Journal Of Business & Economics
Vol. 4 No. 1 (Jan-June 2012) pp.62-88

Examining Data and Measurement Model Specification in SEM: An Illustration from Management Development

Ghulam Dastgeer*
Atiq ur Rehman **
Wali Rahman***

Abstract

Data analysis is highly critical for a value added research output but very tricky to handle by the researchers. Each statistical technique in research methodology has its own nuts and bolts that the researcher has to take care of. The purpose of present study is to present the most important aspects, issues and procedures to examine the characteristics of data and relationships of interest prior to Structural Equation Modeling technique. Through literature review the authors have noted some main issues and procedures in examination of data prior to a SEM analysis. Major issues discussed in the paper are model complexity, sample size, nature of data, and measurement model fit. An example in the field of Management Development (MD) is also presented to explain the procedure of data analysis in SEM. Findings of the research revealed that by devoting considerable time and effort on examining and exploring the nature of data and the relationships among variables, before the application of this technique, can help researchers in resolving procedural issues that eventually lead to better prediction and reliability of results. The present study contributes to literature on SEM by providing a more holistic view of data examination before SEM analysis and practical guidance for researchers to use SEM more effectively.

Key Words: SEM, CFA, Basic Assumptions, Management Development.

---

* Ghulam Dastgeer is an Assistant Professor at International Islamic University, Islamabad, Pakistan. Corresponding author email address: hellodastgeer@gmail.com
** Atiq ur Rehman, PIFRA, Department of Auditor General of Pakistan, Islamabad
*** Wali Rahman, National University of Modern Languages, Islamabad, Pakistan.

1. Introduction

Structural Equation Modeling (SEM) is one of the important multivariate techniques that simultaneously estimates and tests a series of hypothesized inter-related dependency relationships between a set of latent constructs (Reisinger and Mavondo, 2007). SEM has acquired hegemony among multivariate techniques and is the preeminent multivariate method of data analysis among the multivariate techniques, it has been, and continues to be, the technique that is undergoing the most refinements and extensions (Hershberger, 2003). Schumacker and Lomax as cited in Reisinger & Mavondo (2007, p. 42) state "SEM can be used to examine the nature and magnitude of postulated dependence relationships and at the same time assess the direct and indirect relations".

The importance of SEM has practically been recognized through its application in a number of disciplines, including psychology, sociology, economics, cross-cultural research, environmental studies, marketing, tourism and management studies (Reisinger & Mavondo, 2007). Researchers like Dastgeer and Rehman (2012), Bulut and Culha (2010), D'Netto et al. (2008), Garcia-Morales, Llorens-Montes and Verdu-Jover (2008), Chaiburu and Marinova (2006), Cheng (2001) and Tracey et al. (2001) used SEM in the field of management development (MD) and reported the benefits and effectiveness of SEM for research.

The choice and preference of a specific statistical technique is subject to the demand of data. However, once the choice is made the researcher is required to have basic knowledge of that particular technique. SEM cannot be an exception to that necessity. Broadly speaking, besides the basic jargon, it includes pre-SEM technical analysis, model specification, and analysis of a measurement model prior to measuring a structural model. By pre-SEM analysis, authors mean the examination of the characteristics of data and fulfillment of basic statistical assumptions prior to a SEM analysis. Data analysis may appear a time-consuming but very crucial step that helps researchers to get a basic understanding of the data and relationships between variables (Hair, Black, Babin, Anderson & Tatham, 2006). But this important step is overlooked by most of the researchers specially before using SEM techniques. For example, Reisinger and Mavondo (2007) reported that too often researchers use small sample sizes (less than 100) in SEM studies with no discussion of whether the sample is sufficiently large enough to run SEM or not. Similarly, in a large scale analysis using of SEM, Schreiber, Nora, Stage, Barlow, and King (2006) found that researchers provide no discussion concerning basic assumptions of data analysis like normality, outliers, linearity, or multi-collinearity.

Likewise, a researcher is required to specify a model before starting the analysis. In this specification a researcher is usually guided by a combination of theory and empirical results from the previous research (Hox & Bechger, 1998). Analysis of a measurement model means to look into the association between the variables (latent and observed). And before testing the hypothesized relationships among the constructs of the model, the measurement model must hold (Cheng, 2001; Andreson & Gerbing, 1988). Measurement model is tested for validation of the measurement instrument. That means missing any crucial step or wrong specification of a measurement model can lead to potentially catastrophic problems in subsequent process/analysis of a SEM model. After a comprehensive analysis of studies that used SEM analysis, Schreiber et al. (2006) complain that researchers are not fully aware of specification and estimations of a measurement model. In management and human resource (HR) literature, studies like Dastgeer and Rehman (2012), Bulut and Culha (2010), D'Netto, Bakas, and Bordia(2008), Garcia-Morales, Llorens-Montes and Verdu-Jover (2008), Chaiburu and Marinova (2006), and Tracey, Tannenbaum and Mathieu (2001) used SEM analysis but in the first stage of specification and estimation of a measurement model, one or more steps are missing in these studies.

The purpose of the present study is to investigate the critical aspects, issues and procedures to examine the characteristics of data and assessment of measurement model prior to SEM analysis. An extensive literature review was carried out by researchers to highlight the main issues and procedures in data examination and assessment of a measurement model prior to a SEM analysis.

2. Specification of Path and Measurement Model

Model specification process consists of all steps necessary to specify the relationships among the latent constructs and determine how the latent constructs are to be measured. This step is considered as most important, difficult and crucial because everything else follows from it.

2.1 Size of a Path Model

The first step in SEM is developing a theoretical model and converting the theoretical model into a path diagram of causal relationships among constructs/variables. A path diagram helps in depicting a series of causal relationships among variables. But a fundamental question remains: How many variables or constructs should be there in a path diagram? Hair et al. (2006) opine that although there is no specific theoretical limit on the number of variables to use in a model, however, researchers must balance the number of variables included in the model against the practical limitations of the SEM. Practical limitations mainly relate with interpretation of results. As the number of variables in a model increases the interpretation of results and statistical significance becomes difficult to achieve. However, arbitrary deletion of variables is not acceptable. Hair et al. (2006) warn the researchers not to omit a concept just because the number of variables is becoming large; rather researchers should go for parsimonious and concise theoretical models.

From now onward, an example will be used to illustrate how to examine the characteristics of data and estimation of a measurement model prior to SEM analysis. Figure 1 depicts a path diagram of model of management development (MD) effectiveness adopted from Dastgeer (2012). A detailed discussion on theoretical base and concepts of the model is beyond the scope of this paper. The model simply portrays the relationship among variables. There are total five constructs (denoted by rectangles) and six hypotheses (indicated by arrows) in the model. For example, it is hypothesized that line manager support is positively associated with individual initiative, opportunity for skill utilization, and program design. Further it is hypothesized that variables like individual initiative, opportunity for skill utilization, and program design have direct positive association with MD effectiveness.

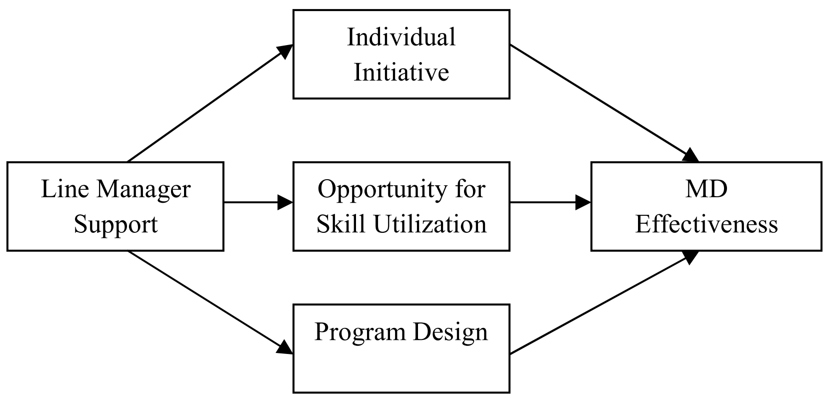

Fig.1 Path Diagram: Predictors of MD Effectiveness

2.2 Specification of a Measurement Model (Determining the Number of Indicators)

In SEM a path diagram is always demarcated in terms of constructs and researchers have to find variables/indicators to measure each construct. Hair et al. (2006) state that a construct can have a single indicator although a single indicator generally does not provide adequate representation of such constructs and hence, creates problem of estimation reliability. A better approach is to obtain multiple indicators of each construct. Each construct should have at least two pure variables/indicators but preferred minimum number of indicators is three (Resinger & Mavondo, 2006; Hair et al., 2006). Although there is no cut-off number of indicators per construct, researchers like Hair et al. (2006) have suggested that five to seven indicators should represent most constructs.

In figure 2, an example of a measurement model or Confirmatory Factor Analysis (CFA) is given that specify relationships between latent constructs and their indicators. The model is put to see if the model fits into the data. Figure 2 depicts that in measurement all latent constructs are connected by double headed curved arrows that represent correlations among constructs. Each of the construct has six to seven indicators/observed variables used to measure it.

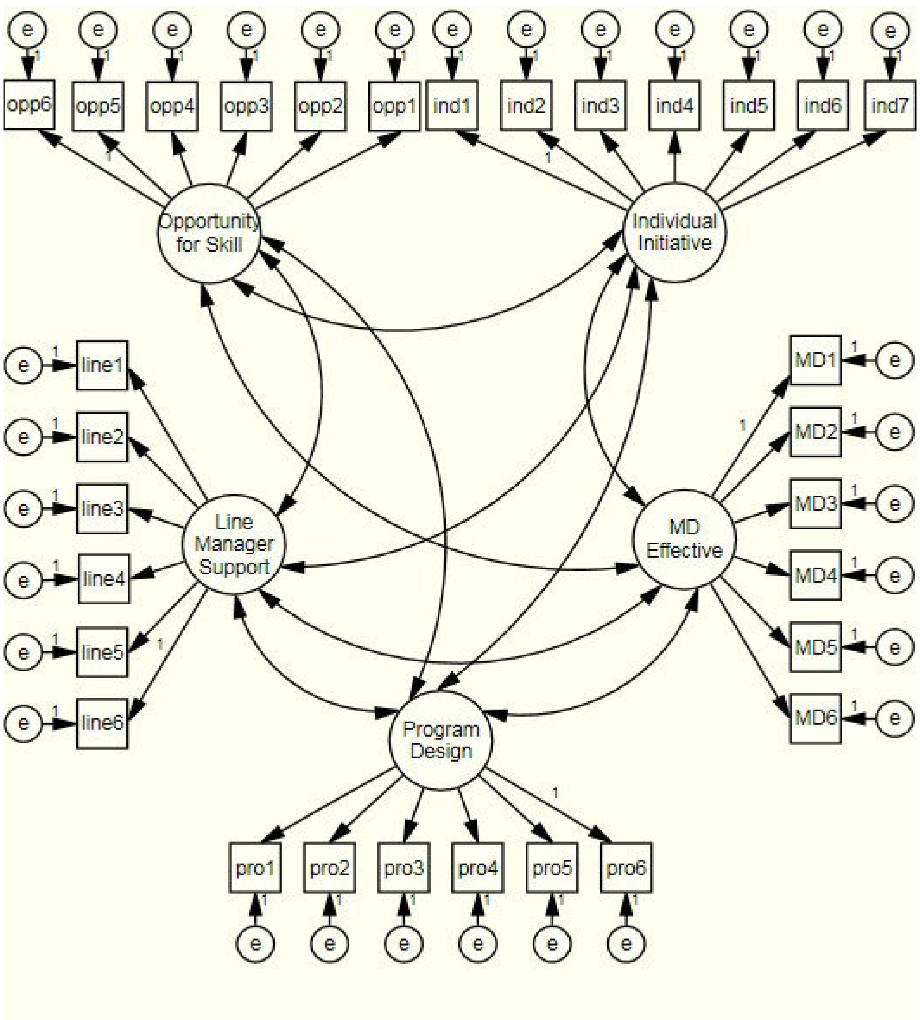

Fig. 2 Path Diagram: CFA of MD Effectiveness

3. Pre-SEM Technical Analysis

3.1 Sample Size

As in any other statistical technique, sample size plays a critical and important role in estimation and interpretation of results and estimation of sampling error (Hair et al., 2006).Reisinger and Mavondo (2007) stated that sample size is critical for achieving acceptable fit measure. Reisinger and Mavondo (2007), Schreiber et al. (2006) and Hair et al. (2006) argued that there is no standard requirement of sample size for SEM but absolute minimum sample size must be at least greater than the number of correlations in the input data matrix. Recommendation from researchers like Reisinger and Mavondo (2007), Schreiber et al. (2006) and Hair et al. (2006) is a minimum ratio of at least five respondents for each estimated parameter, with a ratio of ten respondents per parameter considered most appropriate. SEM analysis of small samples is almost certainly problematic. Sample size requirements increase as complexity of a model increases (Schreiber et al., 2006). It implies that a small sample may not be large enough to support the estimation of a more complex model. While using the most common estimation procedure Maximum Likelihood Estimation (MLE), Hair et al. (2006) recommended sample size ranging from 100 to 200 and sample size of 200 is considered appropriate and critical. The question is; is there any effect of large sample size? As researchers increase the sample size above this value (like 400 or more), the MLE method increases in its sensitivity to detect differences among the data and make indicators of goodness-of-fit measures a poor fit.

In CFA example of MD effectiveness model, the achieved sample size is 177 that fulfill the minimum requirements of SEM as recommended by Hair et al. (2006).

3.2 Statistical Assumptions and Outliers

Reisinger and Mavondo (2007) argued that like other multivariate data analysis techniques, in SEM there are few basic assumptions that need to be satisfied in order to ensure accurate inferences. Some major assumptions are independent observations, random sampling (except for longitudinal studies), linearity of all relationships, multivariate normality, discriminant validity of measures, no extreme cases (i.e. outliers), data measured on interval or ratio scale, and no skewness or kurtosis in the data (Reisinger and Mavondo, 2007; Hair et al., 2006). In the standard use of structural equation modeling, the observations are drawn from a continuous and multivariate normal population; a sufficiently large variation from normality can make all statistical tests invalid (Hair et al., 2006). Before using the data for SEM estimation, researchers should perform all of the diagnostic tests on the data.

Outliers can distort statistical analysis, hence, researchers are strongly recommended to identify any outlier in the data before it is used for SEM. "Outliers are observations with a unique combination of characteristics identifiable as distinctly different from other observations" (Hair et al., 2006).

In the example of MD effectiveness model, the assumptions of normality are evaluated through SPSS 17 using box plots. In total, seven questionnaires were treated as outliers because those were substantially different from the other observations; hence those questionnaires were dropped out.

3.3 Missing Data

Besides all other statistical assumptions, researchers are suggested to report a systematic discussion of handling missing data. In SEM missing data can have a thoughtful effect on calculating the input data matrix and estimation process (Carter, 2006; Hair et al., 2006). Because SEM requires a complete data set, there are numerous methods to deal with the missing data problem, for example:

i) Delete case: It involves deletion of incomplete cases from the dataset (Carter, 2006). That means a researcher deletes the variable/indicator from the data sheet that has record with missing data.

ii) Imputation: It involves placing expected values into the data set in the location of the missing data (Carter, 2006).

Hair et al. (2006) stated that there is no single best method to deal with the missing data problem because every method has its advantages and disadvantages, and researchers should try to use several techniques to assess stability of the results. In sum, SEM studies should indicate the extent to which there is missing data and should describe the technique used to handle missing data (Reisinger and Mavondo, 2007).

In the example of MD effectiveness model, the missing data through SPSS17 is evaluated. It was found that two respondents filled the questionnaires partially that led to the missing data problem. As the missing data was large, those questionnaires were removed from the data sheet to get stable results.

3.4 Reliability and Unidimentionality

3.4.1 Reliability

Sekaran (2006) defines reliability of a measure as “an indication of stability and consistency”. Most commonly method for testing reliability of research instruments is the internal consistency method which involves computation of Cronbach’s alpha. Internal consistency reliability confirms the consistency of respondents’ answers to all the items in a measure and the items are independent measures of the same concept (Sekaran, 2006). Cheng (2001) argued that an indicator/measure has to be deleted if it has extremely low internal consistency. The acceptable threshold of Cronbach’s alpha is 0.70 (Nunnally & Bernstein, 1994).

In the example of MD effectiveness model, Cronbach’s alpha statistics was used to check the reliability of the research instrument (all constructs).

Cronbach values for all constructs are given in table 1.

Table1
Cronbach Values for Constructs

Name of Variable	Cronbach Alpha Value	Number of Items
MD effectiveness	0.894	6
Line manager support	0.899	6
Individual Initiative	0.666	7
MD program design	0.891	6
Opportunity for Skill utilization	0.886	6

Table 1 show that the Cronbach value for all constructs, except individual initiative ranged from 0.83 to 0.92. Only the construct "Individual Initiative" had a value of 0.66 falling some what short of the recommended level. Henceforth, except the construct "Individual Initiative", Cronbach values for all constructs imply that all items of each construct are measuring the same content universe (i.e. construct). The construct Individual Initiative needs revision, but before revision of this particular construct, the unidimentionality of all other constructs is checked.

3.4.2 Unidimentionality

Unidimentionality is a concept similar to the concept of reliability. Hair et al. (2006) defined unidimentionality as "a characteristic of a set of indicators that has only one underlying trait or concept in common" (p. 584). Researchers should perform unidimentionality tests on all constructs that have multiple indicators before estimation of a SEM analysis (Hair et al., 2006). In order to test unidimensionality of a scale, usually, principal component factor analysis that generates eigen values is used. As a rule, eigen values should be greater than one to establish unidimensionality of a scale (Hoe, 2008).

To test the unidimentionality of a measurement instrument, in the example used in this study, at first, a rationale review of item contents was done by researchers to determine 'like items', as suggested by Hall et al. (1999). Secondly, principal component factor analysis was used to test for unidimentionality as suggested by Germain, Droge and Daugherty (1994). All constructs in the current study were separately subjected to principal component analysis and the eigen values are presented in Table 2.

Table 2
Eigen Values of Measures

Construct	Component	Total	Initial Eigen Values	
			% of Variance	Cumulative %
Management Development Effective	1	3.933	65.549	65.549
	2	0.723	12.051	77.600
	3	0.424	7.069	84.669
	4	0.379	6.312	90.981
	5	0.274	4.573	95.554
	6	0.267	4.446	100.000
Line Manager Support	1	4.051	66.912	66.912
	2	0.595	9.909	76.822
	3	0.571	9.511	86.333
	4	0.345	5.756	92.089
	5	0.304	5.073	97.162
	6	0.170	2.838	100.00
Individual Initiative	1	2.765	39.493	39.493
	2	1.153	16.474	55.968
	3	0.919	13.131	69.099
	4	0.751	10.734	79.833
	5	0.604	8.632	88.465
	6	0.452	6.460	94.925
	7	0.355	5.075	100.00
MD Program Design	1	3.889	64.812	64.812
	2	0.592	9.862	74.674
	3	0.574	9.573	84.248
	4	0.367	6.120	90.368
	5	0.319	5.313	95.681
	6	0.259	4.319	100.000
Opportunity of Skill Utilization	1	3.822	63.699	69.699
	2	0.552	9.197	72.896
	3	0.542	9.031	81.928
	4	0.407	6.778	88.706
	5	0.400	6.671	95.377
	6	0.277	4.623	100.00

Extraction Method: Principal Component Analysis

Table 2 depicts that except for Individual Initiative, all other constructs had only the first eigen value greater than 1. This provides support for the unidimensionality of these scales. For Individual Initiative, two eigen values were greater than 1. To enhance the unidimensionality of Individual Initiative, once again a rationale review of items was done, assessed the "Croncach's alpha of item deleted" and incremental modification was carried out to find out the threats. It was found that item "ind4"and item "ind5" are serious threats to the unidimensionality of this construct. These two items were deleted from the data sheet and principal components analysis was rerun on these two constructs to determine the eigen values.

Table 3
Revised Eigen Values of Individual Initiative

Construct	Component	Total	Initial Eigen Values	
			% of Variance	Cumulative%
Individual Initiative	1	2.698	53.958	53.958
	2	0.810	16.194	70.152
	3	0.635	12.709	82.861
	4	0.500	10.002	92.863
	5	0.57	7.137	100.000

Table 3 shows that only first eigen value was greater than 1 for the construct Individual Initiative. This provided support for the unidimensionality of these scales. After deletion of threats we rerun the reliability test on the Individual Initiative construct. Revised Cronbach Alpha value of Individual Initiative is 0.78 that falls under the recommended value of 0.70.

4. Measurement Model Estimation

Cheng (2001) suggests two different ways to evaluate a measurement model's validity. First is a test of the measure of each construct separately. Second is a test of all measures together at one time (p. 653). Cheng (2001) further suggests that the second method of evaluation of measurement model is better than the first one.

In our example, the second method of measurement model's validity is adopted as per Cheng's recommendation. LISREL software was used to estimate the model and constructs' correlations. For parameter estimation number of estimation methods available include maximum likelihood, weighted least squares, instrumental variables, generalized least squares, two-stage least squares, unweighted-least squares, ordinary least squares, and diagonally weighted least squares depending on the data and nature of the model (Reisinger & Mavondo, 2007). Maximum Likelihood Estimation (MLE) procedure is the most commonly used and accepted method for model estimation (Reisinger &Mavondo, 2007). MLE provides valid results even if the sample size is small (Hair et al., 2006). Procedurally, CFA of the combined measurement model was performed to validate the measures of latent constructs. In CFA, overall model fit portrays the degree to which the specified indicators represent the hypothesized constructs. All the indicators of latent constructs were loaded on their specific constructs and all constructs were inter-correlated. To ensure reliability of the indicators by CFA, it was confirmed that the factor loads are higher than 0.4 and significant ($t \geq 1.96$; $p \leq 0.05$), composite reliability of each whole scale, by applying the Cronbach alpha, is (≥ 0.7) and average variance extracted is (≥ 0.5) (Hair et al., 2006).

4.1 Determining Offending Estimates

While estimating a measurement model, first of all, researchers are required to examine the results for offending estimates. There are a number of estimated coefficients that need to fall in acceptable limits. According to Hair et al. (2006), offending estimates refer to any value that exceeds its theoretical limits. Before analyzing the hypothesized relationships among variables and interpreting the results for overall model fit, researchers need to correct the nonsensical or theoretically inconsistent estimates. Cheng (2001) suggests model modification if there is any indicator that does not measure its underlying construct or is not reliable. To modify a model, an offending indictor has to be deleted (Cheng, 2001).The most common examples of offending estimates are the following.

4.1.1 Correlation among constructs (Convergent validity)

Correlation between two variables/constructs is treated as an offending estimate if (in the standardized solution) it exceeds the value of 1.00 or even if the variables are highly correlated. A solution for such type of offending estimates is elimination of one of the constructs or one has to ensure that true discriminant validity is established among the constructs (Hair et al., 2006).

Table 4 present the correlations among all latent constructs of the model. The matrix shows that all the constructs are significantly correlated with each other and none of the correlations are above 0.68. Therefore, multicollinearity problem is fairly low. Values exceeding 0.80 can be indicative of problems and values exceeding 0.90 should always be examined (Hair et al., 2006).

Table 4
Correlation Among all Latent Constructs

Variable	1	2	3	4	5
1.Opportunity for Skill Utilization	1.00				
2.Program Design	.663	1.00			
3.Individual Initiative	.385	.542	1.00		
4.MD Effectiveness	.579	.637	.554	1.00	
5.Line Manger Support	.684	.554	.503	.352	1.00

4.1.2 Standardized Factor Loading

Another type of offending estimate in CFA is the standardized factor loading that exceeds or is very close to 1.00 (Hair et al., 2006).To deal with offending estimates and to achieve the best fitting measurement model, Hair et al. (2006) and Cheng (2001) suggest the deletion of offending variables or setting up a small value (0.005) for corresponding error variance to ensure that loading will be less than 1.0. Segars and Grover (1993) recommend that deletion of offending indicators should be made one by one as deletion of one indicator or measure may affect other parts of the model instantaneously. The model is then required to be re-estimated.

Table 5 and Figure 3 show all the indicator loadings were statistically significant for the proposed constructs and no indicators had loading so low that they should be deleted. The "t" values associated with each of the loadings exceed the critical values for the .05 significance level (critical value = 1.96) as suggested by Hair et al. (2006). All the variables are significantly related to their specified constructs, verifying the posited relationships among indicators and constructs. In summary, the various measures of overall model goodness of fit and standardized regression

Table 5
Measurement Model Results
(Standardized Regression Weights or Construct Loadings)

Constructs (Variables)	Indicators	Standardized Structural Coefficient	t values
Management Development Effectiveness	MD1	0.826***	12.73
	MD2	0.818***	12.54
	MD3	0.732***	10.68
	MD4	0.829***	12.81
	MD5	0.632***	08.78
	MD6	0.746***	10.95
Individual Initiative	Individual1	0.788***	11.26
	Individual2	0.810***	11.69
	Individual3	0.532***	06.86
	Individual6	0.616***	08.18
	Individual7	0.500***	06.30
Line Manager Support	Line1	0.657***	09.30
	Line2	0.652***	09.19
	Line3	0.783***	11.81
	Line4	0.843***	13.21
	Line5	0.890***	14.43
	Line6	0.819***	12.65
Opportunity for Skill Utilization	Opportunity1	0.795***	11.96
	Opportunity2	0.750***	10.99
	Opportunity3	0.776***	11.53
	Opportunity4	0.733***	10.64
	Opportunity5	0.749***	10.97
	Opportunity6	0.694***	09.89
Program Design	Design1	0.763***	11.29
	Design2	0.832***	12.82
	Design3	0.745***	10.90
	Design4	0.741***	10.82
	Design5	0.766***	11.35
	Design6	0.708***	10.16

Notes: ***p 0.001 (two-tailed)

weights lend sufficient support to deem the results as acceptable representation of the hypothesized constructs.

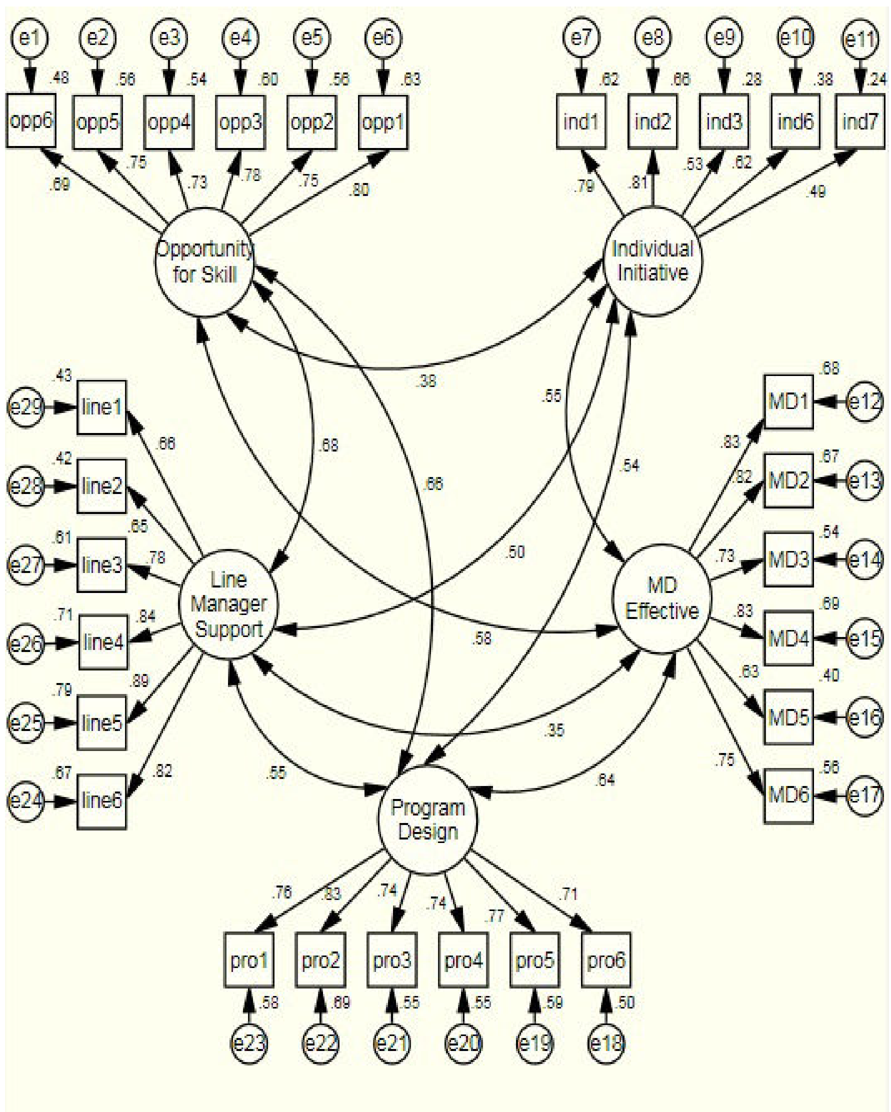

Fig. 3 Results of CFA

4.1.3 Standard Error

A very large standard error associated with any estimated coefficient is also treated as an offending estimation. Remedy for such a problem is deletion of such indicators (Hair et al., 2006). Examinations of the results reveal no instance of such a problem in the model.

In sum, examination of results indicate that there is no offending estimate in the current measurement model, thus authors can proceed to assess the overall model fit (goodness of fit) of the CFA.

4.2 Overall Model Fit (goodness of fit)

Overall structural model fit is also known as evaluation of the goodness of fit measures of structural models. Goodness of fit measures determines whether the researcher should reject or accept the structural model being tested (Hair et al, 2006; Reisinger & Mavondo, 2007). There are four main types of goodness of fit measures to assess the structural model. These are:

1. Incremental Fit Measures
2. Parsimonious Fit Measures
3. Absolute Fit Measures
4. Noncentrality-based Measures

LISREL results of all four types of goodness of-fit measures of the current model of MD effectiveness are given in Table 6. Several researchers including Hair et al. (2006) and Reisinger and Mavondo (2007) are of the opinion that researchers are not required to report all types of goodness of measures, however, different indices from each type of goodness of fit meausre should be reported to assess the structural model. In the current examples different indices from all four types of goodness of fit measures are given. These results indicate that the proposed model has a best fit to the data as all measures achieved the acceptable level and the model is accepted.

Table 6
Comparison of Goodness of Fit Measures for Structural Model

Measurement Model Data
29 indicators for 5 constructs (1 exogenous, 4 endogenous)
Total degree of freedom = 367   Sample Size = 168
Proposed Model: chi-square = 647.87   df =367   *p* = .000
Null or Independent Model: chi-square = 3262   df = 406   *p* = .000

Goodness of Fit Measure	Accepted Value*	Calculation	Adequacy*
1.Absolute Fit Measures			
Likelihood ratio chi-square statistic(χ^2)		χ^2= 647.87 Significance level:000	
Goodness of Fit index (GFI)	Higher values indicates better fit, no established thresholds	GFI = 0.79	Marginal
2.Incremental Fit Measures			
Tucker-Lewis Index (TLI)	Acceptable value: ≥0.90	TLI or NNFI = 0.97	Good
Normed Fit Index(NFI)	Acceptable value: ≥0.90	NFI = 0.93	Good
Incremental Fit Index (IFI)	Acceptable value: ≥0.90	IFI = 0.97	Good
3.Noncentrality Based Measures			
Root mean square error of approximation (RMSEA)	Acceptable values under 0.08	RMSEA = 0.06	Good
Comparative Fit Index (CFI)	Acceptable value: ≥0.90	CFI = 0.97	Good
4.Parsimonious Fit Measures			
Normed chi-square	Lower value: 1.0, Upper value: 3.0 or 5.0	Normed χ^2 = χ^2/d.f = 1.76	Good
Parsimonious Normed Fit Index (PNFI)	Greater value shows well fit	PNFI = 0.84	Marginal
Relative Fit Index (RFI)	Acceptable value: ≥0.90	RFI = 0.92	Good

* Source Hair et al. (2006)

4.3 Measurement Model Fit

According to Hair et al. (2006), after assessing the overall measurement model goodness of fit and examination of the construct loadings, “the reliability and variance extracted measures for each construct need to be computed to assess whether the specified indicators are sufficient in their representation of the constructs" (p. 612). Table 7 and 8 present the computations for reliability and the variance extracted measures.

4.3.1 Composite Reliability

As stated earlier, reliability is a measure of the internal consistency of constructs. In other words, reliability depicts the degree to which the items indicate the common latent construct (Hair et al., 2006). Table 7 shows that all constructs displayed satisfactory levels of reliability ranging from 0.79 to 0.90 and exceed the level of 0.70, except the individual initiative construct having value of 0.68 falling somewhat short of the recommended level.

Reliability of the constructs was measured with the help of following formula (Hair et. al, 2006).

$$\text{Construct reliability} = \frac{(\text{Sum of standardized leading})^2}{(\text{Sum of standardized leading})^2 + (\text{Sum of indicator measurement error})^{*}}$$

Table 7
Reliability for all Constructs

Constructs	Reliability
MD Effectiveness	0.89
Individual Initiative	0.78
Line Manager Support	0.90
Opportunity for Skill utilization	0.89
Program Design	0.89

4.3.2 Variance Extracted

This is another method of assessing fitness of measurement model. The

* Indicator measurement error was calculated as 1- (Sum of standardized leading) 2. Or it can be found as diagonal of the measurement error correlation matrix (theta-delta matrix) in the LISREL output.

variance extracted measure reflects the overall variance of the indicators accounted for by the latent constructs (Hair et al., 2006). Higher variance extracted values occur when the indicators are true representatives of the latent construct. For the variance extracted measures, Table 8 shows all constructs exceeded the recommended level of 0.50 or 50 per cent substantially ranging from 0.56 to 0.74 except the individual initiative. Construct "individual initiative" has a value of 0.47, falling somewhat short of the recommended level. Thus for all the constructs, the indicators are sufficient in terms of how the measurement model is now specified.

Variance extracted of the constructs was measured with the help of following formula.

Variance Extracted =

$$\frac{\text{Sum of squared standardized leadings}}{\text{Sum of standardized leading} + \text{Sum of indicator measurement error}^{*}}$$

Table 8
Variance-Extracted for all Constructs

Constructs	Variance Extracted
MD Effectiveness	0.59
Individual Initiative	0.44
Line Manager Support	0.62
Opportunity for Skill Utilization	0.56
Program Design	0.58

For example in this study, the recommended values of goodness of fit indices (i.e. attainment of best fitting measurement model) were attained and can be entered in the second stage of SEM that is testing the structural model, where the estimated coefficients for both practical and their theoretical

[*] Indicator measurement error was calculated as 1- (Sum of standardized leading) 2. Or it can be found as diagonal of the measurement error correlation matrix (theta-delta matrix) in the LISREL output.

implications are examined.

5. Item Parceling in SEM

Hair et al. (2006) argued that in SEM, a large sample is required if the model is overly large or complex. Reisinger & Mavondo (2007) stated that sample size has a significant influence on the complexity of a model, a simple model is preffered if sample size is small and complex models can be examined if a large sample is available. Hall, Snell and Singer (1999) state that increasing the number of indicators directly affects the sample requirment and further recommend that with a small sample size, the number of indicators per construct should be limited e.g., to three or four (p. 235). Item parceling is a solution for the problem of large number of indicators and small sample size. Item parcels are commonly formed in order to reduce the number of indicators of lengthy scales (Bandalos & Finney, 2001; Bagozzi & Edwards, 1994).

Little, Cunningham and Shahar (2002) stated that "Parceling is a measurement practice that is used in multivariate data analysis approaches, particularly for use with latent variable analysis techniques such as SEM and Confirmatory Factor Analysis (CFA)" (p.152). Bandalos and Finney (2001) define item parceling as "a process by which raw item responses are combined into sub-scales prior to analysis". This process is done by combining or averaging item responses into parcel score, these parcels are used as the observed variables most commonly in CFA or SEM (Bandalos, 2002).

Bandalos (2002) argued that use of item parceling has become common in SEM. Meade and Kroustalis (2005) stated that because of advantageous properties, parcels have been advocated by many authors. These include greater reliability than individual items, a more optimal indicator to sample size ratio, a greater likelihood of achieving a proper model solution and better model fit. Bandalos and Finney (2001) reported that researchers have cited three common reasons for using item parceling, first, it increases the stability of the parameter estimated, second, it improves the variable to sample size ratio and third, it is a remedy to small sample size.

Bagozzi and Edwards (1998) argued that item parceling can reduce the number of parameters estimated, resulting in more stable parameter estimates and proper solution of model fit. Coffman and MacCallum (2005) state "in SEM or CFA as the number of indicators increases so do the number of parameters estimated and the order of correlation matrix. The larger the order of a correlation matrix the less likely the model is to fit well". From this perspective using parcels rather than items as indicators of latent variables involve reduction in the number of measured variables and is likely to fit the model better than the model with items as indicators (Coffman & MacCallum, 2005).

Table 9
Simple Random Parceling

Name of Constructs	Name of Parcels	Aggregated Items
Management Development Effectiveness	MDE1	MD1+MD2
	MDE2	MD3+MD4
	MDE3	MD5+MD6
Line Manager Support	LineM1	Line1+ Line 2
	LineM2	Line 3+ Line 4
	LineM3	Line 5+ Line 6
Opportunity for Skill Utilization	Opportunity1	opp1+ opp2
	Opportunity2	opp3+opp 4
	Opportunity3	opp5+ opp6
Program Design	Program1	pro1+ pro2
	Program2	pro3+ pro4
	Program3	pro5+pro6
Individual Initiative	Individual1	Ind1+Ind2
	Individual2	Ind3+Ind6+Ind7

The use of item parceling is not without controversy. Perhaps most important is determining the dimensionality of the items to be parceled (Bandalos, 2002). Because the dimensional nature of a measured construct can have a serious impact on the accuracy and validity of various parceling techniques (Little et al., 2002). Bandalos and Finney (2001) recommended that researchers should use item parceling only when parceled items are strictly unidimentional. Item parcels work effectively when constructed on unidimensional structures (Little et al., 2002). "It has also been found that the use of parceling can result in biased estimates of model parameters" (Hall et al., 1999). In sum, the amount of arguments in favor of its advantage side far outweigh the disadvantages of item parceling (Little et al., 2002) and researchers will continue to view item parceling as an attractive option (Hall et al., 1999).

After determining the nature of dimentionality of a set of items, one or other techniques for parceling items can be applied. Based on the unidimentional nature of the measures, "simple random assignment" technique was used. In a simple method for constructing parcels, "all items are assigned randomly and without replacement to one of the parcels grouping, and depending on the number of items to be assigned, two, three or, possibly four parcels could be created" (Little et al., 2002, p. 165). As discussed before, keeping in view the recommendations of Hair et al. (2006), Little et al. (2002) and Hall et al. (1999), it was decided to create 3 parcels per latent construct. Table 9 depicts the simple random parceling process and name of parcels along with their aggregated items.

Table 10
Goodness of Fit Indices for CFA with Parcels

Goodness of fit Indices	**Calculation of Measure***
χ^2	100.66
d.f	67.00
χ^2/d.f	1.502
GFI	0.92
NFI	0.97
IFI	0.99
RFI	0.96
CFI	0.99
RMSEA	0.05
NNFI	0.99

5.1 Goodness of Fit indices after Item Parceling

Table 10 represents the goodness of fit indices for the measurement model after item parceling. The table indicates that the χ^2/d.f ratio was 1.502, which was much smaller than the threshold value of 3.00 (Hair et al., 2006). All other indices supported a good fit to the data as compared to the recommended values. (Detailed discussion on goodness of fit indices and recommended values are given in Table 6).

6. Conclusion

This article provides some basic techniques/guidelines for researchers to pre-SEM technical data analysis, specification and estimation of measurement model prior to estimation of a structural model. Guidelines included in this article could not cover all aspects because of the complex nature of SEM. A careful analysis of data leads to better prediction and more accurate assessment of dimensionality. By devoting considerable time and effort on examining and exploring the nature of data and the relationships among variables, estimation of a measurement model before the application of SEM techniques (structural model examinations) can help researchers in resolving the procedural issues and assist in insightful interpretation of the results. Researchers are strongly recommended to hold all these techniques/procedures before they face problems during SEM analysis that force them to do so.

References

Anderson, J.C., and Gerbing, D.W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Phychological bulletin, 103*(4), 411-423.

Bagozzi, R.P., and Edwards, J.R. (1998). A General Approach for Representing Constructs in Organizational Research. *Organizational Research Methods, 1*, 45-87.

Bandalos, D.L. (2002). The Effects of Item Parceling on Goodness-Of-Fit And Parameter Estimate Bias in Structural Equation Modeling.

Structural Equation Modeling, 9(1), 78-102.

Bandalos, D.L., and Finney, S.J. (2001). Item Parceling Issues in Structural Equation Modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New Developments and Techniques in Structural Equation Modeling* (pp. 269-296). Mahwah: Lawerence Erlbaum Associates, Inc.

Bulut, C., and Culha, O. (2010). The Effects of Organizational Training on Organizational Commitment. *International Journal of Training and Development, 14*(4), 309-322.

Carter, R.L. (2006). Solutions for Missing Data in Structural Equation Modeling. *Research & Practice in Assessment, 1*(1), 1-6.

Cheng, E. W. (2001). SEM Being More Effective than Multiple Regression in Parsimonious Model Testing for Management Development Research. *Journal of Management Development, 20* (7), 650-667.

Chiaburu, D.S., and Marinova, S.V. (2005). What predicts skill transfer? An exploratory study of goal orientation, training self-efficacy and organizational supports. *International Journal of Training and Development, 9*, 110-123.

Coffman, D.L., and MacCallum, R.C. (2005). Using Parcels to Convert Path Analysis Models into Latent Variable Models. *Multivariate Behavioral Research, 40*(2), 235-259.

Dastgeer, G. (2012). *A Study of Management Development Effectiveness of Banking Organizations of Pakistan*. Unpublished doctoral thesis, National University of Modern Languages, Islamabad, Pakistan.

Dastgeer, G., and Rehman, A. (2012). Effectiveness of Management Development in Pakistani Corporate Sector: Testing the D'Netto Model. *Journal of Management Development, 31*(8), 740-751.

D'Netto, B., Bakas, F., & Bordia, P. (2008). Predictors of Management Development Effectiveness: An Australian Perspective. *International Journal of Training and Development, 12* (1), 2-23.

Garcia-Morales, V.J., Llorens-Montes, F.J., and Verdu-Jover, A.J. (2008). The Effects of Transformational Leadership on Organizational Performance through Knowledge and Innovation. *British journal of Management, 19*, 299-319.

Germain, R., Droge, C., & Daugherty, P. J. (1994). The Effect of Just-In-Time Selling on Organizational Structure: An Empirical Investigation. *Journal of Marketing Research, 31*(4), 471-483.

Hair, J.F., Anderson, R.E., Tatham, R.L., and Black, W.C. (2006). *Multivariate Data Analysis* (5th ed.). Upper Sadle River: NJ, Pearson Education Incorporation.

Hall, R. J., Snell, A.F., and Foust, M.S. (1999). Item Parceling Strategies in SEM: Investigating the Subtle Effects of Unmodeled Secondry Constructs. *Organizational Research Methods, 2*(3), 233-256.

Hershberger, S.L. (2003). The Growth of Structural Equation Modeling: 1994-2001. *Structural Equational Modeling, 10*(1), 35-46.

Hoe, S.L. (2008). Issues and Procedures in Adopting Structural Equation Modeling Technique. *Journal of Applied Quantitative Methods, 3*(1), 76-83.

Hox, J.J., and Bechger, T.M. (1998). An Introduction to Structural Equation Modeling. *Family Science Review, 11*, 354-373.

Little, T. D., Cunningham, W. A., and Shahar, G. (2002). To Parcel or not to Parcel: Exploring the Question, Weighing The Merits. *Structual Equation Modeling, 9*(2), 151-173.

Meade, A.W., and Kroustalis, C.M. (2005). *Problems of Item Parceling with CFA: Test of Measurement Invariance.* Paper presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Nunnally, J., and Bernstein, I. (1994). *Psychometric Theory.* New York, NY: McGraw-Hill.

Reisinger, Y., and Mavondo, F. (2007). Structural Equation Modeling. *Journal of Travel& Tourism Marketing, 21*(4), 41-71.

Schreiber, J.B., Frances, K.S., Jamie, K., Amaury, N., and Elizabeth, A. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99*(6), 323-337.

Segars, A.H., and Grover, V. (1993). Re-Examining Perceived Ease of Use and Usefulness: A Confirmatory Factor Analysis. *MIS Quarterly, 17*(4), 517-525.

Sekaran, U. (2006). *Research Methods for Business.* India: Wiley India Pvt. Ltd.

Tracey, J.B., Tannenbaum, S., and Mathieu, J.E. (2001). The Influence of Individual Characteristics and the Work Environment on Varying Levels of Training Outcomes. *Human Resource Development Quarterly, 12*(1), 5-23.