at the variable-to-factor ratio (Preacher & MacCallum, 2002). It is critical to have at least three items that are assigned to each factor (Anderson & Rubin, 1956), otherwise a factor is generally weak and unstable. However, it is often acceptable for a model to contain at most one such factor (Costello & Osborne, 2005). It is also important to make sure a large sample is available. There are numerous theories concerning how this can best be defined. One common rule is that there should be 10 people for every variable in the model (Everitt, 1975). Therefore, in order to run a CFA on 20 items, the data must be collected for at least 200 respondents. Once it is reasonable to conclude that model and sample size assumptions have been met, the data should be checked for missing data, univariate outliers, multivariate normality, and collinearity.

The CFA procedure can now be run using PROC CALIS (Covariance Analysis of Linear Structural Equations). This analysis will provide various results to both estimate the parameters in the model and assess model fit. The assessment of model fit is very important in CFA, as this provides evidence to validate the model, so careful interpretation of results is needed.

ASSESSING CFA FIT STATISTICS

When running CFA, many different fit statistics are used to help determine whether the model provides adequate fit for the data. The chi-square test indicates the amount of difference between expected and observed covariance matrices. A chi-square value close to zero and a chi-square *p*-value greater than 0.05 indicate that there is little difference, which is one indicator of good fit. However, the chi-square test is widely recognized to be problematic because it is very sensitive to sample size (Jöreskog, 1969). Therefore, it is often preferred to evaluate model fit based on other fit statistics.

The Root Mean Square Error of Approximation (RMSEA) is related to the residuals in the model. RMSEA values range from zero to one with a smaller RMSEA value indicating better model fit. Good model fit is typically indicated by an RMSEA value of 0.06 or less (Hu & Bentler, 1999), but a value of 0.08 or less is often considered acceptable (Browne & Cudeck, 1993). The Comparative Fit Index (CFI) is an incremental fit index, which assesses overall improvement of a proposed model over an independence model where the observed variables are uncorrelated (Byrne, 2006). CFI values range from zero to one with a larger value indicating better model fit. Acceptable model fit is indicated by a CFI value of 0.95 or greater (Byrne, 2006). The Normed Fit Index (NFI) and Nonnormed fit index (NNFI) are two other indicators that are commonly used to measure model fit (Bentler & Bonett, 1980). For each of these indicators, a larger value specifies better model fit and values above 0.90 are considered acceptable. The RMSEA, CFI, NFI and NNFI are four good indices to verify that a model is adequate.

If the fit statistics are acceptable, the parameter estimates can then be examined. The ratio of each parameter estimate to its standard error is distributed as a *t*-statistic and is significant at the 0.05 level if the value exceeds 1.96 and at the 0.01 level if the value exceeds 2.56 (Hoyle, 1995), for large samples. Since datasets used for CFAs are typically large and the *t*-distribution approaches the *z*-distribution as sample size increases, critical values from the *z*-distribution (1.96 and 2.56) can be used with large samples. For a good-fitting model, most or all parameter estimates should be significantly different from zero. If a parameter estimate is not significant, dropping the corresponding item from the model should be considered. Additionally, where applicable, correlations between the latent factors should be checked to see how the factors relate to each other. If correlations are sufficiently high (Bagozzi & Yi, 1988), consideration should be taken to define the model with fewer factors than originally hypothesized.

STRUCTURAL EQUATION MODELING

SEM models can also be performed by using the SAS[®] procedure PROC CALIS. In order to run a SEM model, all of the same steps should be taken as when running a CFA model.

ASSESSING SEM FIT STATISTICS

When running SEM, all of the same fit statistics should be considered that were discussed for the CFAs. In addition to these, the beta values and corresponding *t*-statistics for each path leading to an outcome variable should be investigated. Again for large samples, the *t*-statistic is significant at the 0.05 level if the value exceeds 1.96 and at the 0.01 level if the value exceeds 2.56. If this is found, it can be said that the respective latent factor is a significant predictor of the outcome variable.

PROC CALIS

The SAS PROC CALIS procedure estimates the parameters and test statistics for adequate fit for both CFA and SEM. The following example will illustrate both the syntax to run CFA and SEM and how to interpret the results that are presented in the corresponding output.

DESCRIPTION OF EXAMPLE USED

This example looks at the Multidimensional Scale of Perceived Social Support (MSPSS; Zimet, Powell, Farley, Werkmen, & Berkoff, 1990), a noncognitive assessment that measures perceived social support from family (e.g., “My family really tries to help me”), friends (e.g., “I can talk about my problems with my friends”), and significant others (e.g., “There is a special person who is around when I am in need”). This assessment consists of 12 items that are rated on a seven-point Likert scale from (1) “Very Strongly Disagree” to (7) “Very Strongly Agree.” Higher scores indicate higher levels of perceived support. Therefore, psychological theory suggests that a three-factor model would be a good fit for this assessment, with each factor representing the facets of family, friends and significant others, within the general construct of social support. The scale has been shown to have adequate reliability as Cronbach’s coefficient alpha statistics have been found to be 0.90, 0.94, and 0.95 for the family, friends, and significant others subscales respectively (Dahlem, Zimet, & Walker, 1991). In this example, the assessment was administered as part of a larger study to 591 college students. Fifty-nine percent of the sample was female and the median age of the students was 20.

CFA SYNTAX

CFA was run to confirm that the hypothesized three-factor model provides a good fit to the data. Syntax for the model run is shown in figure 1.

Figure 1: CFA Syntax

```
proc calis data=mspss;
lineqs
item1 = p1 family + e1,
item2 = p2 family + e2,
item3 = p3 family + e3,
item4 = p4 family + e4,
item5 = p5 friends + e5,
item6 = p6 friends + e6,
item7 = p7 friends + e7,
item8 = p8 friends + e8,
item9 = p9 f_sigoth + e9,
item10 = p10 f_sigoth + e10,
item11 = p11 f_sigoth + e11,
item12 = p12 f_sigoth + e12;
std
e1-e12 = vare1-vare12,
family=1,
friends=1,
f_sigoth =1;
cov
family friends = covflf2,
family f_sigoth = covflf3,
friends f_sigoth = covf2f3;
var
item1 item2 item3 item4
item5 item6 item7 item8
item9 item10 item11 item12
;
run;
```

The 12 lines of code under “lineqs” establish equations for the latent factors. Variable names for latent variables must begin with “f” and variable names for factor error terms must begin with “d” or “e.” The four lines of code under “std” assign a variance of one to each of the latent variables and allow the variances of the factor standard error terms to be freely estimated. The lines of code under “cov” allow the latent variables to covary and be freely estimated.

CFA RESULTS

The PROC CALIS CFA output contains many different pieces. The selections below highlight the most important sections of the results.

Figure 2: CFA Fit Statistics

The CALIS Procedure	
Covariance Structure Analysis: Maximum Likelihood Estimation	
...	
Chi-Square	211.9600
Chi-Square DF	51
Pr > Chi-Square	<.0001
...	
RMSEA Estimate	0.0731
...	
Bentler's Comparative Fit Index	0.9690
...	
Bentler & Bonett's Non-normed Index	0.9599
Bentler & Bonett's NFI	0.9597
...	

The Chi-square statistic is significant, which is shown by the corresponding p -value in figure 2. This is indicative of large differences between the observed and expected covariance matrices. Since the chi-square indicator is highly dependent on sample size, we should look at the other indices for guidance regarding the appropriateness of the model fit. The RMSEA of 0.0731 is greater than 0.06, but is less than 0.08. Therefore, this shows acceptable model fit. The NNFI and NFI (0.90 or larger) and the CFI (0.95 or larger) also all meet the criteria for acceptable fit. When acceptable model fit is found, the next step is to determine if all parameter estimates are significantly different from zero. The parameter estimates, also referred to as estimates for the manifest variable equations, are shown in figure 3.

Figure 3: CFA Parameter Estimates

Manifest Variable Equations with Estimates		
item1 =	0.7844*family	+ 1.0000 e1
Std Err	0.0359	p1
t Value	21.8641	
item2 =	0.8816*family	+ 1.0000 e2
Std Err	0.0339	p2
t Value	26.0111	
...		
item5 =	0.8275*friends	+ 1.0000 e5
Std Err	0.0345	p5
t Value	23.9731	
item6 =	0.8519*friends	+ 1.0000 e6
Std Err	0.0340	p6
t Value	25.0797	
...		
item9 =	0.8564*f_sigoth	+ 1.0000 e9
Std Err	0.0334	p9
t Value	25.6336	
item10 =	0.9240*f_sigoth	+ 1.0000 e10
Std Err	0.0318	p10
t Value	29.0782	
...		

The t -statistics shown are all greater than 2.56. Therefore, all parameters are significant at the 0.01 level. Variances of error terms, referred to as exogenous variables, appear in the output in figure 4.

Figure 4: CFA Variances of Error Terms

Variances of Exogenous Variables				
Variable	Parameter	Estimate	Standard Error	t Value
e1	vare1	0.38468	0.02796	13.76
e2	vare2	0.22282	0.02275	9.79
...				
e12	vare12	0.26992	0.01930	13.98

The t -statistics shown are all greater than 2.56. Therefore, all error variances are significantly different from zero at the 0.01 level. Covariances among latent variables, referred to as exogenous variables are shown in figure 5.

Figure 5: CFA Covariances Among Latent Variables

Covariances Among Exogenous Variables					
Var1	Var2	Parameter	Estimate	Standard Error	t Value
family	friends	covf1f2	0.51914	0.03483	14.91
family	f_sigoth	covf1f3	0.46550	0.03625	12.84
friends	f_sigoth	covf2f3	0.56740	0.03149	18.02

The t -statistics shown are all greater than 2.56. Therefore, covariances are significantly different from zero at the 0.01 level. Because the variances of the latent variables are fixed to be one, the correlations are equal to the covariances. Thus, the latent constructs are moderately correlated between 0.40 and 0.60.

SEM SYNTAX

SEM was run to investigate how well the three-factor MSPSS model predicts a student's GPA. Syntax for the model run is shown in figure 6.

Figure 6: SEM Syntax

```
proc calis data=mspss;
lineqs
gpa = a1 family + a2 friends + a3 sigoth + e13,
item1 = p1 family + e1,
item2 = p2 family + e2,
item3 = p3 family + e3,
item4 = p4 family + e4,
item5 = p5 friends + e5,
item6 = p6 friends + e6,
item7 = p7 friends + e7,
item8 = p8 friends + e8,
item9 = p9 f_sigoth + e9,
item10 = p10 f_sigoth + e10,
item11 = p11 f_sigoth + e11,
item12 = p12 f_sigoth + e12;
std
e1-e13 = vare1-vare13,
family =1,
friends =1,
f_sigoth =1;
cov
family friends = covf1f2,
family f_sigoth = covf1f3,
friends f_sigoth = covf2f3;
var
gpa
item1 item2 item3 item4
item5 item6 item7 item8
item9 item10 item11 item12
;
run;
```

For the SEM syntax, an extra line of code is added to the “lineqs” section where an equation is specified for the outcome variable, GPA. This line of code establishes the three latent variables as predictors of GPA and shows that the proposed model will investigate their relationships with GPA, represented by the a1, a2 and a3 parameters.

SEM RESULTS

The PROC CALIS SEM output is very similar to the CFA output. The following two figures highlight the portions that are unique to the full SEM output. Parameter Estimates, also referred to as estimates for the manifest variable equations, are shown in figure 7.

Figure 7: SEM Parmeter Estimates

Manifest Variable Equations with Estimates			
gpa	=	0.0495*family +	0.1105*friends + -0.0468*f_sigoth + 1.0000 e13
Std Err	0.0543 a1	0.0585 a2	0.0550 a3
t Value	0.9109	1.8881	-0.8513

None of the three t -statistics are significant at the 0.05 level. Therefore, none of the three latent constructs is a significant predictor of the outcome measure, student GPA. Variances of error terms, referred to as exogenous variables, are shown in figure 8.

Figure 8: SEM Variances

Variances of Exogenous Variables				
Variable	Parameter	Estimate	Standard Error	t Value
e13	vare13	0.98550	0.05762	17.10

The t -statistic is greater than 2.56. Therefore, the error variance for the outcome variable is significantly different from zero at the 0.01 level. All other results in the SEM output are similar to the ones shown in the CFA output. Although the inclusion of the outcome variable alters the other results, changes are minimal.

CONCLUSION

This paper has shown the power of PROC CALIS in analyzing a noncognitive assessment using CFA and SEM. These two statistical methods are effective in both confirming that a hypothesized model provides a good fit for collected data and investigating whether a set of nonobserved, or latent, variables can significantly predict an outcome measure. The procedures outlined in this paper show that PROC CALIS has the analytic capabilities to produce informative results.

The results of the CFA showed that the three-factor model for the MSPSS provides a good fit to the data. Therefore, it can be said that the general construct of social support can be described by the three specific facets of support provided by family, by friends, and by significant others. Additionally, the results of the SEM analysis showed that none of the three facets of social support are significant predictors of GPA for a sample of college students.

REFERENCES

- Anderson, T. W., & Rubin, H. (1956). Statistical inference in factor analysis. In J. Neyman (Ed.), *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (pp. 111-150). Berkeley, CA: University of California Press.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16, 74–94.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-600.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage Publications.

- Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum
- Child, D. (1990). *The essentials of factor analysis*, (2nd ed.). London, England: Cassel Academic.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7)
- Dahlem, N. W., Zimet, G. D., & Walker, R. R. (1991). The Multidimensional Scale of Perceived Social Support: A confirmation study. *Journal of Clinical Psychology*, 47, 756-761.
- Everitt, S. (1975). Multivariate analysis: The need for data, and other problems. *British Journal of Psychiatry*, 126, 237-240.
- Hoyle, R. H. (1995). The structural equation modeling approach: Basic concepts and fundamental issues. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 1-15). Thousand Oaks, CA: Sage Publications.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog, K.G. (1969), A general approach to confirmatory factor analysis, *Psychometrika*, 34, 183-202.
- Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. *Behavior Genetics*, 32, 153-161.
- Steinberg, J. (2010, November). *Exploring the dimensionality of large-scale standardized educational assessments using PROC FACTOR*. Paper presented at Northeast SAS Users Group (NESUG), Baltimore, MD.
- Zimet, G. D., Powell, S. S., Farley, G. K., Werkman, S., & Berkoff, K. A. (1990). Psychometric characteristics of the Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment*, 66, 610-617.

ACKNOWLEDGEMENTS

The author would like to thank Ted Blew, James Carlson, Bruce Kaplan, Ed Kulick, Jennifer Minsky, Jonathan Steinberg and Sailesh Vezzu for their support and assistance with this paper.

CONTACT INFORMATION

Your comments and questions are appreciated and encouraged. Please contact the first author at:

Steven Holtzman

Research Data Analyst

Data Analysis and Research Technologies

Educational Testing Service

Rosedale Road – Mail Stop 20-T

Princeton, NJ 08541

(609) 734-1593

sholtzman@ets.org

TRADEMARKS

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.