Chapter V
Theory Development in Information Systems Research Using Structural Equation Modeling: Evaluation and Recommendations

Nicholas Roberts
Clemson University, USA

Varun Grover
Clemson University, USA

ABSTRACT

Structural equation modeling (SEM) techniques have significant potential for assessing and modifying theoretical models. There have been 171 applications of SEM in IS research, published in major journals, most of which have been after 1994. Despite SEM’s surging popularity in the IS field, it remains a complex tool that is often mechanically used but difficult to effectively apply. The purpose of this study is to review previous applications of SEM in IS research and to recommend guidelines to enhance the use of SEM to facilitate theory development. The authors review and evaluate SEM applications, both component-based (e.g., PLS) and covariance-based (e.g., LISREL), according to prescribed criteria. Areas of improvement are suggested which can assist application of this powerful technique in IS theory development.

INTRODUCTION

Structural equation modeling (SEM) has become an important and widely diffused research tool for theory development in the social and behavioral sciences. One reason for the substantive use of SEM is that it enables researchers to conduct a single, systematic and comprehensive analysis
by modeling relationships among multiple independent and dependent variables simultaneously (Kline, 2005). Additionally, in contrast to exploratory methods, SEM allows for the specification of a precise model that is driven by theoretical considerations (Bollen, 1989). Finally, SEM also permits researchers to model higher-order latent variables (Edwards, 2001). These inherent advantages provided by SEM have caused many researchers in the information systems (IS) field to use it for measuring constructs or developing and testing IS theories.

Despite SEM’s numerous advantages, the relative sophistication of SEM also makes it prone to misuse (Anderson & Gerbing, 1988). Moreover, theory development relies upon the effective use of empirical research methods (Van Maanen, Sorensen, & Mitchell, 2007). Invalid theory development could greatly inhibit the building of a cumulative tradition of research. Thus, we believe it is important to take stock of how this powerful technique has been applied in IS research.

To strengthen ties between theory and empirical IS research, this study provides an in-depth review and analysis of a critical mass of SEM applications in three top-tier IS journals. Based on our review, we suggest specific areas for improvement. To the best of our knowledge, no comprehensive survey of contemporary SEM applications in the IS field has been reported in the literature.

OVERVIEW OF STRUCTURAL EQUATION MODELING

To provide a basis for subsequent discussion, we present a brief overview of SEM. SEM is a technique used to specify, estimate, and evaluate models of linear relationships among a set of observed variables in terms of a generally smaller number of unobserved variables. Figure 1 depicts a basic latent variable model. A circle is used to represent each of the four latent variables, and the boxes represent associated manifest or indicator variables. The relationships between the latent variables and their indicators are often referred to as a “measurement” model, in that it represents an assumed process in which an underlying construct determines or causes behavior that is reflected in measured indicator variables.

Within this context, it is important to note that the arrows go from the circles to the boxes, which is consistent with the process noted above. Thus, each factor serves as an independent variable in the measurement model, and the indicator variables serve as the dependent variables. Each indicator is also potentially influenced by a second independent variable in the form of measurement error, and its influence is represented as a cause of the indicator variable through the use of a second arrow leading to each of the indicators. Finally, the model shown in Figure 1 includes correlations (double-headed arrows) among the three exogenous constructs (LV1–LV3) and regression-like structural parameters linking exogenous and endogenous constructs (e.g., LV3, LV4). The model also acknowledges that there is unexplained variance in the endogenous latent variable. The part of the overall model that proposes relationships among the latent variables is often referred to as the structural model.

Often using a maximum likelihood function, covariance-based SEM techniques attempt to minimize the difference between the sample covariances and those predicted by the theoretical model. As a result, the parameters estimated by this procedure attempt to reproduce the covariance matrix of the observed measures. Observed measures are assumed to have random error variance and measure-specific variance components that are not of theoretical interest. Hence, this error variance is modeled separately. Following this, the covariances among the latent variables are adjusted to reflect the attenuation in the observed covariances due to excluded error variance components. Because of this assumption, “the amount of variance explained in the set of observed measures by modeling relationships among multiple independent and dependent variables simultaneously (Kline, 2005). Additionally, in contrast to exploratory methods, SEM allows for the specification of a precise model that is driven by theoretical considerations (Bollen, 1989). Finally, SEM also permits researchers to model higher-order latent variables (Edwards, 2001). These inherent advantages provided by SEM have caused many researchers in the information systems (IS) field to use it for measuring constructs or developing and testing IS theories.

Despite SEM’s numerous advantages, the relative sophistication of SEM also makes it prone to misuse (Anderson & Gerbing, 1988). Moreover, theory development relies upon the effective use of empirical research methods (Van Maanen, Sorensen, & Mitchell, 2007). Invalid theory development could greatly inhibit the building of a cumulative tradition of research. Thus, we believe it is important to take stock of how this powerful technique has been applied in IS research.

To strengthen ties between theory and empirical IS research, this study provides an in-depth review and analysis of a critical mass of SEM applications in three top-tier IS journals. Based on our review, we suggest specific areas for improvement. To the best of our knowledge, no comprehensive survey of contemporary SEM applications in the IS field has been reported in the literature.

OVERVIEW OF STRUCTURAL EQUATION MODELING

To provide a basis for subsequent discussion, we present a brief overview of SEM. SEM is a technique used to specify, estimate, and evaluate models of linear relationships among a set of observed variables in terms of a generally smaller number of unobserved variables. Figure 1 depicts a basic latent variable model. A circle is used to represent each of the four latent variables, and the boxes represent associated manifest or indicator variables. The relationships between the latent variables and their indicators are often referred to as a “measurement” model, in that it represents an assumed process in which an underlying construct determines or causes behavior that is reflected in measured indicator variables.

Within this context, it is important to note that the arrows go from the circles to the boxes, which is consistent with the process noted above. Thus, each factor serves as an independent variable in the measurement model, and the indicator variables serve as the dependent variables. Each indicator is also potentially influenced by a second independent variable in the form of measurement error, and its influence is represented as a cause of the indicator variable through the use of a second arrow leading to each of the indicators. Finally, the model shown in Figure 1 includes correlations (double-headed arrows) among the three exogenous constructs (LV1–LV3) and regression-like structural parameters linking exogenous and endogenous constructs (e.g., LV3, LV4). The model also acknowledges that there is unexplained variance in the endogenous latent variable. The part of the overall model that proposes relationships among the latent variables is often referred to as the structural model.

Often using a maximum likelihood function, covariance-based SEM techniques attempt to minimize the difference between the sample covariances and those predicted by the theoretical model. As a result, the parameters estimated by this procedure attempt to reproduce the covariance matrix of the observed measures. Observed measures are assumed to have random error variance and measure-specific variance components that are not of theoretical interest. Hence, this error variance is modeled separately. Following this, the covariances among the latent variables are adjusted to reflect the attenuation in the observed covariances due to excluded error variance components. Because of this assumption, “the amount of variance explained in the set of observed
The importance of theory to the advancement of scientific knowledge in the IS field cannot be overstated. Theory allows researchers to understand and predict outcomes of interest, even if only probabilistically (Cook & Campbell, 1979). Theory also allows researchers to describe and explain a process or series of events (Mohr, 1982). As Hall and Lindzey (1957) noted, the function of theory “is that of preventing the observer from being dazzled by the full-blown complexity of natural or concrete events” (p. 9). Thus, the purpose of theory is twofold: to organize (parsimoniously) and to communicate (clearly).

Many scholars define theory in terms of relationships between independent and dependent measures is not of primary concern” (Anderson & Gerbing, 1988, p. 412). Thus, covariance-based SEM techniques provide parameter estimates that best explain the observed covariances. Moreover, covariance-based SEM techniques also provide the most efficient parameter estimates (Joreskog & Wold, 1982) and an overall test of model fit. In contrast to reproducing the covariance matrix of the observed variables, a component-based SEM approach focuses on maximizing the variance explained by the structural model (Chin, 1998). Based on a principal-component model, no random error variance or measure-specific variance is assumed. Parameter estimates are obtained based on the ability to minimize the residual variances of dependent variables (Fornell & Bookstein, 1982). Fit is assessed on the basis of the percentage of variance explained in the regression equations. Component-based SEM techniques also estimate latent variables as linear combinations of their observed measures, thereby providing an exact definition of component scores. This advantage of precise definitions in conjunction with explaining a large percentage of the variance in observed measures is valuable in accurately predicting individuals’ standings on the components (Anderson & Gerbing, 1988).
variables. For instance, Bacharach (1989) defines theory as “a statement of relations among concepts within a set of boundary assumptions and constraints” (p. 496). Theories range from guesses, conjectures, or speculation to more formal propositions, hypotheses, or models, with those that are more formal more likely to appear in print (Weick, 1995). Theory development can be conceptualized as a discourse between generating explanations for phenomena and using appropriate methods to evaluate their validity (Van Maanen et al., 2007). Because the advancement of theory partially rests on methodological tools used to help define ideas and test their conclusions (Blalock, 1969), it is important to match research methods to stages of theory development.

SEM techniques are useful for establishing relationships between constructs (Bollen, 1989). Specifically, SEM techniques leverage knowledge gained from case studies because they allow researchers to examine patterns of relationships across individuals, groups or organizations. For example, partial least squares (PLS) is a component-based SEM technique that centers on establishing that relationships exist and maximizing the variance explained by the structural model (Wold, 1982). PLS’s emphasis on relationship building and variance explanation makes it well suited for theory building (Joreskog & Wold, 1982). Covariance-based SEM techniques are also useful in this stage for three reasons: (1) when theory suggests competing models exist, researchers can use SEM techniques to model and test alternative models with the same data set (Anderson & Gerbing, 1988); (2) modification indices can provide insight into plausible alternative explanations for relationships among constructs (MacCallum, 1986); and (3) SEM results are replicable and reusable, thereby providing researchers with opportunities to independently confirm results and evaluate alternative models (Kline, 2005).

Theory building is followed by theory testing, a stage in which rigorous tests are conducted to determine whether hypothesized relationships among constructs actually exist. Appropriate theory testing methods include SEM techniques, regression, and analysis of variance (ANOVA). These methods allow researchers to estimate relationships and make inferences based on statistical analysis. SEM techniques are especially powerful in that they enable researchers to test complete research models (Anderson & Gerbing, 1988). Because covariance-based SEM techniques focus on maximizing the “fit” of the theoretical model to the observed data, they are considered valuable for evaluating nomological nets (Bollen, 1989; Gefen, Straub, & Boudreau, 2000). Table 1 details our discussion of research methods that support various stages of the theory development process.

**LITERATURE REVIEW**

This study assesses the diffusion and use of SEM techniques in the IS field. We reviewed applications of SEM techniques in three well-established

<table>
<thead>
<tr>
<th>Theory Development Stage</th>
<th>Supporting Research Method</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploratory</td>
<td>In-depth case studies, exploratory interviews</td>
<td>Identification of key concepts and issues, evidence that a phenomenon is important</td>
</tr>
<tr>
<td>Theory Building</td>
<td>Multi-site case studies, SEM techniques</td>
<td>Construct definitions, explanations for relationships among constructs, initial tests of relationships, search for alternative models</td>
</tr>
<tr>
<td>Theory Testing</td>
<td>SEM techniques, ANOVA, regression</td>
<td>Omnibus tests of theoretical relationships, evaluation of nomological networks</td>
</tr>
</tbody>
</table>
IS journals: *Information Systems Research* (ISR), *Journal of Management Information Systems* (JMIS), and *MIS Quarterly* (MISQ). Consistent with prior research (Gefen et al., 2000), all issues of these journals published over the period 1990-2007 were searched for empirical SEM applications. Theoretical papers addressing issues related to SEM and papers using ANOVA, exploratory factor analysis, path analysis, and regression were excluded from the sample. Table 2 shows the number of articles collected from each journal. Table 3 displays the number of articles using either component-based or covariance-based SEM.

A total of 171 articles satisfied our selection criteria (see Table 2). One apparent trend is that the use of SEM in IS research steadily increased over the period 1998 to 2007. The SEM applications were evenly distributed across the three major IS journals. Another trend is that use of component-based techniques rose substantially starting in 2003 (see Table 3).

### METHODOLOGICAL SEM ISSUES

Figure 2 depicts the recommended sequence of activities that need to be performed to conduct effective SEM analysis: (1) model specification, (2) data screening, (3) model estimation and assessment, and (4) model respecification (Kline, 2005; Hair et al. 2006). Each stage consists of a set of salient methodological issues. For instance, researchers need to consider indicator selection, formative/reflective approaches, model identification, and sample size issues in the model specification stage. For each SEM analysis stage

**Table 2. Applications of SEM in IS research**

<table>
<thead>
<tr>
<th>Year</th>
<th>ISR</th>
<th>JMIS</th>
<th>MISQ</th>
<th>Total</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>1991</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2.3</td>
</tr>
<tr>
<td>1992</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>1993</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>1994</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>3.5</td>
</tr>
<tr>
<td>1995</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>4.1</td>
</tr>
<tr>
<td>1996</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>2.3</td>
</tr>
<tr>
<td>1997</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>3.5</td>
</tr>
<tr>
<td>1998</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>1999</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>9</td>
<td>5.3</td>
</tr>
<tr>
<td>2000</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>6.4</td>
</tr>
<tr>
<td>2001</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>5.3</td>
</tr>
<tr>
<td>2002</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>16</td>
<td>9.4</td>
</tr>
<tr>
<td>2003</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>15</td>
<td>8.8</td>
</tr>
<tr>
<td>2004</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>15</td>
<td>8.8</td>
</tr>
<tr>
<td>2005</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>18</td>
<td>10.5</td>
</tr>
<tr>
<td>2006</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>21</td>
<td>12.3</td>
</tr>
<tr>
<td>2007</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>22</td>
<td>12.9</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>59</td>
<td>55</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>33.3</td>
<td>34.5</td>
<td>32.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
we describe methodological issues, evaluate the current state of IS research with respect to those issues, and provide recommendations for future research. We also note key differences in these points for component-based and covariance-based SEM approaches.

Model Specification Issues and Recommendations

We cover four issues related to model specification: (1) indicator selection, (2) formative/reflective, (3) model identification, and (4) sample size.

Indicator Selection

A critical decision in survey-based research involves how many indicators can be explicitly related to a latent variable (i.e., construct). Technically, a latent variable may be assessed with only two indicators under certain conditions. However, models with low indicator-to-construct ratios often cause problems with identification and convergence. Hence, scholars recommend that each construct be measured with at least three indicators (Anderson & Gerbing, 1988). In the theory building stage, guidelines suggest that researchers have at least four or five indicators per latent variable, as it is often necessary to drop some indicators in order to achieve construct validity (Churchill, 1979) and arrive at a well-fitting measurement model (Kline, 2005). In either case, it is important that researchers use appropriate indicators to capture the domain space of the construct (Little, Lindenberger, & Nesselroade, 1999).

### Table 3. Applications of SEM in IS research

<table>
<thead>
<tr>
<th>Year</th>
<th>Component</th>
<th>Covariance</th>
<th>Total</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>1991</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2.3</td>
</tr>
<tr>
<td>1992</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>1993</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1.2</td>
</tr>
<tr>
<td>1994</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>3.5</td>
</tr>
<tr>
<td>1995</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>4.1</td>
</tr>
<tr>
<td>1996</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2.3</td>
</tr>
<tr>
<td>1997</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>3.5</td>
</tr>
<tr>
<td>1998</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>1999</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>5.3</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>6.4</td>
</tr>
<tr>
<td>2001</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>5.3</td>
</tr>
<tr>
<td>2002</td>
<td>4</td>
<td>12</td>
<td>16</td>
<td>9.4</td>
</tr>
<tr>
<td>2003</td>
<td>8</td>
<td>7</td>
<td>15</td>
<td>8.8</td>
</tr>
<tr>
<td>2004</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>8.8</td>
</tr>
<tr>
<td>2005</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td>10.5</td>
</tr>
<tr>
<td>2006</td>
<td>14</td>
<td>7</td>
<td>21</td>
<td>12.3</td>
</tr>
<tr>
<td>2007</td>
<td>18</td>
<td>4</td>
<td>22</td>
<td>12.9</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
<td>75</td>
<td>171</td>
<td></td>
</tr>
<tr>
<td>(%)</td>
<td>56.1</td>
<td>43.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evaluation and Recommendation: Table 3 provides descriptive statistics regarding the issue of indicator selection. Overall, the median number of indicators across all SEM applications was 26 and the median number of constructs was 6, resulting in a median ratio of indicators to constructs of approximately 4.3. The median ratios of indicators to constructs were 4.3 and 4.8 for component-based SEM and covariance-based SEM, respectively.

A notable proportion of SEM applications in IS research used single indicator constructs when multiple indicators per construct is desirable. The use of single indicator constructs is not recommended because single indicator constructs ignore measurement unreliability, which is one of the problems that SEM was specifically designed to circumvent (Gefen et al., 2000). Researchers must also consider issues regarding indicator selection and model identification. If a standard confirmatory factor analysis (CFA) model with a single factor has at least three indicators, the model is identified. If a standard CFA model with two or more factors has at least two indicators per factor, the model is identified. Bollen (1989) referred to the latter condition as the two-indicator rule. However, models with factors that have only two indicators are more prone to estimation problems, especially when the sample size is small (Kline, 2005). Hence, a minimum of three indicators per factor is recommended.

Formative/Reflective

The distinction between formative and reflective indicators has recently gained attention in the IS field. Constructs are usually viewed as causes of indicators, meaning that variation in a construct leads to variation in its indicators (Bollen, 1989). Such indicators are termed reflective because they represent reflections, or manifestations, of a construct (Fornell & Bookstein, 1982; Gefen et al., 2000). For example, behavioral intention to use a system is often operationalized with three reflective indicators (e.g., Davis, Bagozzi, & Warshaw, 1989; Venkatesh, Morris, Davis, & Davis, 2003). Hence, an individual’s change in the latent behavioral intention construct results in corresponding changes in each manifest indicator of intention. Constructs can also be viewed as being formed by their indicators (Blalock, 1971; Bollen & Lennox, 1991). Such constructs are termed formative, meaning the construct is formed or induced by its measures (Fornell & Bookstein, 1982; Gefen et al., 2000). Formative constructs are commonly conceived as composites of specific component variables or dimensions (Edwards & Bagozzi, 2000). For example, if information quality is defined in terms of accuracy, completeness, currency and format, its value may vary with changes in any one of its indicators (Nelson, Todd, & Wixom, 2005).
Table 4. Analysis of model specification issues

<table>
<thead>
<tr>
<th>Model Specification</th>
<th>Component (n=96)</th>
<th>Covariance (n=75)</th>
<th>Overall (n=171)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of indicators</td>
<td>29 (20, 37)</td>
<td>24 (17.5, 38)</td>
<td>26 (19, 37)</td>
</tr>
<tr>
<td>Number of constructs</td>
<td>6.5 (4, 8)</td>
<td>5 (3, 7)</td>
<td>6 (4, 8)</td>
</tr>
<tr>
<td>Ratio of indicators to constructs</td>
<td>4.3 (3, 5.7)</td>
<td>4.8 (3.4, 9.8)</td>
<td>4.3 (3.2, 7.3)</td>
</tr>
<tr>
<td>Percentage of models containing at least one single-indicator construct</td>
<td>21</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>Percentage of models containing at least one formative construct</td>
<td>46</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td>Sample size (median)</td>
<td>179 (99, 267)</td>
<td>256 (161, 409)</td>
<td>212 (128, 355)</td>
</tr>
</tbody>
</table>

Evaluation and Recommendation: While an assessment of the proper use of formative constructs is beyond the scope of our investigation, we did find that 36 percent of all SEM applications had at least one formative construct in the research model. Furthermore, 46 percent of the component-based SEM applications had at least one formative construct, while 24 percent of covariance-based SEM applications had at least one formative construct. While SEM scholars have noted the importance of appropriately conceptualizing and testing constructs as either formative or reflective (Jarvis, Mackenzie, & Podsakoff, 2003; Petter, Straub, & Rai, 2007), we also note that formative representations are fraught with problems like interpretational confounding and external consistency (Howell, Breivik, & Wilcox, 2007). Furthermore, properly specifying the nature of relationships between constructs and their indicators is important because these relationships constitute a secondary theory that bridges the gap between abstract theoretical constructs and measurable empirical phenomena. Without this secondary theory, the mapping of theoretical constructs onto empirical phenomena is ambiguous, and theories cannot be meaningfully tested (Blalock, 1971).

Model Identification

Once a theoretical model has been specified, it is necessary to consider whether or not it is statistically identified. A model is said to be identified when it is theoretically possible to derive a unique estimate of each parameter (Kline, 2005). An underidentified model creates problems because it is possible for two distinct sets of parameter values to yield the same population variance-covariance matrix. In other words, two different solutions for the same structure with widely differing theoretical implications can account for the data equally well. Thus, it is important to address identification issues.

Evaluation and Recommendation: Our review found that very few researchers discuss whether they checked for model identification. This might be due to two reasons. First, journal space is a scarce resource, and second, most SEM computer software packages provide a warning message when a model is underidentified. A simple rule to avoid underidentification is that the number of freely estimated parameters should not exceed the number of distinct elements in the variance-covariance matrix of the observed variables (Sharma, 1996). Formally, the number of free parameters
must be less than or equal to the number of observations (i.e., model degrees of freedom ≥ 0). The number of observations equals \( v(v + 1)/2 \), where \( v \) is the number of observed variables.

**Sample Size**

When using covariance-based SEM techniques, small sample sizes can cause nonconvergence and biased parameter estimates (Anderson & Gerbing, 1988; Fornell, 1983), thereby inhibiting theory development. Simulation studies show that small samples are not compatible with maximum likelihood estimation of covariance structure models (Boomsma, 1982). Anderson and Gerbing (1984) suggest that a sample size of 150 or more will typically be needed to obtain parameter estimates that have standard errors small enough to be of practical use. Other scholars recommend a minimum sample size of 200 for SEM analysis (Hair, et al., 2006). Alternately, Bentler and Chou (1987) argue that, under normal distribution theory, the ratio of cases to free parameters should be at least 5:1 to calculate reliable parameter estimates and higher (at least 10:1) to conduct meaningful significance tests.

One advantage of component-based SEM is that it is relatively robust to small sample sizes. Since component-based SEM estimates regression equations sequentially, the sample size needs to meet the demands of the most complex regression equation in the model. The most complex regression will involve either (1) the construct with the greatest number of indicators or (2) the construct with the greatest number of antecedent constructs. Sample size requirements can be calculated by multiplying ten times (1) or (2), whichever is greater (Barclay, Higgins, & Thomson, 1995).

**Evaluation and Recommendation:** Table 3 shows that the median sample size for all SEM applications is 212. Furthermore, the median sample size for covariance-based SEM applications is 256, which exceeds the conservative minimum sample size of 200 recommended by SEM scholars. The median sample size for component-based SEM applications is 179, which provides evidence that component-based SEM may be used more often than covariance-based SEM when sample size is smaller. However, researchers should not blindly use component-based SEM in all cases when sample size is small (Marcoulides & Saunders, 2006). IS researchers should take into account a number of factors when determining the appropriate sample size, such as the psychometric properties of the constructs, the strength of the relationships among the constructs, the complexity and size of the model, and the amount of missing data.

**Data Screening Issues and Recommendations**

The second stage in SEM analysis involves screening the data for missing values, normality, and outliers. We discuss data screening issues in the following sections.

**Missing Data**

The raw data should be carefully screened before a variance-covariance matrix is computed. It is important to ensure that there are no coding errors and that missing values have been appropriately addressed. Virtually all methods of statistical analysis are plagued by problems with missing data, and SEM is no exception. The use of inappropriate methods for handling missing data can lead to bias in parameter estimates (Jones, 1996), bias in standard errors and test statistics (Glasser, 1964), and inefficient use of the data (Afifi & Elashoff, 1966).

**Evaluation and Recommendation:** We found that IS researchers do not report missing data issues very often. Only 19 percent of all SEM applications discussed missing data. Again, this may be due to limited journal space. There are a number of ways to handle missing data. Conventional “deletion” methods include listwise
deletion and pairwise deletion. Another general approach to missing data is to make some reasonable guesses for the values of the missing data and then proceed to a traditional analysis of the real and imputed data. However, conventional imputation methods that use some form of the mean for data imputation lead to underestimates of standard errors (Little & Rubin, 1987). We recommend that IS researchers consider recent advances in missing data imputation techniques, such as maximum likelihood imputation and multiple imputation. These advanced methods have much better statistical properties than traditional methods (e.g., listwise deletion, pairwise deletion, mean imputation) (Allison, 2003).

Normality

Outlier detection and normality of the data distribution should also be assessed. The most widely used estimation methods in covariance-based SEM applications assume multivariate normality, which means that “(1) all the univariate distributions are normal, (2) the joint distribution of any pair of the variables is bivariate normal, and (3) all bivariate scatterplots are linear and homoscedastic” (Kline, 2005, pp. 48-49). Using data that severely departs from the assumption of normality may lead to one of two problems (Cohen, Cohen, West, & Aiken, 2003). First, parameter estimates may be biased. Specifically, estimates of the path coefficients, $R^2$, significance tests, and confidence intervals may all be incorrect. Second, the estimate of the standard error of the path coefficients may be biased. In such cases, the estimated value of the path coefficient is correct, yet hypothesis tests may not be correct.

Component-based SEM is robust to non-normal data distribution (Chin, 1998). Similar to the sample size issue, researchers should not consider component-based SEM as a silver bullet for handling non-normal data (Marcoulides & Saunders, 2006). However, we note that EQS 6.1 provides analyses which are robust to non-normal data distribution (Byrne, 2006). In particular, EQS provides a robust chi square statistic called the Satorra-Bentler scaled statistic (Satorra & Bentler, 1988) and robust standard errors (Bentler & Dijkstra, 1985), both of which have been corrected for non-normality in large samples. These robust estimates have also been applied in IS research (Swanson & Dans, 2000). We may find that future advances in other covariance-based SEM software packages, such as AMOS and LISREL, provide options that allow researchers to conduct reliable covariance-based analyses of non-normal data.

Evaluation and Recommendation: Only 11 percent of all SEM applications tested for multivariate normality. Fifteen percent of covariance-based SEM papers tested for normality, and only 7 percent of component-based SEM papers tested for normality. While minor, this distinction may be due to the fact that component-based SEM is robust to non-normality. We found a number of papers that, after testing for normality and finding a non-normal distribution of their data, decided to conduct a component-based SEM analysis instead of a covariance-based SEM analysis. We recommend that IS researchers report tests for multivariate normality when using SEM applications.

Model Estimation and Assessment Issues and Recommendations

The next stage in SEM analyses involves estimating the model and assessing the results. We discuss four methodological issues: (1) estimation approach, (2) assessment of model fit, (3) assessment of measurement model, and (4) assessment of structural model.

Estimation Approach

SEM techniques allow researchers to simultaneously estimate measurement and structural models (Bollen, 1989). The ability to do this in a one-step analysis approach, however, does not imply that it
Theory Development in IS Research Using Structural Equation Modeling

is the best way to conduct SEM analyses. Scholars recommend that researchers take a two-step estimation approach, where the measurement model is assessed prior to the estimation of the structural model (Anderson & Gerbing, 1988). The measurement model provides an assessment of convergent validity and discriminant validity (Campbell & Fiske, 1959). Once convergent and discriminant validity meet required thresholds, the test of the structural model then constitutes an assessment of nomological validity (Cronbach & Meehl, 1955).

**Evaluation and Recommendation:** Table 5 provides descriptive statistics on model estimation and assessment. Our review finds that a majority of SEM applications in IS take a two-step approach. However, 35 percent do not take a two-step approach, opting instead to simultaneously estimate measurement and structural models in a one-step approach. There is much to gain in theory development from separate estimation of the measurement and structural models (Anderson & Gerbing, 1988; Gefen et al., 2000). Hence, we recommend that IS researchers use a two-step approach in estimating SEM applications.

Assessment of Model Fit

Fit refers to the ability of a model to reproduce the data, that is, the observed variance-covariance matrix. Model fit indices are either absolute or incremental. Absolute fit indices, such as chi square statistics and root mean square error of approximation, evaluate the degree to which the model reproduces the observed covariance matrix (Kline, 2005). Incremental fit indices, such as comparative fit index and normed fit index, assess the relative improvement in fit when the model is compared with a restricted, nested baseline model (Hu & Bentler, 1998). There is no single “magic index” that provides a gold standard for all models; thus, researchers should report multiple model fit indices in order to adequately assess how well the model fits the observed data.

**Evaluation and Recommendation:** Table 5 shows that 88 percent of covariance-based SEM applications reported multiple measures of model fit. Thus, a small minority (12 percent) failed to report two or more measures of model fit. Component-based SEM applications do not attempt to reproduce the observed variance-covariance matrix. As a result, component-based SEM does not provide measures of “model fit” similar to its covariance-based SEM counterpart.

The availability of so many different fit indexes presents a number of problems: (1) different fit indexes are reported in different articles; (2) different reviewers of the same manuscript may request indexes that they know about or prefer; (3) a researcher may report only those fit indexes

---

**Table 5. Analysis of model estimation and assessment issues**

<table>
<thead>
<tr>
<th>Model Estimation &amp; Assessment</th>
<th>Component (n=96)</th>
<th>Covariance (n=75)</th>
<th>Overall (n=171)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of two step applications*</td>
<td>68</td>
<td>61</td>
<td>65</td>
</tr>
<tr>
<td>Percentage of applications that reported multiple measures of model fit</td>
<td>NA</td>
<td>88</td>
<td>NA</td>
</tr>
<tr>
<td>Percentage of models for which construct reliability were reported</td>
<td>94</td>
<td>97</td>
<td>95</td>
</tr>
<tr>
<td>Percentage of models for which convergent and discriminant validity were reported</td>
<td>75</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td>Percentage of models for which structural coefficients were provided</td>
<td>98</td>
<td>85</td>
<td>92</td>
</tr>
<tr>
<td>Percentage of models for which R² for structural equations were reported</td>
<td>92</td>
<td>52</td>
<td>74</td>
</tr>
</tbody>
</table>

* These statistics include only SEM applications with structural models.
with favorable values; and (4) a preoccupation with overall model fit may distract researchers from other important information, such as whether or not parameter estimates and variance explained values make sense. SEM scholars recommend a minimal set of fit indexes that should be reported and interpreted when reporting the results of SEM analyses (Boomsma, 2000; Gefen et al., 2000; Kline, 2005; McDonald & Ho, 2002). These statistics include: (1) the model chi-square, (2) the Steiger-Lind root mean square error of approximation (RMSEA, Steiger, 1990) with its 90% confidence interval, (3) the Bentler comparative fit index (CFI, Bentler, 1990), and (4) the standardized root mean square residual (SRMR). Finally, we also recommend IS researchers to be wary of recommended cutoff values for fit indices as universal “golden rules” that must be strictly adhered to (Marsh, Hau, & Wen, 2003).

Assessment of Measurement Model

The measurement model is usually assessed in terms of construct reliability and construct validity (Bagozzi, Yi, & Phillips, 1991). Reliability assesses the internal consistency of construct indicators (Nunnally & Bernstein, 1994). In the IS field, construct reliability is often assessed by computing Cronbach’s alpha (Straub, Boudreau, & Gefen, 2004).

Construct validity refers to the extent to which an instrument measures what it is supposed to measure (Cronbach & Meehl, 1955). Construct validity consists of convergent validity and discriminant validity. Convergent validity assesses the extent to which different indicators for the measure refer to the same construct, and discriminant validity assesses the extent to which a measure is adequately distinguishable from related constructs (Campbell & Fiske, 1959). Convergent validity can be assessed with a number of techniques, such as multi-trait multi-method (MTMM), exploratory factor analysis (EFA), and confirmatory factor analysis (CFA) (Straub et al., 2004). When using CFA, one assesses convergent validity by testing the significance of estimated factor loadings. Discriminant validity is also assessed with MTMM, EFA and CFA techniques. When using CFA, discriminant validity can be obtained through the use of a chi square difference test (Venkatraman, 1989).

Evaluation and Recommendation: As indicated in Table 5, 95 percent of all SEM applications reported assessments of construct reliability. However, 20 percent of all SEM applications did not report evidence of construct validity (i.e., both convergent and discriminant validity). Twenty-five percent of component-based SEM applications did not report evidence of construct validity, while only 15 percent of covariance-based SEM applications failed to report evidence of construct validity. Theory development in IS research rests on the effective conduct of our empirical research. Furthermore, the conceptual domain of a construct must be effectively converted into the operational domain in order to conduct empirical research (Straub et al., 2004). Thus, assessing convergent and discriminant validity is critical to the advancement of our understanding of IS-related phenomena. We encourage IS researchers to effectively and adequately assess construct reliability and validity in their empirical work, especially when using SEM applications.

Assessment of Structural Model

The structural model is assessed by investigating the sign, size and statistical significance of the structural coefficients. It is also important to assess the level of variance explained in the dependent variables (i.e., predictive validity). As noted earlier, the structural model should also be assessed in terms of model fit when using covariance-based SEM techniques.

Evaluation and Recommendation: The vast majority of SEM applications in IS research (92 percent) assess the sign, size and statistical significance of the structural coefficients.
Furthermore, 74 percent of all SEM applications report the variance explained ($R^2$) by the model. However, one interesting finding is that 52 percent of covariance-based SEM applications reported $R^2$ values, while 92 percent of component-based SEM applications reported $R^2$ values. Thus, in addition to reporting model fit and confirmation issues, IS researchers should still report $R^2$ values regardless of SEM technique used.

Model Respecification Issues and Recommendations

The final stage in SEM analyses involves specifying alternative models. When an initial model of interest does not provide a good approximation of real world phenomena (as evidenced by good model fit), researchers often alter the model to improve its fit to the data. Modification of a hypothesized model to improve its parsimony and/or fit to the data is termed a “specification search” (Long, 1983; MacCallum, 1986). A specification search is designed to identify and eliminate errors from the original specification of the hypothesized model.

Jöreskog and Sörbom (1996) describe three model specification (and evaluation) strategies: (1) strictly confirmatory, where a single a priori model is studied, (2) model generation, where an initial model is fit to the data and then modified (frequently with the use of modification indices) until it achieves adequate fit, and (3) alternative models, where multiple a priori models are studied. The “strictly confirmatory” approach is quite restrictive and does not leave the researcher any leeway if the model does not work. The model generation approach is troublesome because of the potential for abuse, results that lack validity (MacCallum, 1986), and high susceptibility to capitalizing on chance (MacCallum, Roznowski, & Necowitz, 1992). The third approach, specifying alternative models, is useful in theory building because it gives the researcher alternative perspectives concerning the focal phenomena.

Evaluation and Recommendation: In our review, 19 percent of all SEM applications reported making model comparisons. Of the models that were changed post hoc, only 30 percent of these model changes were supported with theoretical arguments. Hence, 70 percent of the model changes were performed based on data-driven reasons (e.g., recommendations based on specification searches). We recommend that IS researchers compare alternate a priori models to discover the model that the observed data support best rather than use specification searches. Moreover, all model changes should be guided by theory since there is always the danger of capitalizing on chance (MacCallum et al., 1992).

RECOMMENDATIONS

The popularity of SEM techniques has rapidly increased in recent years. We take stock of the current state of SEM research in the IS field and suggest ways in which IS researchers can better use these powerful techniques to improve IS theory. Based on our review of SEM applications in three leading IS journals between 1990 and 2007, we provide some general guidelines in which IS researchers can take full advantage of SEM techniques:

Specification: Model specification issues should be considered prior to conducting data collection efforts. Specifically, researchers should consider the number of indicators necessary to ensure construct validity, formative/reflective approaches, model identification, and requisite sample size. We recommend a minimum of three indicators per construct in order to conduct reliable analyses of parameter estimates, standard errors, and model identification. IS researchers should also undertake a careful examination of whether each construct should be modeled as formative or reflective. Since covariance-based SEM techniques require large sample sizes, researchers should consider component-based SEM
as an alternative when they are confronted with small sample sizes.

**Screening:** As in most empirical research, careful screening of raw data is necessary before estimating and testing the hypothesized model. Since missing values, outliers, and non-normal data could cause several problems in estimating structural models (e.g., nonconvergence and biased parameter estimates), data screening should be carefully considered. Since conventional data deletion and mean-based imputation methods often produce unreliable results, IS researchers should consider advanced data imputation techniques (e.g., maximum likelihood imputation, multiple imputation) when faced with missing data problems. We also note that, while component-based SEM is robust to non-normal data, recent advances in covariance-based SEM software packages provide estimates which are robust to non-normal data.

**Estimation/Assessment:** In order to gain greater confidence in their empirical findings, IS researchers should follow a two-step approach when estimating SEM applications. A two-step approach estimates the measurement model and the structural model separately. We also recommend that IS researchers report multiple measures of model fit. Doing so provides greater insight into how well the model reproduces the observed variance-covariance matrix. Theory development depends upon the valid translation of the conceptual realm to the operational realm. Hence, IS researchers should always assess reliability, convergent validity, and discriminant validity. Finally, we recommend that IS researchers report the sign, size and statistical significance of all structural coefficients, the variance explained in endogenous variables by exogenous variables, and multiple measures of model fit. Our review showed that covariance-based SEM research often concentrates on assessment of model fit, much to the expense of findings related to predictive validity (i.e., variance explained).

**Respecification:** One advantage of SEM is the ability to compare alternative models, thereby providing multiple views of a phenomena and enhancing theory development. However, post hoc changes to an initial model should be guided by theory rather than data-driven considerations.

We note that while our objective is to provide guidelines to improve the quality of future SEM applications in IS research, we do not recommend blind adherence to the individual guidelines. It is possible that other evidence presented in individual studies (e.g., strong theoretical support) can more than offset any single quality criteria. Furthermore, these criteria could be argued to vary depending on the purpose of the study. Yet consistent violation of multiple criteria will likely result in a poor application of the technique, subsequently hindering theory development in the IS field.

**CONCLUSION**

It appears as though SEM is here to stay in the IS field, at least for the foreseeable future. SEM’s ability to improve measurement reliability in multi-indicator constructs and investigate theoretical frameworks with complex relationships in a single analysis make it a powerful technique. These attributes, along with sophisticated yet relatively easy-to-use software, make it highly probable that the use of SEM will persist. Additionally, SEM is another tool that helps IS researchers strengthen the link between the conceptual realm and the operational (statistical) realm when developing theory. In particular, SEM techniques are useful for (1) establishing relationships exist between constructs, (2) modeling and testing alternative models with the same data set, and (3) confirming and replicating results. By maximizing the fit of the theoretical model to the observed data, covariance-based SEM techniques are especially useful for evaluating comprehensive nomological nets. Thus, SEM is a powerful technique which
can be used throughout the theory development process.

However, as our results suggest, astute use of this technique is necessary in order to take full advantage of its ability. In particular, IS researchers should be wary of issues related to model specification, data screening, model estimation and assessment, and model respecification. Based on our review of the IS literature and extant SEM methodological research, we provided guidelines to aid IS scholars in addressing these issues. We hope that our review of prior SEM applications will further enhance the quality of empirical research in the IS field and ultimately contribute to better development of IS theory.

REFERENCES


Theory Development in IS Research Using Structural Equation Modeling


**KEY TERMS AND DEFINITIONS**

**Construct Validity:** The extent to which a given test is an effective measure of a theoretical construct

**Factor analysis:** A statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions (factor)

**Latent Variable:** Research construct that is not observable or measured directly, but is measured indirectly through observable variables that reflect or form the construct

**Measurement Model:** Sub-model in structural equation modeling that specifies the indicators for each construct and assesses the reliability of each construct for estimating the causal relationships

**Reliability:** Extent to which a variable or set of variables is consistent in what it is intended to measure

**Structural Equation Modeling:** Multivariate technique combining aspects of multiple regression (examining dependence relationships) and factor analysis (representing unmeasured concepts with multiple variables) to estimate a series of interrelated dependence relationships simultaneously

**Structural Model:** Linkages between research constructs (or variables) that express the underlying structure of the phenomenon under investigation

**Theory:** A statement of relations among concepts within a set of boundary assumptions and constraints