

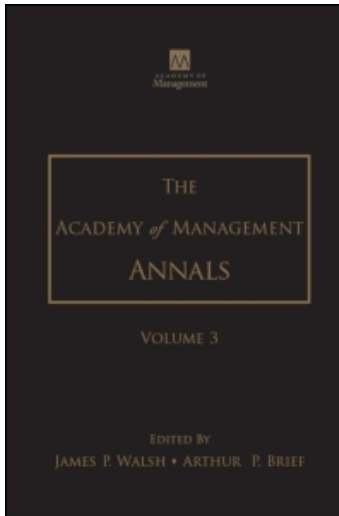
This article was downloaded by: [Academy of Management]

On: 18 August 2009

Access details: Access Details: [subscription number 791784088]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Academy of Management Annals

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t791720496>

12 Structural Equation Modeling in Management Research: A Guide for Improved Analysis

Larry J. Williams ^a; Robert J. Vandenberg ^b; Jeffrey R. Edwards ^c

^a Center for the Advancement of Research Methods and Analysis, Virginia Commonwealth University, ^b Terry College of Business, University of Georgia, ^c Kenan-Flagler Business School, University of North Carolina,

First Published on: 01 June 2009

To cite this Article Williams, Larry J., Vandenberg, Robert J. and Edwards, Jeffrey R. (2009) '12 Structural Equation Modeling in Management Research: A Guide for Improved Analysis', *The Academy of Management Annals*, 3:1, 543 — 604

To link to this Article: DOI: 10.1080/19416520903065683

URL: <http://dx.doi.org/10.1080/19416520903065683>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

12

Structural Equation Modeling in Management Research: *A Guide for Improved Analysis*

LARRY J. WILLIAMS*

*Center for the Advancement of Research Methods and Analysis,
Virginia Commonwealth University*

ROBERT J. VANDENBERG

Terry College of Business, University of Georgia

JEFFREY R. EDWARDS

Kenan-Flagler Business School, University of North Carolina

Abstract

A large segment of management research in recent years has used structural equation modeling (SEM) as an analytical approach that simultaneously combines factor analysis and linear regression models for theory testing. With this approach, latent variables (factors) represent the concepts of a theory, and data from measures (indicators) are used as input for statistical analyses that provide evidence about the relationships among latent variables. This chapter

*Corresponding author. Email: ljwilli1@vcu.edu

ISSN 1941-6520 print/ISSN 1941-6067 online

© 2009 Academy of Management

DOI: 10.1080/19416520903065683

<http://www.informaworld.com>



first provides a brief introduction to SEM and its concepts and terminology. We then discuss four issues related to the measurement component of such models, including how indicators are developed, types of relationships between indicators and latent variables, approaches for multidimensional constructs, and analyses needed when data from multiple time points or multiple groups are examined. In our second major section, we focus on six issues related to the structural component of structural equation models, including how to examine mediation and moderation, dealing with longitudinal and multilevel data, issues related to the use of control variables, and judging the adequacy of models and latent variable relationships. We conclude with a set of recommendations for how future applications of SEM in management research can be improved.

Introduction

Models of the research process presented in research methods textbooks emphasize two related components (Schwab, 2005). The first reflects the need to have operational measures of the conceptual variables described by the theory being examined, while the second reflects the use of these measures to test for the relationships between the conceptual variables as hypothesized by the theory. Historically in management research, a typical approach to this research process is to separate these two components. For example, a researcher might use exploratory factor analysis to develop evidence that the measures properly reflect the underlying constructs, and then subsequently create scales that might be used in linear or logistic regression to identify significant predictors as proposed by the theory. Within such an approach, there is a separation of the model and analysis that links the measures to their proposed underlying constructs and the model and analysis that examines relationships between the underlying constructs.

Since the early 1980s a large segment of management research has used structural equation modeling (SEM) as an analytical approach that combines these two components and considers them simultaneously. Thus, SEM is often described as combining factor analytic and regression models into a single data analysis tool. Using the language of SEM, latent variables (factors) represent the concepts of the theory, and data from measures (indicators) are used as input for statistical analyses that provide evidence about the relationships of the latent variables with their indicators and relationships among the latent variables. Reviews of this technique have appeared periodically, including James and James (1989), Harris and Schaubroeck (1990), Medsker, Williams, and Holahan, (1994), Williams and James, (1994), Williams, Edwards, and Vandenberg (2003), Williams, Gavin, and Hartman (2004); and Shook, Ketchen, Hult, and Kacmar (2004). Management studies using SEM have increased dramatically in frequency since the early 1980s. There are actually several types of SEM models, one of which will be the focus of this chapter, and

of this type of application the number of published studies in selected areas of management has increased from nine articles during 1978–1987, to 28 during 1988–1993, to 91 during the most recent time period of 2001–2008 (James & James, 1989; Medsker et al., 1994; Williams & O’Boyle, 2009).

A key motivation underlying the present chapter is our beliefs that (a) there are many misunderstandings about various issues concerning SEM among management researchers; (b) there are many shortcomings in the execution of SEM in articles in our top substantive outlets; and (c) substantive researchers in management will benefit from having one source that, at least at the time of publication, will describe what they need to do to get the most out of their SEM analyses. Unlike a previous publication by the same authors (Williams et al., 2003), which was basically descriptive, the present chapter will be much more prescriptive, providing technical guidance based on our experiences in conducting and reviewing SEM papers. Our hope is that with this chapter management researchers will be better able to realize the full benefits of the powerful SEM approach.

To accomplish our objectives, we will first present a simple Example Model for those less familiar with SEM terminology and techniques. This model will be described as consisting of two components, a measurement model and a structural model. Next, several current topics related to the measurement model will be discussed, including the selection and/or development of indicators to represent the latent variables, the types of relationships possible linking latent variables to indicators, approaches for examining latent variables that are multidimensional in nature, and the evaluation of measurement models when data from more than one group or more than one time period are being analyzed. From there we will switch to several current topics related to the structural model, including how to approach mediation and moderation hypotheses, analyses useful for investigating change with longitudinal data, models appropriate with data that are nested by level consisting of individuals organized within groups, how to incorporate control variables in an SEM analysis, and how best to evaluate the adequacy of an SEM model of latent variable relationships.

Example Model

A basic latent variable structural equation model that will be used to introduce methodological issues associated with SEM applications in management research is shown in Figure 12.1. Several aspects of the traditional notation and terminology are illustrated with this figure using the labels associated with the popular LISREL program (Jöreskog & Sörbom, 1996). A circle is used to represent each of four latent variables; the boxes represent associated indicator variables. The relationships between the latent variables and their indicators are often referred to as a *measurement* model, in that it represents or depicts an assumed process in which an underlying construct determines or

causes behavior (e.g., response to a survey question) 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. Thus, each factor serves as an independent variable in the measurement model (since the arrow goes from the circle to the box). The indicator variables serve as the dependent variables, and paths from latent variable to indicators are often referred to as factor loadings (λ). Each indicator is also potentially influenced by a second independent variable in the form of measurement error that contributes to its unique variance, which is comprised of two parts: systematic variance and random error variance. This measurement error influence (ϵ, δ) is represented as a cause of the indicator variable through the use of a second arrow leading to each of the indicators. It should be noted that while our emphasis is on models with multiple indicators for each latent variable, models with only a single indicator are available (referred to as total aggregation with reliability correction models—see Rogers and Schmitt (2004)).

In addition to proposing links between the indicators and the latent variables, the model shown in Figure 12.1 also depicts relationships among the four latent variables. The part of the overall model that proposes relationships among the latent variables is often referred to as the *structural* model. The model includes a correlation (double headed arrow— ϕ) between the two exogenous latent variable (ξ_1, ξ_2), two regression-like structural parameters (γ) linking the exogenous latent variables with two endogenous latent variables (η_1, η_2), and a similar regression-like structural parameter (β) linking the two endogenous latent variables. Finally, the model also acknowledges that there is unexplained residual variance in the two endogenous latent variables (ζ).

To evaluate the model shown in Figure 12.1, a management researcher would begin with a covariance matrix from a given data set among the measures being used as indicators. Given the model specification, which describes

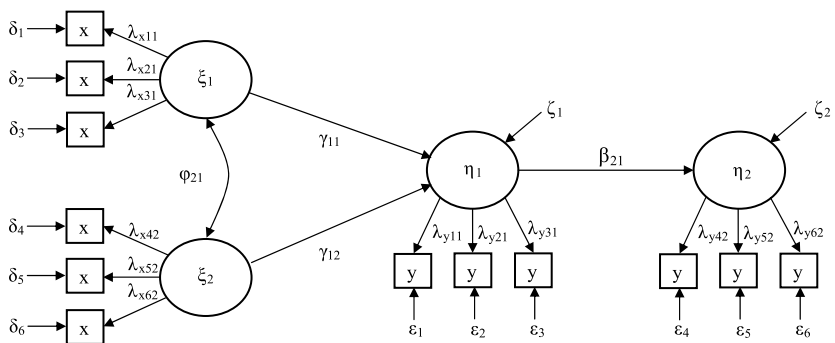


Figure 12.1 Example Model.

specific relationships based on theory among latent variables and indicators and among latent variables, estimates of the parameters mentioned previously (λ , ϵ , δ , ϕ , γ , β , and ζ) are obtained using specialized software (e.g., LISREL, MPlus, AMOS, EQS, etc). The most commonly used parameter estimation procedure is maximum likelihood, which also provides standard errors of the parameter estimates which can be used for testing null hypotheses that the estimates are statistically different from zero. For these estimates to be obtained, scales must be set for the latent variables, and this is typically achieved by fixing values of either factor variances or of one indicator for each latent variable to the value of 1.0. The software program must also converge through an iterative estimation process and arrive at a final set of parameter estimates, which must be within acceptable ranges to be interpretable (e.g., the results from a model with negative error variance estimates would generally not be trusted).

Based on the model specification and the parameter estimates, a variety of measures of model fit reflecting the adequacy of the model and a selection of model diagnostics are also obtained, including a chi-squared statistic with degrees of freedom equal to the number of unique elements in the covariance matrix minus the number of parameters estimated. These various fit measures ultimately reflect the similarity between the sample covariance matrix and a predicted covariance matrix calculated by the software program using the model estimates obtained with the analysis. Finally, the researcher also might compare the model with an alternative model using the same data, such as an alternative model that includes additional direct paths from the two exogenous variables to η_2 , and then compare the two models using a chi-squared difference test. This test directly examines the null hypothesis that the two direct paths originally not included are statistically different from zero. More details on the basics of SEM can be found in any of several introductory texts, such as Kline (2005).

Measurement Model Issues

Developing Indicators: Items and Parcels

A researcher wanting to examine a model such as the one shown in Figure 12.1 would need first to consider the types of measures to be used as indicators of the latent variables. In terms of the types of data, there are some published studies from the management literature in which objective measures (e.g., age, income) have been used as indicators (Baum & Locke, 2004; Frese et. al., 2007; Henderson, Berry, & Matic, 2007; Mesquita & Lazzarini, 2008; Walters & Bhuian, 2004) . However, a more typical application of SEM would be based on questionnaire data, in which existing established scales are completed by managers and/or their employees to obtain their perceptions about various aspects of their work environment. For these

types of studies, the researcher needs to decide how the indicators will be established or developed for each latent variable.

In these situations, the researcher can choose to use individual questionnaire items as indicators (referred to as *total disaggregation*), or can instead combine items from each scale into subsets called parcels and use these as indicators of the latent variable (referred to as *partial disaggregation*). Examples of disaggregation include studies of entrepreneurship, innovativeness, learning, and cycle time (Hult, Ketchen, & Nichols, 2002) and entrepreneurship, innovativeness, and market/customer orientation (Hult, Snow, & Kandemir, 2003), as well as research on dimensions of organizational justice, leader–member exchange, perceived organizational support, pay satisfaction, job control, and job strain (Roch & Shanock, 2006; Elovainio, Kivimäki, & Helkama, 2001). Recent examples of partial disaggregation and the use of parcels include studies of psychological empowerment (Alge, Ballinger, Tangiraola, & Oakley 2006), and procedural justice (Aryee, Chen, Sun, & Debrah, 2007). Parcels have also been used to study job search behaviors (Brown, Ferris, Heller, & Keeping (2007), performance appraisal (Elicker, Levy, & Hall, 2006), and team empowerment (Mathieu, Gilson, & Ruddy, 2006).

Relevant for management research is a comparison of the use of items as indicators (i.e., total disaggregation) with combining sets of items to form parcels (i.e., partial disaggregation). Coverage of relevant statistical and measurement issues have been provided by Bagozzi and Heatherton (1994) and Bagozzi and Edwards (1998), along with more recent writing by Little, Cunningham, Shahar, and Widaman (2002), Rogers and Schmitt (2004), Coffman and MacCallum (2005) and Williams and O’Boyle (2008). The use of items has the key advantage that it provides information about each individual item or question used in the SEM analysis. The most important information about the items focuses on the strength of the relation between the latent variable and the item indicator, including the estimates of the standardized factor loadings and error variances, as well as the squared multiple correlation for the item that reflects the amount of variance of the item associated with the latent factor. These three diagnostics are related to each other by psychometric theory and will converge to the same perspective as to whether an item is a good indicator. From a more philosophical perspective, items may be preferred because they are as close to the response of the individual as possible (as compared to combinations of items). This has been referred to as the empiricist–conservative position by Little et al (2002).

However, management researchers should be aware of some of the problems associated with using items as indicators of latent variables. As summarized by Little et al. (2002), Coffman and MacCallum (2005), and Rogers and Schmitt (2004), items will typically have lower reliability and communality, and a smaller ratio of common-to-unique variance than parcels or scales, which can limit their effectiveness as indicators. Second, the use of items can

increase the chances for correlations among uniqueness estimates due to shared error variance (for example due to item wording effects) and for measurement contamination that could cause dual factor loadings to emerge during preliminary exploratory factor analysis. Third, items may be more likely than parcels or scales to be non-normally distributed, which may violate normality assumptions associated with most SEM parameter estimation procedures. Another concern associated with the use of items as indicators includes the number of parameters required to represent a total disaggregation model as compared to partial aggregation models, given the number of factor loadings and error variance estimates (sample size recommendations often are based on the number of parameters estimated). Finally, using items (as compared to parcels) will result in a larger covariance matrix, which makes it less likely that the model will fit well even if it closely matches the process being studied.

As noted earlier, one alternative to the use of items and the total disaggregation approach involves combining items to create parcels which are then used as indicators. This approach has become increasingly popular in recent years (see review by Williams & O'Boyle, 2008). Little et al. (2002) describe arguments for the use of parcels from a pragmatic-liberal position, which asserts that solid and meaningful indicators (as obtained by combining items) of core constructs will replicate well across samples and studies. Advantages of the use of parcels described in previous reviews (Coffman & MacCallum, 2005; Little et al., 2002; Rogers & Schmitt, 2004; Williams & O'Boyle, 2008) include the fact that parcels will have more intervals between scale points as compared to items. Second, fewer parameters will be estimated when parcels are used, which can be important if the sample size is small. Third, it has also been suggested that there will be fewer chances for correlations among uniqueness estimates when parcels are used. Finally, the use of parcels may be preferred to data transformations or the use of more complicated parameter estimation procedures if there are normality problems or dichotomous items. Among the disadvantages of parcels discussed in previous reviews is that if they are used with multidimensional constructs, the parcels may be multidimensional, resulting in the latent variable being multidimensional, which can create problems in interpreting estimated relations among the latent variables (since the meaning of the latent variable is less clear given its multidimensional nature). The use of parcels has also been described as potentially hiding important sources of model misspecifications, and in fact, mis-specified models based on parcels may occasionally show better fit than correctly specified models assessed with items as indicators.

A management researcher using parcels would need to decide how to combine the original items from each scale to form the multiple parcel indicators used to represent the latent variable. Williams and O'Boyle (2008) reviewed strategies that have been used by management researchers in creating parcels,

drawing in part on earlier discussions by Little et al. (2002). Before describing these strategies, it should be emphasized that before forming parcels a researcher must consider evidence that the total set of items from a scale that will be grouped to form parcels for a latent variable is unidimensional (i.e., is composed of one factor). The use of parcels is most appropriate when such evidence is favorable (although we will discuss approaches to forming parcels given multidimensionality in this section and a preferred approach to multidimensionality in a later section).

Perhaps the most attractive approach to forming parcels has been labeled *item-to-construct balance* by Little et al. (2002) or the *factorial algorithm* by Rogers and Schmitt (2004). This approach is based on a goal of using parcels that are equally balanced in terms of their difficulty and discrimination, and it uses standardized factor loadings from a single-factor model for each latent variable that includes all scale items. This approach works toward balancing the best and worst items across the parcels so that no single parcel has all the good or bad items. Recent research comparing this approach to others that are available has found that with a unidimensional set of items, the various approaches result in similar parameter estimates, but the factorial method may be slightly preferred because it is superior on model fit criteria (Rogers & Schmitt, 2004).

Management researchers often investigate latent variables that are multidimensional with multiple facets (e.g. organizational commitment). In this situation, using parcels instead of single items may be preferred by some management researchers with models that include a large number of latent variables, or when researchers are not interested in investigating differential relationships involving antecedents or consequences of the multiple facets. A first approach to forming parcels in this type of situation has been attributed to Kishton and Widaman (1994) and discussed by Little et al. (2002) as the *internal-consistency approach*; it would be implemented by combining items from each facet to form the parcels. A second strategy for forming parcels with this multidimensional example is defined as the *domain-representative approach*, through which parcels are created by combining items from across facets to form the parcels. For example, using an organizational commitment scale, each parcel would include one or more items measuring belief in organizational goals, willingness to exert effort for the organization, and desire to remain in the organization.

In terms of a comparison of these two approaches to forming parcels among a set of multidimensional items, the internal consistency parcels would tend to have high internal consistency reliability estimates, because the items that load together would be fairly highly correlated, thus maximizing the value of alpha. However, one problem with this approach is that the resulting indicators only represent one component of the latent variable, which is counter to the objective of having indicators that can serve as stand-alone measures of

the entire latent variable. A second problem with this approach is that the resulting correlations among the indicators are relatively lower, because the sets of items forming the indicators were distinct enough that they factored separately in the initial exploratory factor analysis. Williams and O'Boyle (2008), among others, recommend that in this context the domain representative model be strongly considered, in which indicators are formed so that each contain one or more items from each of the multiple dimensions.

Indicator–Latent Variable Relationships: Formative vs. Reflective Measurement

A second aspect of the development of the measurement model involves the direction of the relationship between the latent variables and their indicators, and two general approaches are available. In management research, measures are usually treated as *reflective*, meaning they are observed manifestations of unobserved latent constructs (Diamantopoulos & Siguaw, 2006; MacKenzie, Podsakoff, & Jarvis, 2005; Podsakoff, Shen, & Podsakoff, 2006). Reflective measurement is firmly rooted in management research, as evidenced by the widespread use of factor analysis and structural equation models that specify measures as outcomes of latent variables, as shown in Figure 12.1. To facilitate our presentation, another such model is shown in Figure 12.2a, in which the latent variable has three reflective indicators labeled x_1 , x_2 , and x_3 . Substantively, ξ might signify overall job satisfaction, and x_1 , x_2 , and x_3 could represent scores on the items “In general, I am satisfied with my job”, “All in all, the job I have is great”, and “My job is very enjoyable” (Edwards & Rothbard, 1999). A key point is that the arrows representing the factor loadings (λ) go from the latent variable to the indicator.

Reflective measures can be contrasted with *formative* measures, which are assumed to form or produce an underlying construct. A formative measurement model is shown in Figure 12.2b, where the three measures x_1 , x_2 , and x_3 are specified as correlated causes of the latent variable η . This model is based on an example of formative measurement offered by MacKenzie et al. (2005), in which the latent variable η is contextual job performance (CJP) and x_1 , x_2 , and x_3 are measures of task performance, job dedication, and interpersonal facilitation. A key point is that now the arrows go from the indicator to the latent variable (γ_1 – γ_3). The distinction between formative and reflective measures has attracted increasing attention in the structural equation modeling literature (Bollen & Lennox, 1991; Edwards & Bagozzi, 2000; MacCallum & Browne, 1993), and some researchers have argued that many organizational measures treated as reflective are better specified as formative (Diamantopoulos & Siguaw, 2006; MacKenzie et al., 2005; Podsakoff et al., 2006).

Within the area of leadership, Podsakoff, MacKenzie, Podsakoff and Lee (2003) have argued that measurement model misspecification is widespread. Specifically, they examined the measurement approaches taken with 138 leadership constructs from 47 studies published in three leading journals that

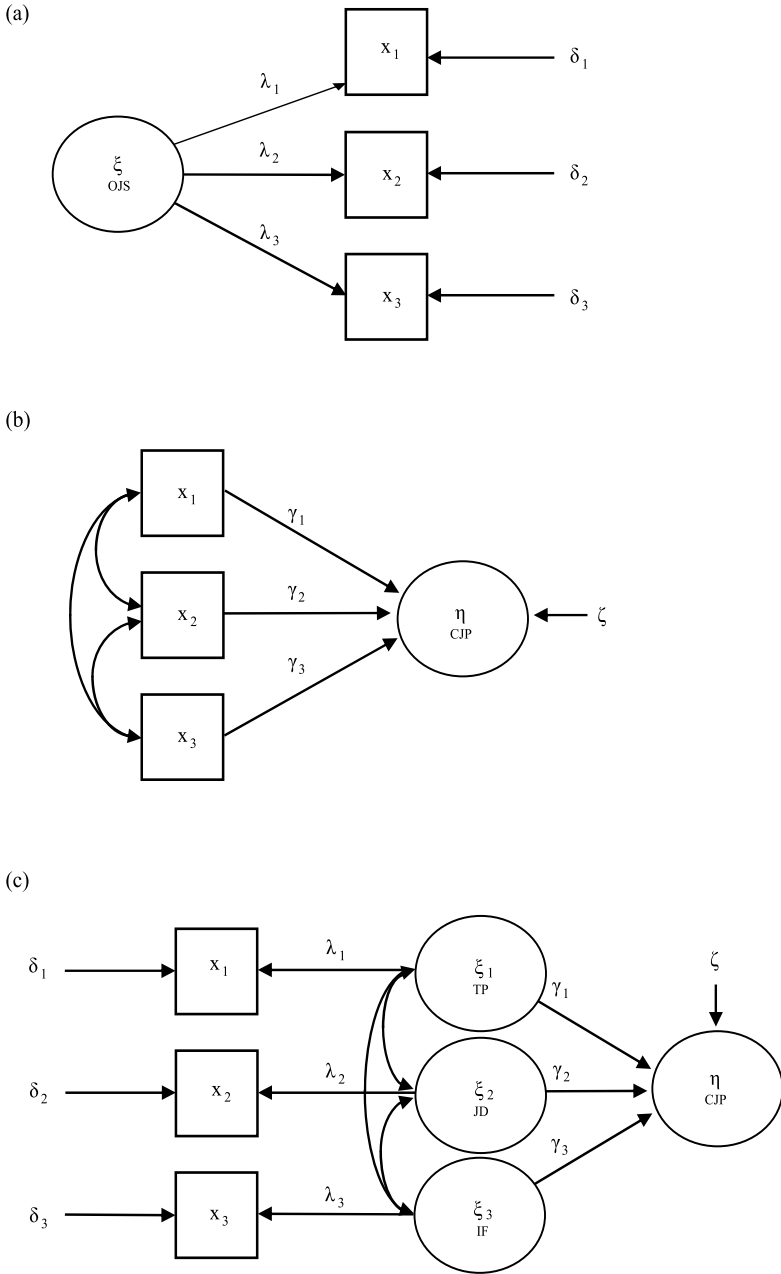


Figure 12.2 (a) Reflective Measures; (b) Formative Measures; (c) Formative Measurement Using Latent Variables.

Note: OJS, overall job satisfaction; CJP, contextual job performance; TP, task performance; JD, job dedication; IF, interpersonal facilitation.

published leadership research. Using several criteria, Podsakoff et al. (2003) concluded that 47% were incorrectly specified by treating formative indicators as reflective indicators. Similarly, in a review of strategic management studies published in top journals between 1994–2003, Podsakoff et al. (2006) examined 257 constructs from 45 published articles, focusing on conceptual issues, correlations among indicators, factor loadings for indicators, and reliabilities of indicators. Based on this information, Podsakoff et al. (2006) claim that 62% of the constructs were inappropriately modeled as having reflective indicators, when they should have been modeled as having formative indicators.

A management researcher must choose between reflective and formative measures, and thorny issues are raised concerning model specification and interpretation. In particular, identification problems will occur with formative measurement models such as the one shown in Figure 12.2b, unless the latent variable directly or indirectly causes at least two indicators in addition to the ones used as antecedents of the latent variable (Bollen & Davis, 1994; MacCallum & Browne, 1993). This is problematic because, as noted earlier, a unique set of parameter estimates cannot be obtained unless the model and its parameters are identified. To address identification concerns, the three measures of contextual job performance could be supplemented by adding general measures such as “My subordinate performs his or her assigned tasks well” and “My subordinate fulfills his or her job responsibilities”, with these measures added as reflective indicators of the contextual job performance construct η . Alternately, the model in Figure 12.2b could be expanded to include outcomes of contextual job performance, such as salary and advancement, each with their own reflective indicators. In either case, a management researcher who wants to estimate formative measurement models must revisit the theoretical justification for such models to locate reflective measures or causally dependent constructs, even if none were included in the original conceptualization of the model.

The parameters of reflective and formative measurement models also require different interpretations by management researchers. For reflective models, the loadings of the measures on the construct (e.g., λ_1 – λ_3) are driven by the covariances among the measures, such that higher loadings mean the measures have more in common with one another. The error terms (δ_1 – δ_3) represent aspects of each measure not related to the construct (e.g., random measurement error, item specificity), and lower error variances indicate higher reliabilities for the measures. In formative models, the relationships among the measures are absorbed by covariances that directly relate the measures to one another, as indicated by the curved arrows among the indicators in Figure 12.2b (MacCallum & Browne, 1993), and the loadings of the measures on the construct (γ_1 – γ_3) are heavily influenced by the relationships between these measures and the reflective indicators added to identify the

model. As a result, the loadings for formative measures can vary dramatically depending on the reflective indicators chosen for a particular study (Howell, Breivik, & Wilcox, 2007). In addition, with formative measurement the residual assigned to the construct (ζ) does not represent measurement error, but instead signifies aspects of the construct not associated with its measures (Diamantopoulos, 2006; Jarvis, MacKenzie, & Podsakoff, 2003; MacKenzie et al., 2005). Measurement errors in the formative indicator measures themselves are not taken into account by the model.

Perhaps the most fundamental issue underlying the choice between reflective and formative measurement models involves the causal relationship between constructs and measures (Edwards & Bagozzi, 2000). As indicated by Figure 12.2a, reflective measurement models specify constructs as causes of measures. This notion makes sense from a critical realist perspective (Borsboom, Mellenbergh, & van Heerden, 2004; Cook & Campbell, 1979; Loevinger, 1957; Messick, 1981), which claims that constructs refer to real entities but recognizes they cannot be measured directly or infallibly. This perspective readily applies to constructs examined in management research. Returning to Figure 12.2a, if ξ is job satisfaction and x_i is a score on the item "In general, I am satisfied with my job", it stands to reason that job satisfaction is a real experience in the mind of the respondent that causes him or her to report a particular score. More generally, when management researchers develop theories to explain relationships among constructs, the constructs presumably refer to real entities that exist regardless of whether they happen to be measured by researchers. If constructs did not exist in some real sense, they could not influence one another, and attempts to develop theories that explain relationships among constructs would be pointless. When management researchers collect measures, they gather scores about events and conditions that exist in organizations. These scores then become input into data analysis, such as structural equation modeling, and constitute the boxes in measurement models such as those in Figures 12.1 and 12.2a. From this perspective, constructs refer to real phenomena that produce the scores collected by researchers, and causality runs from constructs to measures, consistent with reflective measurement.¹

In contrast to reflective measurement, formative measurement implies that measures cause constructs, as indicated in Figure 12.2b. The idea that measures cause constructs is difficult to reconcile with the notion that measures are merely observed scores (Edwards & Bagozzi, 2000; Howell et al., 2007; Messick, 1995). When measures are specified as causes of constructs, as in formative measurement models, a subtle form of operationalism is invoked in which measures are equated with specific constructs believed to cause a more general construct (Edwards & Bagozzi, 2000). Said differently, scores are data that reside in the files researchers use for analysis. These scores have no causal potential of their own, but instead are quantitative evidence of causal processes

that empirical research attempts to capture (Borsboom et al., 2004; Howell et al., 2007). Thus, when measures of task performance, job dedication, and interpersonal facilitation are specified as formative indicators of contextual job performance, the scores on these measures are treated as if they have causal potential.

However, contextual job performance does not result from these scores, but instead arises from the constructs these scores represent. From this perspective, the measurement of contextual job performance should be recast as the model in Figure 12.2c, where x_1 , x_2 , and x_3 are reflective measures of the constructs task performance (TP), job dedication (JD), and interpersonal facilitation (IF), as represented by ξ_1 , ξ_2 , and ξ_3 , which in turn are causes of contextual job performance as symbolized by η . This perspective applies to virtually every case of formative measurement we have encountered, in which measures are better conceived as reflective indicators of specific constructs that in turn cause general constructs. The effects of specific constructs on general constructs are best described within the realm of structural models that relate constructs to one another (Figure 12.2c), as opposed to measurement models that relate constructs to measures (Figure 12.2b).

If causality runs from constructs to measures, then what explains the appeal of formative measurement models among some current management researchers? Perhaps the attraction of formative measurement lies in the apparent parsimony of combining distinct measures into larger wholes, along with the sense that such measures are necessary to adequately cover the content domain of a construct (Diamantopoulos & Winklhofer, 2001; Nunnally & Bernstein, 1994). However, if these measures describe distinct aspects of a construct, then each measure represents a construct in its own right, and each of these constructs should be represented by multiple reflective measures (Bollen & Lennox, 1991; Edwards & Bagozzi, 2000; MacKenzie et al., 2005; Podsakoff et al., 2006). Such relationships between the specific and general constructs underlying these measures can be examined within the context of multidimensional construct models, which we now consider.

Relationships among Multidimensional Constructs

A third issue related to measurement models of importance to management researchers concerns the treatment of multidimensional constructs. As noted in the earlier section on parcels, although it is possible to model a multidimensional construct as a single latent variable, in most situations we feel there are advantages to maintaining distinctions between dimensions and modeling relationships among latent variables representing the multiple dimensions. Frequently used examples of multidimensional constructs include organizational citizenship behavior defined in terms of conscientiousness, altruism, sportsmanship, courtesy, and civic virtue (Organ, 1988), transformational leadership conceptualized as idealized influence, inspirational motivation,

intellectual stimulation, and individual consideration (Bass & Riggio, 2006), empowerment defined as meaning, competence, self-determination, and impact (Spreitzer, 1995), and organizational justice distinguished according to distributive, procedural, interpersonal, and informational justice (Colquitt, 2001). The dimensions of these types of multidimensional constructs can themselves be conceived as constructs that are more specific than the broader multidimensional construct itself. As such, multidimensional constructs and their dimensions are latent variables that can be easily accommodated by structural equation models (Edwards, 2001).

Management researchers considering the use of multidimensional constructs in their models confront several issues as they decide how to best represent these constructs with latent variables. One concern is the direction of the relationships between the multidimensional construct and its dimensions. It is possible that the multidimensional construct is a higher-order factor, represented with paths running from the multidimensional construct to the dimensions. Such constructs have been called latent (Law, Wong, & Mobley, 1998) or superordinate (Edwards, 2001), although we prefer the latter term because a multidimensional construct is a latent variable regardless of how it relates to its dimensions. An example superordinate construct model is shown in Figure 12.3a. This model is consistent with the conceptualization of empowerment (EMP) developed by Spreitzer (1995), in which meaning (MNG), competence (COM), and self-determination (DET) might be treated as three of its dimensions assigned to a higher-order empowerment construct. The model contains paths (γ_1 - γ_3) running from the construct (ξ) to its dimensions (η_1, η_2, η_3), and a residual (ζ) is assigned to each dimension. These residuals represent specific aspects of each dimension that are unrelated to the construct. The dimensions, now serving as endogenous latent variables, have indicators specified as reflective measures (y_1 - y_9), each of which has a residual that represents measurement error (ϵ).

Another possibility for management researchers to consider is that the multidimensional construct is a function of (rather than a cause of) its dimensions, as depicted by the aggregate construct model in Figure 12.3b (Edwards, 2001; Law et al., 1998). This model is consistent with arguments that transformational leadership (TLD) should be viewed as a higher-order construct that, in this case, results from inspirational motivation (MOT), intellectual stimulation (IST), and individual consideration (CON), each of which represents a first-order subdimension measured by multiple items (MacKenzie et al., 2005; Podsakoff et al., 2003). In this model, the paths (γ_1 - γ_3) run from the dimensions (ξ_1, ξ_2, ξ_3) to the aggregate construct (η_1), which is assigned a residual (ζ) that represents aspects of the construct not related to its dimensions. The dimensions are allowed to covary with one another (as represented by the curved double head arrows connecting them), and as before, each dimension has reflective measures (x_1 - x_3) that contain random measurement error (ϵ).

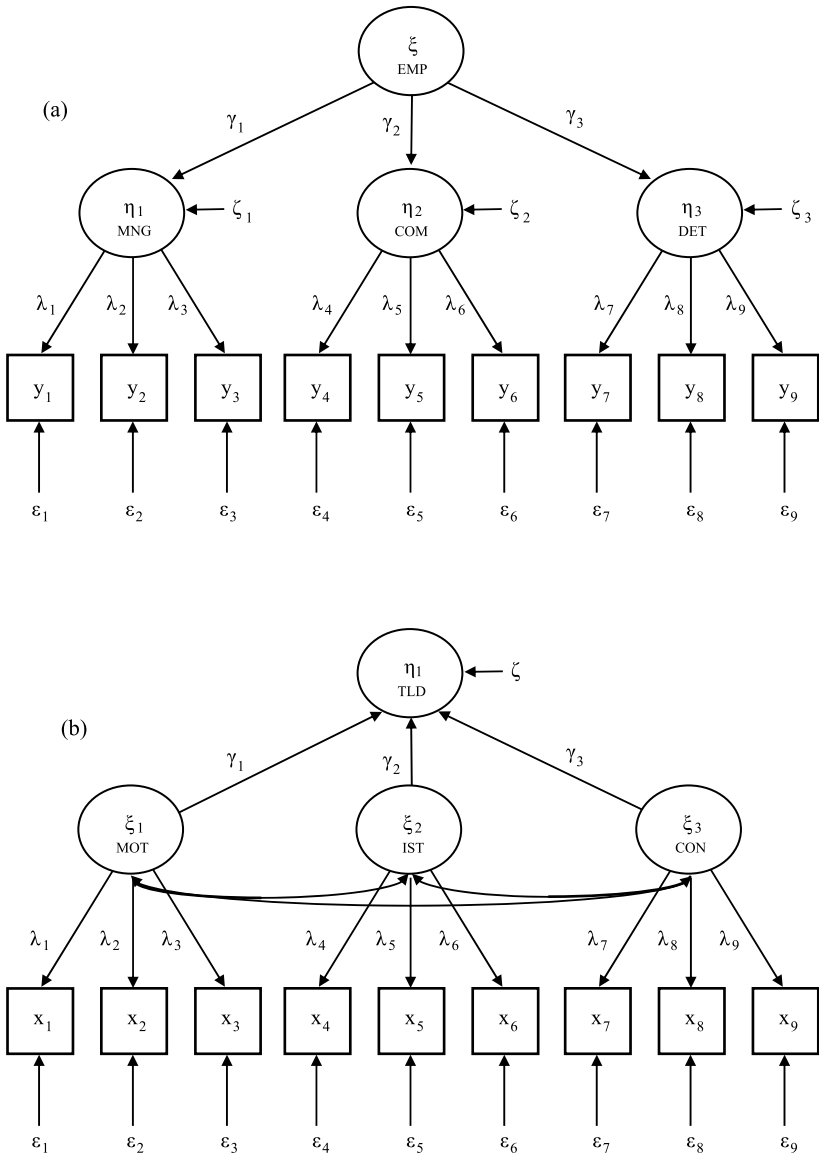


Figure 12.3 (a) Superordinate Construct Model; (b) Aggregate Construct Model.

Note: EMP, empowerment; MNG, meaning; COM, competence; DET, self determination; TLD, transformational leadership behavior; MOT, inspirational motivation; IST, intellectual stimulation; CON, individual consideration.

The choice between superordinate and aggregate construct models has fundamental implications for the meaning of the multidimensional construct and can dramatically impact parameter estimates linking the multidimensional construct to its dimensions (Edwards, 2001).

Another issue important to management researchers as they investigate multidimensional latent variables involves the ontological status of the constructs in their models. In many cases, the multidimensional construct is *defined* in terms of its dimensions (Edwards, 2001). In such cases, the multidimensional construct does not exist apart from its dimensions, but instead is an abstract concept that refers to the dimensions as a set. When multidimensional constructs are conceived in this manner, the relationships between the construct and its dimensions are not causal, because a causal relationship requires that the entities involved are distinct (Edwards & Bagozzi, 2000). Instead, the relationships are functional, specified according to the definition of the multidimensional construct. Alternatively, if the multidimensional construct refers to an entity that exists separately from its dimensions, then the relationships linking the multidimensional construct to its dimensions can be considered causal, and the directions of the relationships can be deduced by applying conditions for causality (Cook & Campbell, 1979; Pearl, 2000; Sobel, 1996).

For instance, if the multidimensional construct is a general personality trait that leads to specific individual differences (Costa & McCrae, 1995; Murtha, Kanfer, & Ackerman, 1996), then the relationships between the construct and its dimensions are captured by the superordinate construct model as in Figure 12.3a. In contrast, if the multidimensional construct refers to overall job satisfaction that results from satisfaction with specific job facets (Aldag & Brief, 1978; Locke, 1976), then the relationships between the construct and its dimensions follow the aggregate construct model as in Figure 12.3b. When multidimensional constructs exist separately from their dimensions, it is often possible to supplement measures of the dimensions with direct measures of the multidimensional construct, as when measures of job facet satisfaction are accompanied by measures of overall job satisfaction (Ironson, Smith, Brannick, Gibson, & Paul, 1989) or measures of distributive, procedural, interpersonal, and informational justice (Colquitt, 2001) are supplemented by measures of overall justice (Ambrose & Schminke, 2009). With respect to Figure 12.3a, direct measures of EMP could be added, while for Figure 12.3b direct measures of TLD could be added. In these cases, the researcher can investigate causal relationships among the multidimensional construct and its dimensions, which is often desirable.

Other issues pertain to the identification and interpretation of multidimensional construct models. In general, superordinate construct models such as the model in Figure 12.3a are identified if the construct has at least three dimensions (in this case MNG, COM, DET), each of the dimensions has at

least three indicators, and identification is achieved by setting the scales for the superordinate construct and its dimensions. In contrast, the aggregate construct model in Figure 12.3b is not identified unless the aggregate construct (η) is specified as a cause of at least two additional constructs with reflective measures or the aggregate construct is assigned at least two reflective indicators of its own (Edwards, 2001). Conditions for identification are more easily met when multidimensional constructs have their own indicators, in which case conventional rules for identification can be applied (Bollen, 1989; Rigdon, 1995).

Superordinate and aggregate construct models also yield different interpretations of the multidimensional construct and its dimensions. A superordinate construct should be viewed as the commonality of its dimensions, given that the paths linking the construct to its dimensions are driven by the covariances among the dimensions, as in a higher-order confirmatory factor analysis. In contrast, an aggregate construct represents a weighted linear combination of its dimensions, as with a regression model where predicted values of y are obtained using the predictor x variables. If the residual for the aggregate construct is fixed to zero, then the construct has no meaning beyond its dimensions, whereas if the residual is freely estimated, the construct can include variance not represented by its dimensions. In some cases, the variance contributed by the residual can exceed that associated with the dimensions, which obscures the meaning of the aggregate construct. For a management researcher such a finding would make it difficult to interpret the relationship of the aggregate construct with other constructs in the model.

A final issue important for management researchers involves the evaluation of models that contain multidimensional constructs. Typically, a multidimensional construct serves as the proxy for its dimensions, such that the only paths linked to the dimensions are those that connect the dimensions to the multidimensional construct. In most cases, paths connecting the dimensions to other constructs can be examined by adding these paths to determine whether the relationships channeled through the multidimensional construct adequately capture the association between the dimensions and other constructs (which would not be the case if the added paths were significant). An exception involves models in which an aggregate construct is specified solely as an effect of other constructs, in which case the model obscures rather than omits the relationships between the dimensions and other constructs in the model (Edwards, 2001).

We have often found that models containing multidimensional constructs are inferior to models that treat the dimensions as a set and omit the multidimensional construct itself. Indeed, if a multidimensional construct is merely conceptual shorthand for discussing the dimensions as a set, then there is no need to invoke a multidimensional construct at all, and instead the dimensions can be viewed as elements of a category for which the “multidimensional

construct” is simply a convenient label. In this case, paths that capture relationships for the dimensions as a set can be represented by multivariate parameters, such as the total coefficient of determination (Cohen, 1982; Edwards, 2001; Jöreskog & Sörbom, 1996), thereby allowing inferences for the dimensions individually and jointly. However, if a multidimensional construct refers to an entity that exists separately from its dimensions, then it should be retained in the model, and researchers should consider assigning measures directly to the multidimensional construct, which would minimize the identification and interpretational issues previously discussed.

Indicator, Groups, and Time: Measurement Invariance

The fourth issue related to measurement models we will consider involves situations in management research where multiple groups are utilized or data are collected within the same sample across time. The three most frequent instances in which this occurs are: (a) when theory dictates that between-group differences exist and thus, the goal of the SEM analyses is to eventually test for those differences (Ployhart & Oswald, 2004; Vandenberg & Scarpello, 1990); (b) there are no theoretical expectations for group differences, but the same data are collected from samples representing fundamentally different populations such as different organizations or cultures (Riordan & Vandenberg, 1994); or (c) the same data are collected from the same group or groups on multiple occasions (Bentein, Vandenberg, Vandenberghe, & Stinglhamer, 2005; Lance, Vandenberg, & Self, 2000; Vandenberg & Self, 1993). In these instances, the equality or invariance of measurement model parameters across time or between groups is of great importance, and investigators should strongly consider additional steps for their analyses.

Over the years there have been many management studies that have included an examination of measurement invariance. In a review and synthesis of this literature, Vandenberg and Lance (2000) identified 67 studies where invariance was investigated in conjunction with tests of substantive hypotheses. The topical areas of these studies included organizational change, measurement development, test administration modes, and cross-cultural generalizability of measures and models. More recent examples have investigated differences between applicants, incumbents, and students on personality measures (Smith, Hanges, & Dickson, 2001), and differences in performance ratings provided by self, peers, subordinates, and supervisors (Woehr, Sheehan, & Bennett, 2005; Fecteau & Craig, 2001). Cross cultural differences in transformational leadership (Schaubroeck, Lam, & Cha, 2007) and creativity (Hirst, Knippenberg, & Zhou, 2009) have also been examined.

As noted by Vandenberg and Lance (2000), and recently reinforced by Schmitt and Kuljanin (2008), and Thompson and Green (2006), additional steps beyond the normal ones involved with SEM (e.g., assessing model fit, significance of factor loadings, etc.) are needed where invariance is a concern.

The goal of these steps is to ensure that the properties of the underlying measurement model representing the constructs and indicators are equivalent or invariant across groups or time. These steps by necessity need to be completed prior to examining structural paths between the latent variables, and thus, are typically completed at the measurement model level with models including factor correlations (rather than structural paths).

For example, consider a management researcher who was interested in testing a model similar to the one shown in Figure 12.1 that proposes that high-involvement work processes (HIWP) mediate the effects of leader-member exchange (LMX) and perceived organizational support (POS) on turnover intentions (TI), and is examining the model with data collected from two different organizations. Before testing the full structural model and examining the three structural paths in the two groups, a multi-sample confirmatory factor analysis model should be examined, as shown in Figure 12.4. In addition to reflecting that data are being tested from two groups, the model in Figure 12.4 is different from the Example Model in Figure 12.1 in that it proposes that all four latent variables are simply correlated with each other (no causal structure is proposed among the latent variables), with the factor covariances represented as before (e.g. ϕ_{12}). This model is also different in that it includes parameters representing the factor variances (e.g., ϕ_1) and the item indicator intercepts (e.g., τ_{x11}), the latter of which can be interpreted as the value of the observed variable when the value for the latent variable is zero. It should be noted that most SEM software packages now provide estimates of these intercepts, as they play a key role in evaluating advanced types of advanced models (e.g., latent means, latent growth modeling).

Within this context, it is important for a management researcher to examine evidence as to whether the various properties of the measurement model involving the four latent variables are equivalent or invariant between the two groups. If evidence suggests that the measures are *not* equivalent, this can have negative implications for the validity of the inferences when the full model and its structural paths are examined. As to the analyses needed to investigate measurement invariance, there are six potential invariance or equivalence tests that can be conducted, and these have been described by Vandenberg and Lance (2000), Schmitt and Kuljanin (2008), and Thompson and Green (2006). These tests focus on whether parameter estimates are equal across the two groups, including (a) configural invariance (i.e., does the factor pattern matrix have the same form across groups); (b) metric invariance (i.e., are the factor loadings equal across groups); (c) intercept invariance (i.e., are the intercepts equal across the groups); (d) are the error-uniqueness estimates equal across groups; (e) are the factor variance estimates equal across groups; and (f) are the factor covariances equal. Each of these invariance tests would typically be undertaken on all parameters as a set in a particular test; that is, all 12 factor loadings would be constrained to be equal in the metric invariance test.

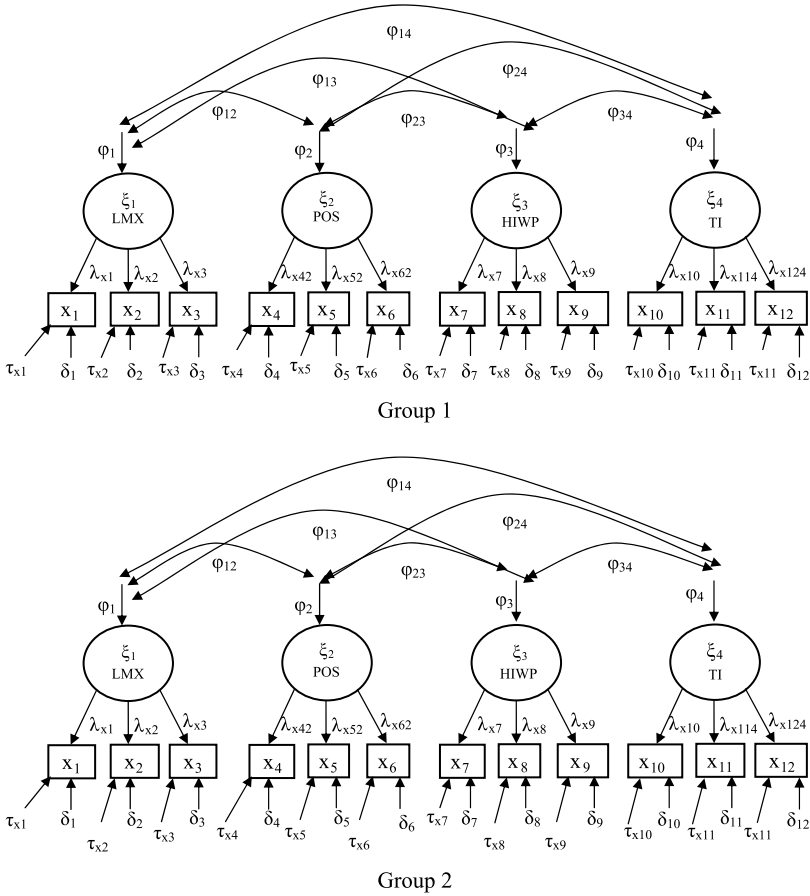


Figure 12.4 Two-group Measurement Model for Measurement Invariance.

Note: LMX, leader–member exchange; POS, perceived organizational support; HIWP, high-involvement work practices; TI, turnover intentions.

The configural (a) and metric invariance (b) tests are most critical, and as noted before they must be examined before investigating the structural parameters (γ , β) of the two groups. We will focus our comments now on issues that can emerge with these tests. A test of configural invariance, also referred to as a test of a weak factorial invariance (Horn & McArdle, 1992) examines whether the same a priori pattern of fixed and free factor loadings holds in each group. If configural invariance is not indicated, this means that one or more of the item sets represent different constructs between the groups. With respect to Figure 12.4, the test for configural invariance addresses whether members in Groups 1 and 2 use the same conceptual frames of reference to respond to the items representing each of the four latent

variables (LMX, POS, HIWP and TI) in the measurement model. Operationally, the same measurement model is specified for each group's data and are jointly analyzed using a multi-sample approach (i.e., a simultaneous CFA of both group's data). Only one set of fit indices is produced, and they are interpreted to infer whether the measurement model represents both groups' data equally well. If results for the fit indices are favorable, the conclusion is that the items were interpreted and responded to using the same constructs in each group. If the indices are unfavorable, it can no longer be safely assumed that items were interpreted in both groups using the same constructs (Vandenberg & Lance, 2000). If configural invariance is not supported, then the researcher should conclude that the hypotheses of differences for the structural parameters across groups cannot be tested using the current data because the evidence indicates that the latent variables and the indicators are not linked the same way in the two groups.

However, the decision about configural invariance should take into account the depth of the problem in the aberrant group. For example, assume upon further diagnosis that HIWP in Group 2 was problematic (i.e., its items had low loadings, an exploratory factor analysis of its items shows that the items load onto two factors, etc.). The researcher should then consider if, from a theoretical perspective, the test of the hypotheses (testing for group differences) can proceed without the HIWP mediator variable. The prospect of not being able to test the hypotheses of group differences increases if more than one latent variable is not configurally invariant (e.g., HIWP and POS) or there is lack of configural invariance in both groups (HIWP in both Group 1 and 2).

Assuming configural invariance has been achieved, the next key test is for metric invariance and examines the equality of factor loadings across groups. Using the same model specification as that for configural invariance (Figure 12.4), this test is undertaken by equating the factor loadings (λ_x) of like items between groups (Vandenberg & Lance, 2000) and it examines the null hypothesis that factor loadings for like items are invariant across groups. The term "metric" is used because the factor loading represents how strongly responses are calibrated to the latent variable. Specifically, factor loadings represent the strength of association between the observed item and its underlying latent variable. If a given loading has a value of 1.0, then it may be presumed that there is a one-to-one correspondence between the response scale of the item and the unobserved scale underlying the latent variable. However, to the degree the coefficient deviates from being 1.0, there is less correspondence between the response scale (or metric) of the observed item and that of the latent variable. Continuing with the example, assume that we tested for metric invariance in Figure 12.4, and failed to support it. Assume further that upon further analysis it was learned that the problem was isolated to λ_{x11} of Group 1 (0.85) not being statistically equal to λ_{x11} of Group 2 (0.50). These results would indicate that even though a Group 1 individual may give the same

response as someone from Group 2, the response value has a different meaning in terms of the corresponding value for the latent variable.

Testing the appropriateness of metric invariance is relatively straightforward (Vandenberg & Lance, 2000). It encompasses the use of both the fit indices of the metric invariance model itself and the chi-squared difference test between this model (with equality constraints between loadings of like items) and the configural invariance model (without the equality constraints). Support for metric invariance is observed when the fit indices for the metric invariance model are favorable *and* the chi-squared difference test between this model and the configural invariance model is not statistically significant. However, the chi-squared difference test is sensitive to the overall sample size, and it may be difficult to observe a non-significant difference when sample sizes are large across the groups. Therefore, the chi-squared difference test by itself is insufficient to conclude that metric invariance is not supported, and additional information should be considered. If the chi-squared difference is statistically significant, but all of the fit indices for the metric invariance model are favorable, then metric invariance is probably an acceptable conclusion. If, however, the chi-squared difference is relatively large and statistically significant, and the fit indices of the metric invariance model are less than favorable (even by a small amount), then the conservative and appropriate conclusion is to assume that metric invariance does not exist between the groups.

Assuming metric invariance has not been supported, the next step in the analysis is to determine the severity of the problem and consider the implications for the researcher's ability to test the model representing the structural hypotheses such as with the model in Figure 12.1. We recommend that management researchers consider the severity of the issue and determine if the lack of invariance is due to one item indicator from among the many items, or nearly all items. Also to be considered is whether it is limited to one particular measure or to most of the measures, and whether it is limited to one group, or whether there is some degree of invariance among items in all groups (if there are more than two groups). If the severity is relatively minor, then a partial metric invariance strategy may be undertaken (Vandenberg & Lance, 2000). For example, if the problem is limited to one item in one of the groups, the item could potentially be removed from the analysis without jeopardizing the construct validity of the measure to which it belongs. Alternatively, the equality constraint for that item could be relaxed and the loadings freely estimated for that item in each group. As noted by Vandenberg and Lance (2000), such a partial metric invariance strategy is not without controversy because it involves judgment calls and thus should be heavily driven by theoretical considerations.

As to implications for hypothesis testing purposes, this is dependent upon the degree of and pattern of invariance as discussed previously. If the lack of invariance is relatively minor, then the researcher can proceed with the tests of

the theoretical differences between groups (if differences are the focus of the hypotheses), or with collapsing the data from both groups together for hypothesis testing purposes (if group differences are not a focus of the research). If the problem is severe; that is, a large number of the 12 items in Figure 12.2, for example, are not metrically invariant or the lack of invariance is due to items of one scale in particular, then the next step depends upon how the groups were to be used for hypothesis testing purposes. If the data from both groups were to be collapsed together, then the researcher may need to simply use the group with the strongest loadings of the observed scores onto the latent variables and discard the other group. If, though, the hypotheses require a comparison between groups, then the prospect emerges that the researcher may have to collect additional data.

Structural Equation Model

Mediation and Latent Variable Relationships

Moving beyond models that focus on measurement issues, we now focus on structural relations representing proposed causal structures among latent variables, as exemplified by our Example Model in Figure 12.1. A first special model of this type we will consider is used often by management researchers and incorporates mediation, such that the effect of one latent variable on another is proposed to be transmitted through one or more additional latent variables. For example, the model in Figure 12.1 proposes that η_1 is a mediator of the effects of ξ_1 and ξ_2 on η_2 . This model proposed full mediation, in that ξ_1 and ξ_2 do not directly influence η_2 , but instead operate only indirectly through η_1 .

Mediation models are prevalent in nearly all areas of management research. Recent SEM examples include work that has examined psychological empowerment as a mediator of the effects of information privacy on organizational citizenship behavior and creative performance (Alge et al., 2006), and research on procedural justice as a mediator of effects of information privacy and outcome favorability on test taking motivation, organizational attraction, and intentions toward the organization (Bauer et al., 2006). Positive and negative sentiments have been examined as mediators of effects of subordinate charisma on interpersonal and informational justice (Scott, Colquitt, & Zapata-Phelan, 2007), as has leader-member exchange as a mediator of the impact of transformational leadership on organizational citizenship behavior and task performance (Wang, Law, Hackett, Wang, & Chen, 2005). Finally, perceived organizational support has been examined as a mediator linking perceived organizational support with commitment and performance (Pazy & Ganzach, 2008).

To further discuss mediation, we refer to a set of models in Figure 12.5.² A model without mediation is shown in Figure 12.5a, which indicates that ξ_1 ,

directly causes η_1 . To illustrate, consider the relationship between value congruence (VC) and organization commitment (OC). Most research has treated this relationship as a direct effect, consistent with the model in Figure 12.5a in which ξ_1 has a direct path to η_1 . This model can be contrasted with the model in Figure 12.5b, in which ξ_1 (VC) causes η_1 (OC) both directly and indirectly through the mediating variable η_2 (PRE) via the paths γ_{21} and β_{12} . This second model is based on the suggestion that the effect of value congruence is mediated by predictability, arguing that value congruence helps employees predict the behavior of others and coordinate their actions, which in turn increases commitment to the organization (Kalliath, Bluedorn, & Strube, 1999; Meglino & Ravlin, 1998).

A more complex mediated model is shown in Figure 12.5c, in which the effects of ξ_1 (VC) on η_1 (OC) are mediated by η_2 (PRE) as well as η_3 , which signifies communication (COM). This model is based on the proposal that value congruence fosters communication by giving employees a common frame for classifying and interpreting events (path γ_{31}), which in turn enhances

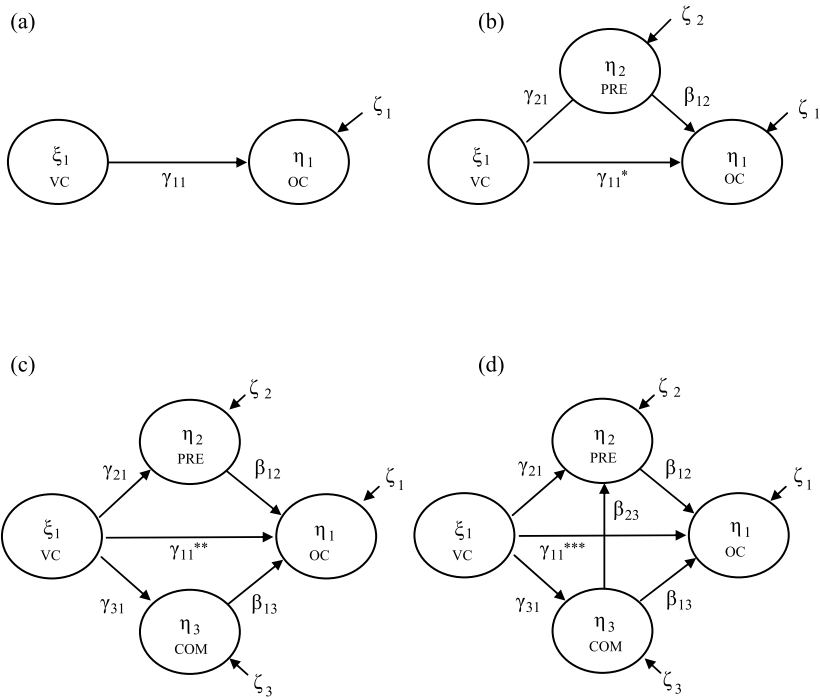


Figure 12.5 (a) No Mediation; (b) One Mediator; (c) Two Mediators; (d) Three-stage Mediation Model.

Note: VC, value congruence; OC, organizational commitment; COM, communication; PRE, predictability.

organizational commitment via the path β_{13} (Erdogan, Kraimer, & Liden, 2004; Meglino & Ravlin, 1998). Finally, the model in Figure 12.5d adds a path (β_{23}) from η_3 to η_2 , thereby incorporating a three-stage mediated effect from ξ_1 to η_3 (COM) to η_2 (PRE) to η_1 . This model is supported by additional research which suggests that open communication facilitates predictability (Reilly & DiAngelo, 1990; Schuler, 1979). These four models show that mediation can be partial or complete, involve multiple mediators, and can transmit effects through multiple paths. Thus, mediated models such as those in Figure 12.5 can be used to depict and test increasingly refined explanations of the relationships between theoretical constructs (Whetten, 1989).

Various procedures for analyzing mediation have been developed (Baron & Kenny, 1986; James & Brett, 1984; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). The most popular among management researchers is the causal steps procedure outlined by Baron and Kenny (1986), which states that mediation is established when four conditions are met: (a) the independent variable is related to the dependent variable; (b) the independent variable is related to the mediator variable; (c) the mediator variable is related to the dependent variable; and (d) when the mediator variable is statistically controlled, the independent variable is no longer related to the dependent variable. Referring to the models in Figure 12.5, the first condition stipulates that γ_{11} in Figure 12.5a is significant, the second and third conditions require that γ_{21} and β_{12} , respectively, in Figure 12.5b are significant, and the fourth condition specifies that γ_{11}^* in Figure 12.5b is not significant. Satisfying all four conditions means that η_2 fully mediates the effect of ξ on η_1 , whereas satisfying the first three conditions but not the fourth indicates that mediation is partial, such that ξ affects η_1 both directly and indirectly through η_2 (Baron & Kenny, 1986).

Despite its popularity, the causal steps procedure has several limitations that compromise the quality of inferences based on its use. First, requiring an initial relationship between the independent and dependent variable (step (a) earlier) can rule out models that have mediated and direct effects with opposite signs (MacKinnon et al., 2002). This point is clarified by noting that γ_{11} in Figure 12.5a, which is used to establish that the independent variable and dependent variable are related, is algebraically equivalent to $\gamma_{21}\beta_{12} + \gamma_{11}^*$ in Figure 12.5b. The term $\gamma_{21}\beta_{12}$ is the product of the paths that carry the mediated effect of ξ on η_1 through η_2 , and γ_{11}^* is the direct effect of ξ on η_1 after η_2 is controlled. Thus, if the mediated effect represented by $\gamma_{21}\beta_{12}$ is accompanied by a direct effect γ_{11}^* with opposite sign, the two effects can combine to yield a nonsignificant value for γ_{11} , even when the mediated effect is substantial. In this case, the absence of an overall relationship would invite the researcher to conclude that no mediation exists when, in fact, a mediated effect is present.

Second, the test of γ_{11}^* in the fourth step of Baron and Kenny (1986) has no direct bearing on the evaluation of the mediated effect represented by $\gamma_{21}\beta_{12}$. As

noted previously, γ_{11} in Figure 12.5a is equal to $\gamma_{21}\beta_{12} + \gamma_{11}^*$ in Figure 12.5b, which in turn implies that $\gamma_{21}\beta_{12} = \gamma_{11} - \gamma_{11}^*$. Thus, for a particular value of $\gamma_{21}\beta_{12}$, γ_{11}^* can take on any value, provided it is accompanied by a value of γ_{11} that satisfies the equality $\gamma_{21}\beta_{12} = \gamma_{11} - \gamma_{11}^*$. This equality further shows that the mediated effect $\gamma_{21}\beta_{12}$ is not captured by either γ_{11} or γ_{11}^* alone, which are the focus of the first and fourth steps, but is instead represented by the difference between γ_{11} and γ_{11}^* . So, an improper focus on the fourth step can also lead to an incorrect conclusion regarding mediation. Third, the tests involved in the causal steps procedure do not themselves yield a test of the mediated effect indicated by the product $\gamma_{21}\beta_{12}$ (MacKinnon et al., 2002), although Baron and Kenny (1986) describe how $\gamma_{21}\beta_{12}$ can be tested separately from the causal steps procedure itself. Finally, because it is framed around the basic mediated model in Figure 12.5b, the causal steps procedure provides little guidance for assessing mediation in more complex models, such as those in Figures 12.5c and 12.5d. The general conclusion is that reliance on the four steps of Baron and Kenny (1986) can lead to incorrect inferences regarding mediation and may not be effective with complex models common to management research.

Limitations of the causal steps procedure presented by Baron and Kenny (1986) have been addressed in subsequent work that management researchers should be aware of. For instance, in a restatement of the causal steps procedure, Kenny, Kashy, and Bolger (1998) noted that the first and fourth steps are unnecessary, leaving the second and third steps as necessary and sufficient to establish mediation. Methods for testing the mediated effect indicated by $\gamma_{21}\beta_{12}$ have also been refined. The test proposed by Baron and Kenny (1986) adapted work by Sobel (1982) to obtain the standard error of $\gamma_{21}\beta_{12}$, which can be expressed as:

$$\sqrt{\gamma_{21}^2 s_{\beta_{12}}^2 + \beta_{12}^2 s_{\gamma_{21}}^2 + s_{\beta_{12}}^2 s_{\gamma_{21}}^2}$$

Dividing $\gamma_{21}\beta_{12}$ by this quantity yields a test statistic that can be referred to the standard normal distribution to determine whether $\gamma_{21}\beta_{12}$ is statistically significant, providing a test of whether the mediated effect is statistically different from zero. This procedure is widely used (MacKinnon et al., 2002), and variations of the previous formula are incorporated into structural equation modeling programs to compute standard errors of indirect effects (MacKinnon et al., 2002; Shrout & Bolger, 2002). However, this test relies on the assumption that the product $\gamma_{21}\beta_{12}$ is normally distributed, which is necessarily violated when γ_{21} and β_{12} are themselves normally distributed (Bollen & Stine, 1990; MacKinnon et al., 2002; Shrout & Bolger, 2002). This problem can be overcome by testing $\gamma_{21}\beta_{12}$ using nonparametric procedures such as the bootstrap (Bollen & Stine, 1990; Efron & Tibshirani, 1993), in which cases are randomly drawn with replacement from the original data to construct bootstrap samples, the model

is estimated for each bootstrap sample, and the results from the bootstrap sample are used to construct sampling distributions and confidence intervals of parameter estimates. Simulation studies evaluating the bootstrap indicate that accurate confidence intervals for $\gamma_{21}\beta_{12}$ can be obtained using the bias-corrected percentile method (Cheung & Lau, 2008; Cheung, 2007; MacKinnon, Lockwood, & Williams, 2004; Taylor, MacKinnon, & Tein, 2008; Williams & MacKinnon, 2008).

Techniques for analyzing the basic mediated model in Figure 12.5b can be extended to more complex models typical of management research. For instance, the model in Figure 12.5c has two mediated effects, one through η_2 and another through η_3 . As stated before, the mediated effect through η_2 is evidenced when γ_{21} and β_{12} both differ from zero, and the magnitude of this effect is given by the product $\gamma_{21}\beta_{12}$. Similarly, the mediated effect through η_3 is supported when γ_{31} and β_{13} both differ from zero, and the size of this effect equals $\gamma_{31}\beta_{13}$. The sum of these two mediated effects and the direct effect indicated by γ_{11}^{**} gives the total effect of ξ on η_1 (Alwin & Hauser, 1975; Bollen, 1987), which represents to total impact of ξ on η_1 . With the addition of a path, the model in Figure 12.5d has three mediated effects, including the two effects represented by $\gamma_{21}\beta_{12}$ and $\gamma_{31}\beta_{13}$ along with the three-stage effect captured by the product $\gamma_{31}\beta_{23}\beta_{12}$. This effect is supported if γ_{31} , β_{23} , and β_{12} each differ from zero. Adding these three mediated effects to the direct effect γ_{11}^{***} yields the total effect of ξ on η_1 . For both of these models, the coefficient products involved in the mediated and total effects can be tested using bias-corrected confidence intervals constructed from results produced by the bootstrap.³

Models with several latent variables can involve numerous mediated effects, and algorithms for identifying these effects as combinations of direct and indirect effects have been developed (Alwin & Hauser, 1975; Bollen, 1987; Brown, 1997; Fox, 1980, 1985; Greene, 1977). These algorithms are essential for theory testing by management researchers, because different mediated effects within a model typically represent distinct theoretical processes. As a result, while current structural equation programs can report the total effects of one latent variable on another, they do not report the individual mediated effects that constitute these sums. These individual effects are important because they represent distinct mediated processes by which one variable can influence another.

As a final caveat, we recommend to management researchers that mediated effects be considered as synonymous with indirect effects (Alwin & Hauser, 1975). This position is consistent with most discussions of mediation (Baron & Kenny, 1986; James & Brett, 1984; Kenny et al., 1998; MacKinnon et al., 2002; Shrout & Bolger, 2002). However, some authors distinguish between mediated and indirect effects by arguing that a mediated effect requires an initial relationship between the independent and dependent variable, claiming that this relationship establishes that there is a relationship to be mediated

(Holmbeck, 1997; Preacher & Hayes, 2004). To the contrary, we argue that mediation can exist irrespective of the relationship between the independent and dependent variable (as when a mediated effect is accompanied by a direct effect of opposite sign that nullifies the overall (i.e., total) effect of the independent variable on the dependent variable, as previously discussed). Other researchers have suggested that an indirect effect exists only when the independent and dependent variables are unrelated (Mathieu & Taylor, 2006). This position seems difficult to defend given that, when the independent and dependent variable are unrelated, an indirect effect can occur only when it is accompanied by countervailing effects of equal magnitude. Returning to the model in Figure 12.5b, the relationship between the independent and dependent variable is represented by the equality $\gamma_{11} = \gamma_{21}\beta_{12} + \gamma_{11}^*$. If this relationship equals zero, then we have $0 = \gamma_{21}\beta_{12} + \gamma_{11}^*$, or $\gamma_{21}\beta_{12} = -\gamma_{11}^*$. We see no reason to restrict consideration of indirect effects to cases in which this equality holds. Rather, we believe the analysis of mediation by management researchers would be greatly facilitated by equating mediated effects with indirect effects and applying effect decomposition procedures as previously described to identify the various components that constitute the relationships among latent variables (Alwin & Hauser, 1975; Bollen, 1987; Brown, 1997; Fox, 1980, 1985; Greene, 1977).

Moderation and Latent Variable Relationships

A second type of latent variable relationship we will consider involves moderation, in which the effect of one variable on another depends on the level of a third variable, usually termed a moderator variable (Zedeck, 1971). Contemporary examples of moderation in management research include gender as a moderator of the effects of work–family conflict on career satisfaction (Martins, Eddleston, & Veiga, 2002), emotional intelligence as a moderator of the effects of job insecurity on coping behavior (Jordan, Ashkanasy, & Hartel, 2002), organizational structure as a moderator of the relationship between justice and social exchange (Ambrose & Schminke, 2003), social status as a moderator of the relationship between organizational citizenship behavior and mistreatment in the workplace (Aquino & Bommer, 2003), and strategic orientation as a moderator of the effect of strategy formulation capability and firm performance (Slater, Olson, & Hult, 2006).

Methods for analyzing moderation with multiple regression are well established and involve the use of product variables created by multiplying the two variables involved in the interaction to create another variable that is used to test the interaction hypothesis (Aguinis, 2003; Aiken & West, 1991; Jaccard, Turrisi, & Wan, 1990). In contrast, methods for analyzing moderation in structural equation models are continuing to evolve. Early management applications focused on categorical moderators and used the multi-sample capabilities discussed previously in the section on measurement invariance to test the

equality of structural paths across groups (Podsakoff, Williams, & Todor, 1986). More recently, several studies have used a version of the total aggregation with reliability correction approach (briefly mentioned previously as an alternative to items and parcels), which was extended to testing continuous moderators by Mathieu, Tannenbaum, and Salas (1992) and was reviewed by Cortina, Chen, and Dunlap (2001). Examples of this approach include a study of empowerment readiness as a moderator of leadership empowerment behavior on sales performance (Ahearne, Mathieu, & Rapp, 2005), research on perceived organizational value of creativity moderating the effects of creative role identity on employee creativity (Farmer, Tierney, & Kung-McIntyre, 2003), and a study of various moderators of cognitive, motivational, and emotional processing underlying active learning approaches (Bell & Kozlowski, 2008).

Management researchers attempting to apply SEM methods for studying interactions face numerous choices and unresolved issues. Although the total aggregation approach to testing interactions has been used and has some advantages, we will focus on relatively newer approaches that are based on the use of multiple indicators. Historically, most work on moderated structural equation modeling with multiple indicators can be traced to Kenny and Judd (1984), who examined the following moderated structural equation: $\eta_1 = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta$. Referring to the model in Figure 12.1, this equation would be appropriate if an interaction was proposed between ξ_1 and ξ_2 because it involves creating a product variable involving the exogenous latent variables ($\xi_1\xi_2$), the significance of which provides the statistical test of interaction and moderation.

A key issue concerns how the indicators for this product latent variable are created. Kenny and Judd (1984) formed indicators of $\xi_1\xi_2$ by computing all pair-wise products of the indicators of ξ_1 and ξ_2 . However, Kenny and Judd (1984) also showed that use of product indicators introduces dependencies among the item loadings, measurement errors, and variances and covariances of ξ_1 , ξ_2 , and $\xi_1\xi_2$, which they noted could be handled by incorporating nonlinear constraints into the model. In this context, these constraints result in the value of one factor loading estimate being constrained to take on the value of the products of two other factor loadings from the same model. Kenny and Judd (1984) imposed these constraints using the early software program COSAN (Fraser, 1980), and Jaccard and Wan (1995) showed how the constraints can be imposed in LISREL (Jöreskog & Sörbom, 1993). Subsequently, Jöreskog and Yang (1996) extended the Kenny and Judd (1984) model to include parameters representing means of manifest and latent variables and intercepts in the measurement and structural equations, which are needed to properly specify the model.

More recently, Algina and Moulder (2001) modified the Jöreskog and Yang (1996) approach by using mean-centered indicators (deviation values instead of raw scores) for ξ_1 and ξ_2 and fixing their intercepts to zero, which reduces

convergence problems and makes it more likely that acceptable parameter estimates will be obtained. Wall and Amemiya (2001) derived a generalized appended product indicator (GAPI) model that is similar to the Algina and Moulder (2001) approach but makes no distributional assumptions concerning the variables in the model, which in turn lessens the need for some of the nonlinear constraints. Also recently, Marsh, Wen, and Hau (2004) went further by dropping all nonlinear constraints of the Jöreskog and Yang (1996) approach, with the intent of simplifying model specification. In addition to these main approaches, other methods that involve product indicators include the two-step procedure presented by Ping (1995, 1996) and the two-stage least squares (TSLS) approach derived by Bollen (1995; Bollen & Paxton, 1998).

The various methods for analyzing moderated structural equation models with product indicators have been critically evaluated (Cortina, Chen, & Dunlap, 2001; Jöreskog, 1998; Li et al., 1998; Marsh et al., 2004; Moulder & Algina, 2002; Yang-Wallentin & Jöreskog, 2001) and management researchers should be aware of their limitations. Overall, this work indicates that the two-step procedure (Ping, 1995, 1996) and TSLS method (Bollen, 1995; Bollen & Paxton, 1998) are less effective than other approaches in terms of parameter bias, efficiency, and statistical power. Methods that disregard means and intercepts (Jaccard & Wan, 1995; Kenny & Judd, 1984) should also be avoided, not only because these parameters are nonzero when models include product terms (Jöreskog & Yang, 1996), but also because means and intercepts are required to locate scale values of the latent variables, as needed to interpret moderating effects (Aiken & West, 1991). A drawback of the unconstrained approach of Marsh et al. (2004) is that it estimates parameters that are known to be functions of other parameters even when all distributional assumptions are dropped (Wall & Amemiya, 2001), which needlessly consumes degrees of freedom while testing the model. On the positive side, models consistent with the Jöreskog and Yang (1996) approach generally perform well, provided the $\xi_1\xi_2$ product term has multiple indicators, the indicators of ξ_1 and ξ_2 are mean-centered, and constraints that result from normality assumptions are relaxed (Algina & Moulder, 2001; Marsh et al., 2004; Wall & Amemiya, 2001; Yang-Wallentin & Jöreskog, 2001).

The foregoing approaches to moderated structural equation modeling involve product indicators of $\xi_1\xi_2$, which greatly increase the complexity of the measurement model for ξ_1 , ξ_2 , and $\xi_1\xi_2$ due to the additional factor loadings and error variances required and the interdependencies among some of these parameters. This added complexity is avoided by methods that do not require product indicators. These methods include the latent moderated structural equations (LMS) approach proposed by Klein and Moosbrugger (2000), the quasi-maximum likelihood method (QML) derived by Klein and Muthén (2007), the two-stage method of moments (2SMM) procedure proposed by Wall and Amemiya (2000, 2003), and Bayesian methods developed by

Arminger and Muthén (1998) and Lee and colleagues (Lee, Song, & Tang, 2007; Lee & Zhu, 2000; Song & Lee, 2006). Simulation studies indicate that these methods perform well relative to methods that involve product indicators discussed in the preceding paragraph (Dimitruk, Schermelleh-Engel, Kelava, & Moosbrugger, 2007; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lee, Song, & Poon, 2004; Schermelleh-Engel, Klein, & Moosbrugger, 1998; Wall & Amemiya, 2000, 2003).

Although these methods for estimating moderated structural equation models without product indicators are promising, the estimates they provide are not appreciably better than those yielded by the GAPI model (Wall & Amemiya, 2001) and the Jöreskog and Yang (1996) approach with mean-centered indicators for ξ_1 and ξ_2 and multiple indicators assigned to the $\xi_1\xi_2$ product term (Algina & Moulder, 2001). Moreover, the LMS, QML, and Bayesian methods rest on the assumption that the indicators of ξ_1 and ξ_2 follow a multivariate normal distribution, and it remains unclear how well these methods perform when this assumption is violated. In addition, these methods are difficult to implement because they have not been incorporated into structural equation modeling programs. Currently, the LMS approach is available in *MPlus* (Muthén & Muthén, 2007), and the 2SMM can be implemented with customized code written for SAS (Wall & Amemiya, 2003).

Beyond these options, researchers must obtain, learn, and implement the required software from researchers who developed the methods. Moreover, the literature that describes these methods speaks to statisticians rather than applied management researchers, and understanding the details of these methods requires advanced statistical training. For these reasons, methods for estimating moderated structural equation models without product indicators can be viewed as promising but technically demanding alternatives to methods that use product indicators, which may be preferred by management researchers. For now, as noted earlier, models based on the Jöreskog and Yang (1996) approach seem to be best, provided the $\xi_1\xi_2$ product term has multiple indicators, the indicators of ξ_1 and ξ_2 are mean-centered, and constraints that result from normality assumptions are relaxed.

Latent Growth Models for Examining Change with Latent Variables

We now consider a third special type of latent variable relationship relevant for management researchers, who often collect data on the same variables from the same source and from at least three occasions across time to circumvent the criticisms of working with cross-sectional designs. With this type of data, regression analysis has typically been used to examine the associations among variables across time. With such an analysis, differences *between* observations on one variable are being associated with the differences between observations on another variable. However, in many instances the researcher is interested in whether *change* in the independent variable will cause a *change*

in an outcome variable. A problem with traditional approaches is that associating the differences between observations on one variable with the between observation differences on another variable, such as in regression analysis, does not analytically operationalize change (Burr & Nesselroade, 1990; Collins & Sayer, 2001; Cronbach & Furby, 1970; Duncan, Duncan, Strycker, Li, & Alpert, 1999; Lance et al., 2000; Rogosa, 1995; Rogosa, Brandt, & Zimowski, 1982). Change in a variable means that across time we will observe an increase or decrease in it that may be linear or curvilinear in some fashion, and that we may associate with outcome variables (Bentein et al., 2005; Chan, 1998; Chan, 2002; Chan & Schmitt, 2000; Lance et al., 2000).

The analysis of change within SEM is referred to as growth modeling or latent growth modeling (LGM). Historically, growth modeling has its roots in two statistical camps—the random coefficient modeling (RCM) approach best represented by Singer and Willett (2003) and the SEM approach best represented by Bollen and Curran (2006). The latter is the emphasis in this chapter because it incorporates latent variables and multiple indicators. Recent examples of the use of LGM in management research include Bentein, et al. (2005) who examined the impact of changes in affective, normative, and continuous forms of commitment on change in turnover intention across 18 months, and how the change in intention predicted actual turnover behavior at 24 months. Also, Chan and Schmitt (2000) examined how change in organizational newcomers' proactive behaviors impacted their initial workplace adaptation to the organization during the socialization period. In yet another study, Day and Lance (2004) examined how individuals developed (i.e., changed) as leaders over time and the implications this had for their own, as well as the unit's effectiveness. A final example includes Marathe, Wan, Zhang, and Sherin (2007) who examined the impact of change in contextual and organizational structure factors upon both the changes in organizations' technical and cost efficiencies.

To illustrate latent growth modeling, we present an Example Model in Figure 12.6 which is based upon the Bentein et al. (2005), and Lance et al. (2000) studies. The example includes four latent variables, η_1 to η_4 (near the bottom of Figure 12.6) that represent the construct of affective commitment (AFF) measured at four points in time, AFF_1 – AFF_4 , and each is operationalized with three items (y_{11} , etc.). The correlations among the same uniqueness estimates for items across time (represented by double-headed arrows at the bottom of Figure 12.6) represent shared error variance associated with using the same items across time.

A key advantage with this design is that using the four repeated measurements of AFF allows for the creation of an initial status or intercept latent variable (η_1) and a slope or change latent variable (η_2). Conceptually, the initial status latent variable (Initial Status-AFF) represents the status of affective commitment at the first occasion (Time 1). The change latent variable (Slope/

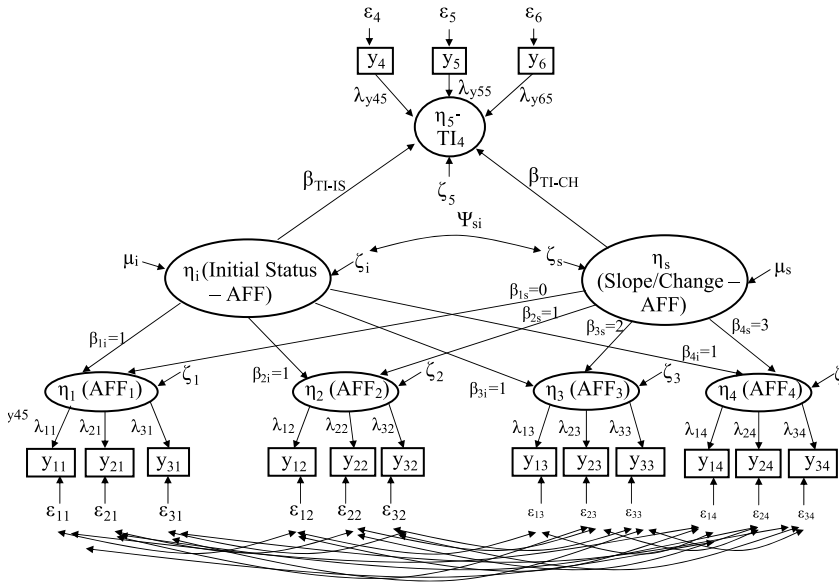


Figure 12.6 Example of Latent Growth Model.

Note: AFF₁, affective commitment Time 1; AFF₂, affective commitment Time 2; AFF₃, affective commitment Time 3; AFF₄, affective commitment Time 4.

change-AFF) represents the vector or slope of change in affective commitment from the initial status or first occasion. Operationally, implementing the LGM model requires the step of fixing to prespecified values certain parameters that otherwise might be freely estimated. Specifically, the initial status latent variable is associated with the first occasion of affective commitment by fixing the betas from the initial status variable to each of the four commitment latent variables (β_{i1} to β_{4i} in Figure 12.6) to the value of one, and by fixing the beta (β_{1s}) from the slope or change latent variable to the first occasion of affective commitment to zero. Similarly, by fixing the betas from the change latent variable to the four affective commitment variables (β_{1s} to β_{4s} in Figure 12.6) at the values of 0, 1, 2 and 3, we can represent linear change across time and the fact that the four occasions of measurement were obtained at equally spaced intervals. However, nonlinear change may also be estimated. The model as described up to this point would be considered as a Level-1 LGM because it focuses on change within a single latent variable over time.

However, a management researcher may also be interested in investigating whether the latent variable representing change in affective commitment is related to some outcome variable, such as individual turnover intention (η_5) measured at the fourth measurement occasion (TI₄). The addition of latent variables and related parameters would result in a new model that would be

labeled as a Level-2 LGM. This model would be consistent with theory that there should be a negative association between affective commitment and turnover intention (Bentein, et al., 2005; Lance, et al., 2000). As shown in the figure, the path from the initial status latent variable to the turnover intention latent variable ($\beta_{\text{TI-IS}}$) represents the cross-sectional test (i.e., traditional test) of the latter hypothesis; that the between-observation differences on the initial status of affective commitment is associated with the between-observation differences of turnover intention. In contrast, the path from the change or slope latent variable to the turnover intention latent variable ($\beta_{\text{TI-CH}}$) represents a fundamentally different test of the hypothesis. A statistically significant and negative estimate of this path coefficient ($\beta_{\text{ti-c}}$) would indicate that higher levels of positive *change* in affective commitment across time would be associated with lower levels of individual's turnover intentions at Time 4. Such a conclusion could not be achieved in a cross-sectional design with latent variables or with a longitudinal design using traditional analyses such as regression.

Management researchers have some important issues to consider before undertaking an LGM, and going through the expense of resources to collect data over multiple time periods. First among these issues is the question of from whom the data will be collected. Even if the researcher has a well specified LGM that is anchored tightly to an appropriate conceptual framework, supportive results may not be possible if there is no change in the substantive variables of interest during the data collection. The application of LGM requires the researcher to carefully select a sample with a focus on the substantive latent variables along which change is required to validly test the hypotheses. A traditionally obtained convenience sample may not be adequate if change does not occur in the key variables as needed to test the theory.

Assuming the sample issue has been addressed, another issue of importance is measurement quality. It is essential that researchers use measures of the substantive variables that have very strong psychometric properties and known validity. This is key because configural and metric invariance (see review of these in earlier part of the chapter) must be present for the variables along which change is expected (affective commitment in Figure 12.6) before one proceeds with the actual test of the LGM (Bentein et al., 2005; Bollen & Curran, 2006; Chan, 1998; Lance et al., 2000). If configural invariance is not supported, then as noted earlier the observed responses to the underlying question were being determined by different constructs at the different time points, which precludes further analysis of change. Metric invariance or strong partial metric invariance must also be present, given that the intercept of the response (observed items) is a function of its lambda (i.e., factor loading onto the latent variable), and the fact that the mean of the substantive latent variable is a function of the item intercepts associated with the substantive variable. Therefore, if metric invariance is not strongly supported, one cannot be fully confident that the mean change represents real change, as it may be an

artifact of the measures not possessing metric invariance (creating artifactual change). These problems are least likely to occur if the measures of the substantive latent variables possess strong measurement qualities, including configural and metric invariance.

A management researcher could examine more complicated hypotheses by extending the research design associated with Figure 12.6. Since four waves of data are involved, nonlinear forms of change (Bollen & Curran, 2006) such as a quadratic change function could be examined. The model could also have included more than one variable along which change was expected. For example, turnover intention could also have been measured at the four time points, and a researcher could evaluate whether change in affective commitment is associated with change in turnover intention (Bentein, et al., 2005). In addition, assuming theoretical differences exist between groups or samples with respect to baseline and change (e.g., an intervention group vs. a control group; an organization undergoing change vs. one not changing; etc.), the groups could be compared via multi-sample analysis to test, for example, whether they started at the same place, or changed at the same rate or in the same direction. Further, moderation tests may be undertaken whereby the impact of change in one latent variable on some outcome is a function of the change in another latent variable. Finally, a lagged design could be used to test causal order.

One final recommendation we provide for management researchers is that they conduct Level-1 tests separately on each of the variables along which change is expected before conducting Level-2 tests that include outcome variables (Bollen & Curran, 2006; Duncan et al., 1999). As noted earlier, the model in Figure 12.6 is technically a Level-2 or conditional LGM because there is a “causal” path from the initial status and change in affective commitment to turnover intention. Thus, intention is conditional upon the initial status and change latent variables. Figure 12.6 would be a Level-1 or unconditional model if turnover intention were removed from the analysis and only the significance of the initial status and change latent variables were evaluated. The Level-1 LGM tests are recommended to: (a) determine whether change did indeed occur; (b) ascertain the form of change (i.e., linear, quadratic, etc.); and (c) identify the direction of change (i.e., upward or downward). If no change had occurred, for instance, it would not be appropriate to attempt to model the slope or change latent variable as an antecedent to turnover intention in Figure 12.6. Further, learning the shape and direction of change from a Level-1 analysis would facilitate the interpretation of the findings from the Level-2 or conditional LGM.

Multilevel Issues: Latent Variable Relationships at Individual and Group Levels

The fourth special type of latent variable relationship we consider involves multiple levels of analysis. In the preceding examples, the relationships

between latent variables examined have been at the individual level of analysis and an unmentioned statistical assumption of independence of observations was made. In many situations common to management research this assumption may be inappropriate, such as when data are collected from individuals nested within work units (e.g., teams or departments). Further, in some of these instances researchers may be interested in investigating models about group-level processes, such as when differences between teams on their commitment is proposed to influence team-level performance. A traditional approach to the analysis of multilevel data involves RCM with observed indicator variables, as was mentioned in discussions of analysis of change.

We will focus here on SEM models for multilevel analysis, and because of the relative newness of this advanced application of SEM, only a few examples are currently available. For one, Richardson and Vandenberg (2005) examined how a team leader's perceptions (a team-level variable) of how well employees could work in a high-involvement work environment impacted the team employees' collective perceptions (an aggregated individual-level variable) of the team climate for high-involvement work processes. Further, they examined how the latter perceptions influenced team absence (based on organizational records) and performance (again a team-level variable). In another application, Cheung and Au (2005) examined whether the same measurement and structural path models (predicting intention from commitment, job security etc.) at the individual level were appropriate to 27 different cultures (i.e. the between-unit level).

The example we will use to discuss issues related to multilevel SEM can be found in Figures 12.7a and 12.7b. This model is based on previous work by Richardson and Vandenberg (2005) and Vandenberg, Richardson and Eastman (1999). To investigate this model it should be noted that data must be collected from a large number of individual employees who are grouped in an adequate number of work teams. In brief, the first conceptual premise related to multilevel SEM is that at the individual or within-unit level, individuals' perceptions of their managers' transformational leadership behaviors (TLB) directly influences their perceptions of the degree to which their work unit is a high involving one (HIW); that is, one in which they are given autonomy and are provided the information, training and other resources to work in an independent manner. This part of the proposed model is shown in Figure 12.7a. In addition to these individual-level processes, the second conceptual premise is that since the same manager directs several members in a work unit, these units are expected to have varying climates for transformational leadership and high-involvement work processes due to differences between managers. Thus, a researcher might want to model these between-group processes, based on the assumption that members within the same unit would be expected to share similar beliefs concerning their manager's transformational style which, in turn, would be expected to influence employees' shared beliefs

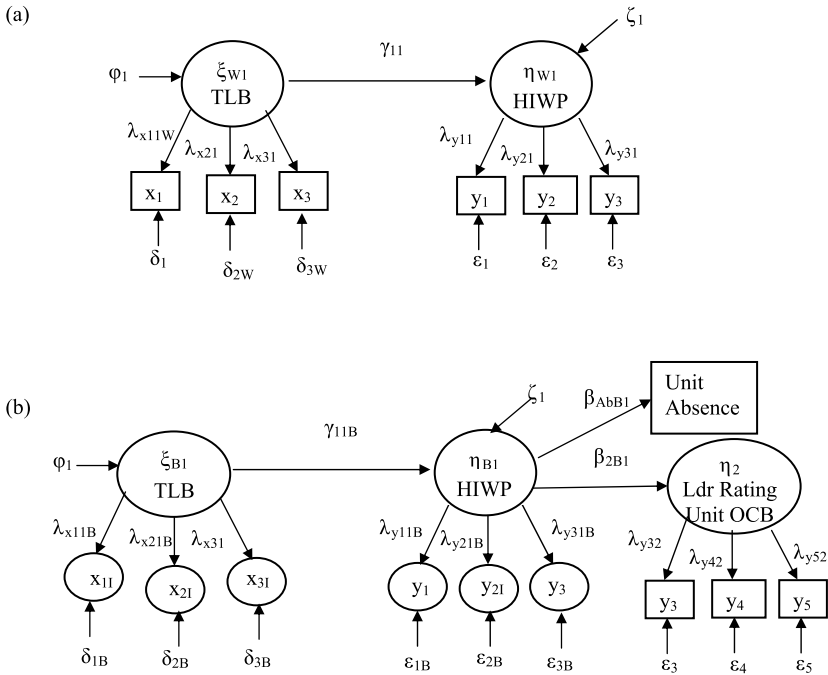


Figure 12.7 (a) Individual or Within-level Model; (b) Group or Between-level Model.

Note: TLB, transformational leadership behavior; HIWP, high-involvement work practices.

concerning the work unit's climate for high involvement. Those hypotheses should be considered if there is something theoretically meaningful about the higher-order unit (in this case the team), and as such individual responses are conditional on those observations belonging to one unit vs. another.

For example, an individual's rating of his/her manager's transformational leadership behaviors (TLB) is proposed to be due to both his/her independent assessment of those behaviors (Figure 12.7a), and a shared belief among all members in the unit as to how those behaviors are exhibited in the work unit (Figure 12.7b). Therefore, regardless of whether one is taking a regression analytical approach (i.e., RCM) or an SEM approach, the key to successfully testing the multilevel hypotheses is separating the individual contributions to the response (also referred to as the within component) from the unit-level contributions to the response (also referred to as the between component). Finally, the model in Figure 12.7b proposes that the climate for high involvement is expected to negatively influence absences in the work unit, but positively influence the degree to which the work-unit managers view their units as engaging in organizational citizenship behaviors.

Management researchers wanting to test multilevel models such as the example shown in Figures 12.7a and 12.7b should begin with a strong theoretical rationale for undertaking a multilevel test that helps define the measures used and how they should be worded (Kozlowski & Klein, 2000). For example, if a researcher is eventually aggregating individual perceptions of unit processes to create work-unit climate indices (such as TLB and HIWP in Figure 12.7b), then the items constituting the measure should make reference to the work unit (Chan, 2005). Second, a strong theory helps define the types of variables to be analyzed at the higher-order level (e.g., work unit, department, organization), which determines how individuals will be grouped. For example, theory should lead to the determination that the work team is the appropriate entity to focus on, as compared to the higher-level unit of department. Defining the type of higher-level entity or group is quite important for specifying the associations to be expected at the higher-order level. Finally, theory helps define how the variables from the lower-order level (i.e., individual) will manifest at the high-order level.

For example, there is the concept of isomorphism in multilevel research (Bliese, 2000; Bliese, Chan & Ployhart, 2007; Chan, 2005; Kozlowski & Klein, 2000). Variables are isomorphic if the conceptual content of the latent variables remains the same across levels as higher-order variables are composed from lower-order ones. This is illustrated in Figures 12.7a and 12.7b where transformational leadership and high-involvement work processes are expected to be the same conceptually across levels. However, isomorphism does not have to be case under all circumstances. Finally, it should be noted that any measures obtained at the lower level (e.g., individual perceptions of high-involvement work processes) that will also be used in some form at the higher-order level (e.g., work unit climate for involvement) should meet all of the assumptions regarding within-group agreement (see LeBreton & Senter (2008) for an excellent review of agreement procedures).

For a management researcher wanting to implement multilevel SEM, the analysis begins with the overall covariance matrix among the individual-level indicators, and then this matrix is separated into two other matrices that are used in the analysis, the within- (Σ_w) and between-level (Σ_b) covariance matrices. For the model shown in Figure 12.7a, the within-level (Σ_w) covariance matrix would be among the 6 indicators (x_1 to x_3 and y_1 to y_3), and would consist of 21 elements (15 covariances and 6 variances). These variances and covariances are adjusted in their computation for the fact that individuals belong to the same work unit. The between-level (Σ_b) covariance matrix in Figure 12.7b would be among the six intercept variables (x_{11} to x_{31} , and y_{11} to y_{31}), the three observed variables for the leader's ratings of the units OCBs (y_4 to y_6), and the measure of unit absence. The latter matrix would consist of 55 elements (10 variances and 45 covariances). An important attribute of this matrix is that the six intercept variables only represent between-unit

differences at this point, and no part of their variances may be attributed to the individual. This feature is what makes it possible to test complex multilevel SEM path models as presented in Figures 12.7a and 12.7b. We mention the two covariance matrices because some researchers want to average individual responses and use this average to represent the between-unit latent variables. This would be inappropriate because such an average is a composite of the individual and unit contributions to the score, and thus, may result in erroneous conclusions about group-level processes.

Another unique aspect of multilevel SEM as shown in Figure 12.7b is the use of boxes to represent the items at the within-group level, but circles are used to represent the between-groups equivalent of those items. This is necessary because these group-level indicators are actually intercepts of the observed scores within each group. Given that they are intercepts, they are predicted values and thus, are latent variables in their own right. Thus, like other latent variables, they are represented through circles (Muthén and Muthén, 2007). Finally, yet another unique aspect of the Example Model in Figure 12.7b is the incorporation of both an observed (absence) and a latent (leader's rating of work unit OCB) endogenous variable at the between level. Work-unit absence rates could be obtained from organizational records, and the OCB ratings could be perceptual in nature. Thus, the OCB latent variable has its own measurement model. Note that since the ratings are obtained from managers, its indicators are observed scores and thus, are represented as boxes in Figure 12.7b.

For management researchers, there are several distinct advantages of taking an SEM approach to multilevel modeling vs. RCM procedures which are often implemented with the software program HLM (Bryk & Raudenbush, 1987). Foremost among those advantages is the ability to test complex multivariate path models such as that in Figures 12.7a and 12.7b. This can be done in an SEM framework but would be very complex to do using RCM procedures. A second main drawback to RCM is that the between-level variables that are created from the individual-level variables (e.g., TLB and HIWP in Figure 12.7b) may only be used as criterion variables at the between-level or unit-level analysis. Thus, RCM procedures could neither treat TLB as a latent exogenous variable nor HIWP as a latent endogenous variable predicting absences and leader OCB ratings (since these are not criterion variables). Another advantage of the SEM approach over the RCM procedure is the use of a measurement model underlying the latent variables. As such, measurement error is accounted for in the SEM approach. With RCM, variables such as transformational leadership and high-involvement work processes would be typically represented as means, and thus, there is the implicit assumption that they are measured without error. Finally, if the models in Figures 12.7a and 12.7b were going to be evaluated to test for differences between two different organizations, then tests of invariance such as those applied to measurement models

discussed in a previous section of this chapter could be undertaken using the SEM approach but that are not readily available with the RCM procedure.

Some general issues need to be considered by management researchers planning a multilevel SEM research study. One very important consideration involves the sample size at the within- and between-unit levels. Heck and Thomas (2000) advocate maximizing sample size at the between level (e.g., having 200 work units vs. 20) in order to produce stable standard errors for examining parameter estimates at this level. However, one should not neglect the sample size at the individual level as well. If, for example, a researcher had 200 work units with each unit consisting of 20 people, but only two people from the 20 responded per unit, then it is very doubtful that any aggregated scores from the individual level would adequately represent a unit climate process at the between level. Thus, we recommend maximizing the number of units but also monitoring how many observations in a unit responded, and who responded. If it is only the highly tenured unit members who responded, for example, then one has to question whether their responses reflect the shared perceptions of all members of that unit. Please note that one could obtain a high within-group agreement supporting their aggregation in the latter case, but conceptually the score may not be a valid representation of the substantive team process of interest (Chen Mathieu, & Bliese, 2004).

Use of Control Variables and Latent Variable Relationships

Our fifth special topic involving latent variable relationships is one that applies to all of the types of models considered in the second part of this chapter. Many management researchers include “control” variables in their designs and analyses with the goal of obtaining better estimates of the relationships between predictors and outcomes based on the theories underlying their research. Referring back to the model in Figure 12.1, a researcher whose theory led to the inclusion of ξ_1 as a predictor of η_1 might also include a control variable that is of less substantive interest but that is added so that the γ_1 parameter estimate will not be “biased”. The intent of the researcher would be that by including the control variable in the model, the γ_1 parameter estimate will better represent the real relationship between ξ_1 and η_1 and will not reflect the fact that the control variable was not included in the model. A researcher might justify this practice with the perspective that it is better to be conservative and include the control variable, even if it does not turn out to be significantly related to the dependent variable, than to leave it out and be subject to the criticism that the significance of the γ_1 estimate reflected an omitted variable. Beyond the Example Model in Figure 12.1, control variables might be considered for models involving mediation and moderation, as well as latent growth and multilevel models.

In implementing a regression analysis with a control variable, a management researcher would likely follow the common practice of including any of

the control variables as a first step in a regression analysis, followed by adding the variables of substantive interest in subsequent steps. In essence, this practice accounts for the variance of the control variables in both the independent and dependent variables. Within an SEM, the approach would be to treat the control variables as exogenous latent variables, allow them to covary with the exogenous variables of substantive interest (ξ_1 and ξ_2), and they would also have direct paths to all of the endogenous variables (η_1 and η_2). In this manner, the variance of the control variables shared with substantive variables of interest is accounted for when testing the significance of γ_{11} , γ_{12} , and β_{21} , the paths of interest since they are associated with the key hypotheses of the model. Examples of the use of control variables in the context of SEM management applications include research on antecedents of adjustment and performance (Kraimer, Wayne, & Jaworski, 2001), a model that links work-life benefits to organizational citizenship behavior (Lambert, 2000), and an investigation of connections between strategic human resource management and firm performance. Additional examples include a study of antecedents of management involvement and acquired knowledge (Tsang, 2002) and research on geographic score and multinational enterprise performance (Goerzen & Beamish, 2003). Finally, Bentein et al. (2005) controlled for gender when examining the impact of change in commitment upon the change in turnover intention, and Richardson and Vandenberg (2005) controlled for industry when conducting their multilevel test on 120 work teams.

Management researchers should consider several issues when deciding to include control variables in their SEM analyses, and we would like to emphasize potential concerns related to the indiscriminate inclusion of control variables, even given the good intentions behind their use. First, the inclusion of control variables can meaningfully change the substantive meaning of the constructs of interest (Breugh, 2006; Edwards, 2008). Edwards (2008, p. 480) illustrates this nicely through the use of findings from Colquitt, Conlon, Wesson, Porter, and Ng (2001) regarding the relationship between procedural justice and job satisfaction obtained after removing (i.e., controlling for) the effects of distributive, interactional and interpersonal forms of justice. As discussed by Edwards, after the removal of the variance of the latter three forms of justice, what remains from the analysis is the relationship of one residual (the residual of procedural justice) to another residual (that of job satisfaction). As Colquitt et al. (2001) did, researchers often interpret these residuals *in toto* as if the complete construct of job satisfaction was regressed onto the complete construct of procedural justice. However, as noted by Edwards (2008), while the residual variances can be interpreted in various ways, they are certainly not the same as the “whole” constructs before control variables were instituted. Indeed, it is difficult to know what the residuals are in the absence of further studies examining the construct validity of the residual itself.

A second concern is that the psychometric properties of the measures operationalizing the control variables are frequently ignored by management researchers. Edwards (2008, p. 480, Equation 5) demonstrates the complications this causes with respect to unambiguously interpreting the regression coefficients representing the associations among the variables of substantive interest. In short, as measurement error of the control variable increases, the control variables have a larger impact on the variables of interest (even if the substantive variable is without error). The problem is that this influence may be largely an artifact of using a control variable characterized by a great deal of measurement error. The issue of unambiguously interpreting the influence of the control variables becomes increasingly complicated with the addition of more control variables each measured with error, combined in the analysis with substantive variables which are also measured with error. The key point is that even though SEM allows for researchers to examine control variables as latent variables to accommodate their measurement error, researchers will ultimately have better success with control variables that are measured well.

A third issue related to use of control variables is that management researchers often ignore possible causal relationships among the control variables and variables of substantive interest in the models they examine. As noted earlier, within an SEM approach control variables are typically modeled as exogenous latent variables that are simply correlated with other substantive exogenous variables of interest. However, as noted by Edwards (2008), there could be any number of possible relationships linking the control and substantive variables. For example, it could be that the control variable is a cause of both the substantive exogenous and endogenous variables of interest. Alternatively, the substantive exogenous variable of interest could be a cause of both the control and the endogenous variables. Or a researcher might be using a control variable as a variable that is better conceptualized as having a moderating effect on relationships of interest. The point, simply stated, is that depending on which of these alternative models is the appropriate one, the researcher may need to examine more complicated models involving tests of mediation, moderation or some combination of mediation–moderation to fully understand the role the control variable has in the model.

A final point we would like to note is the seemingly indiscriminate inclusion of control variables in management studies. According to Becker (2005), over 63% of articles included in his review provided no to very unclear reasons for control variable use. Further, he found that nearly 50% of authors failed to explain how the control variable was operationalized, and nearly as many (48%) failed to discuss the quality of the operationalization's psychometric properties (e.g., reliability, validity, etc.). The net result is the readership not being given a firm conceptual understanding as to why a given control variable was included and why its absence would hinder an unambiguous interpretation of the underlying results if its influence was not removed. Further,

readers are exposed to operationalizations of the control variables that possess unknown or poor measurement qualities. Thus, the real possibility exists that the measures of the control variables are conceptually invalid and thus, are not representing the control variable construct as stated by the researcher. All of these problems result in the potential benefits that lead researchers to include control variables not being realized.

Model Evaluation and Latent Variable Relationships

Our final topic of model evaluation is also one that, like control variables, applies to all of the latent variable models considered in this part of the chapter. As noted earlier, management researchers investigating models such as the Example Model shown in Figure 12.1 will obtain a variety of types of information that can be used to judge the adequacy of the model. Over the years there have been considerable changes in the approaches used by management researchers to evaluate multiple indicator models (see previous reviews by James & James, 1989; Medsker et al., 1994). A more recent review provided by Williams and O'Boyle (2009) has summarized the practices followed in 91 articles published in top management journals between 2001–2008, and from their review several recommendations emerge that should guide future SEM studies. It should be emphasized that the model evaluation process is complex, with several types of information being important, and researchers should be thorough in their use and reporting of this information. Readers are encouraged to consult additional sources for more information on the fit measures and model diagnostics mentioned in the following sections (Kline, 2005).

The first type of information that should be examined for a given model is the goodness of fit measures that are obtained with SEM software packages. These measures ultimately summarize and reflect the difference between the sample covariance matrix used in the analysis and a predicted covariance matrix based on the parameter estimates obtained for a specific model. In general, the more similar the sample and predicted covariance matrices are, the more favorable is the evaluation of the model. Regarding specific indices, management researchers have transitioned away from the goodness of fit index (GFI) and the normed fit index (NFI) to the comparative fit index (CFI) and the root mean square error approximation (RMSEA). These are positive changes as the use of the CFI and RMSEA fit indices are better at assessing model fit (Brown, 2006, pp. 83–86; Kline, 2005, pp. 128–141; Loehlin, 2004, p. 83). A model can be considered favorably if the CFI value exceeds 0.95 and/or the RMSEA is below 0.08, and some SEM experts recommend using these two indices together in judging a model. Another positive trend among management researchers is the increased use of the standardized root mean residual (SRMR), which has risen in popularity and for which a value less than 0.10 is considered to reflect a good model. Other indices showed a reduction

in use in recent management research relative to earlier reviews, such as parsimonious fit indices (PFI, PNFI, and PGFI). Given the problems of these latter indices, their decrease in use is a positive development. However, not all trends from the recent review period were positive. The chi-squared statistic divided by the degrees of freedom (χ^2/df) has almost unanimously been criticized in the SEM methodological literature (see Brown (2006, p. 89) and Kelloway (1998, p. 28) for a summary), however, its presence in the top journals persists even though this statistic tells researchers very little about model fit.

The second set of information that should be examined includes evidence about specific paths that are in the model, including the parameter estimates and their statistical significance. The null hypothesis test for a given structural parameter (γ, β) provides evidence as to whether the proposed relationship between two latent variables is supported. Management researchers typically report this information, which is a good practice. On the other hand, some researchers continue to use modification indices to decide if paths should be added to the model, and this is not a good practice as models revised in this manner typically do not cross-validate well. In addition to the structural parameters, the measurement parameters (factor loadings and error variances) should also be examined. Values for these parameters are used to obtain the squared multiple correlation for each indicator, which summarizes how much of its variance is associated with the latent variable of which it is an indicator. These parameter estimates should also be used to calculate diagnostics for each latent variable that summarize the adequacy of its measures, including the composite reliability and variance extracted estimates. Management researchers need to increase their reporting of these values for the models they examine. Finally, the R^2 statistic for each dependent variable provides researchers with important information regarding how much variance is accounted for by its proposed antecedents, which is different from the issue of how much of the covariances among variables are accounted for by the model (Jöreskog & Sörbom, 1996). Management researchers should also increase their reporting of R^2 statistics for the endogenous latent variables in their models.

In addition to the previous issues, a key question of future importance to management researchers concerns how much the overall fit (or misfit) of a multiple indicator latent variable model is due to limitations of the measurement model that links the latent variables to their indicators. This question is relevant even if one has conducted proper evaluation of a measurement model prior to conducting analyses aimed at testing relationships or causal structure among latent variables, and this issue has been discussed for several years (Mulaik, James, Van Alstine, Bennett, Lind, & Stillwell, 1989; Sobel & Bohrnstedt, 1985; Medsker et al., 1994). As noted by McDonald and Ho (2002), it is possible that the adequacy of a multiple indicator latent variable

model as reflected in an overall GFI can be concluded to be good even when most of this fit is due to the measurement model, or that such adequacy can be concluded to be bad even when the structural model is good due to a poor measurement model. In both cases, a researcher ends up making an inference about relations among latent variables, which is of key theoretical interest, which is wrong owing to the fact that overall fit indices combine information about measurement and structural portions of a model into a single value.

McDonald and Ho (2002) conducted re-analyses that included 14 samples from published between 1995 and 1997 in several psychological journals. Their results showed that in most cases, those models which demonstrated good fit achieved this fit due to good measurement models, even though the structural part of the model may have been unacceptable. Williams and O'Boyle (2009) have replicated the findings of McDonald and Ho (2002) using results from a sample of 34 studies published in top management journals, and found that in most of these studies the overall fit of the full model was being driven by the measurement model, and the structural components did not show good fit. While there are approaches to revising fit indices to reflect the differentiation between the structural and measurement components (e.g., the RNFI of Mulaik et al., 1989), with these approaches it is still not possible to tell if any lack of structural fit is due to a few minor problems of misspecifications vs. one big misspecification. To overcome this fact, McDonald and Ho (2002) recommend a two-step approach that organizational researchers should consider for future use.

With their two-stage procedure, researchers would first fit a basic confirmatory factor analysis using all of the indicators and their latent variables. Next they would fit a structural model using as input the correlations among the latent variables obtained from the original confirmatory factor analysis. The overall fit from the second step yields information about the adequacy of the structural part of the overall theoretical model, and also allows for analysis of residuals at the latent variable level that shows specifically where a model is working well and where it is breaking down. It is difficult, if not impossible, to reach such a conclusion with analyses that begin at the indicator level. While these traditional analyses with indicators as inputs result in residuals that show patterns of misfit at the indicator level, this does lead to direct conclusions about the structural model as can be obtained with the two-step approach of McDonald and Ho (2002) (using factor correlations as input).

A final issue to be discussed emphasizes the importance of looking at a variety of fit indices and diagnostics in judging the adequacy of a model. There does seem to be agreement among organizational researchers using SEM techniques in the value of the use of global fit indices such as the CFI and RMSEA. However, these indices only give an overall assessment of a model, and it cannot be known whether any given value reflects the impact of multiple small shortcomings in a model as compared to a single major problem (as discussed

by McDonald and Ho (2002)). Distinguishing between these two types of model misfit can be better accomplished through the use of residuals, including their magnitude and pattern, which can yield a much more refined answer to the question of where a model is working, as compared with the more typical simple answer to how much it is working.

Additionally, while analysis of residuals at the indicator level can be important, we also believe more attention should be focused on how well the relationships among latent variables are being accounted for. The use of the two-step approach of McDonald and Ho is promising and can help researchers move beyond the question of how good a model is to the question of where it is good and where it is bad, at the level of latent variables directly associated with the theoretical constructs of the model. We feel this approach is consistent with the main reason researchers use SEM, which is to test theories and models of relations among latent variables.

Conclusions

As noted in the introduction, the use of SEM analyses by management researchers has increased steadily over the past 20 years. This rise in popularity reflects the fact that the technique allows researchers to simultaneously implement two key aspects of the research process, linking latent variables associated with concepts of theory to indicators used to represent these concepts and estimating relationships among latent variables as proposed by theory. During this time period, nearly all areas of management research have had their theoretically based propositions tested using SEM, as exemplified by the studies mentioned in the various sections of this chapter. The key areas of organizational behavior, human resources, and strategic management have seen a wide range of processes and relationships examined using SEM, which has become the standard recommended analytical approach. It should be emphasized that there are many aspects of SEM that are still under development and are being investigated by methodologists from management, psychology, sociology, and other social science disciplines. During this developmental period, standards and recommendations have emerged, being guided by the conceptual and analytical work of these methodologists. This work is technical and appears in outlets outside the normal readership outlets of management researchers, and often times there are translation difficulties as management researcher attempt to understand the implications of this work for their own substantive research. Within this context, the goal of our chapter was to present a review of 10 areas of special interest, given the popularity of basic models, such as what we used in our Example Model.

We organized our review of these two groups of topics into one set related to indicators and their relationships to latent variables. This set included: discussion of the development of indicators to represent the latent variables; the types of relationships possible linking latent variables to indicators;

Table 12.1 Summary of Recommendations for Management Researchers Using Latent Variable Models

Topic	Issue/challenge	Recommendations
Developing indicators	Items vs. parcels?	<ol style="list-style-type: none"> 1. Consider items for scale development, if number of latent variables is low, and sample size is adequate. 2. Consider parcels if item distributions non-normal, number of latent variables is high, and sample size relatively small 3. Use factorial method for unidimensional latent variable and domain representative for multidimensional latent variable, but also consider separate latent variables for dimensions
Indicator–latent variable relationships	Reflective vs. formative measurement?	<ol style="list-style-type: none"> 1. Be aware of identification and interpretation issues with formative measurement 2. Be aware of philosophical limitations of formative measurement 3. Consider representing formative indicators with separate latent variables using reflective indicators, model these latent variables as causes of more general constructs, measured with reflective indicators
Multidimensional constructs	How to represent multidimensional constructs with latent variables?	<ol style="list-style-type: none"> 1. Consider modeling multidimensional constructs as separate latent variables 2. Use theory to choose between aggregate and superordinate constructs and be aware of identification and interpretation issues 3. If multidimensional construct exists separately from dimensions, add it to model with its own reflective indicators
Measurement invariance across groups and time	How to establish equality of indicators?	<ol style="list-style-type: none"> 1. If multiple samples or longitudinal data used, investigate equality of measurement parameters 2. Obtain evidence for configural and metric invariance before examining structural parameters 3. If only partial invariance obtained, consider excluding items or latent variables, collapsing across samples, or collecting more data

Table 12.1 Summary of Recommendations for Management Researchers Using Latent Variable Models (Continued)

Topic	Issue/challenge	Recommendations
Mediation and latent variable relationships	How to test for mediation?	<ol style="list-style-type: none"> 1. Be aware of revision to Baron and Kenny's (1986) causal steps approach 2. Test for mediated effects as products of path estimates using bootstrap approach 3. In complicated models with multiple mediation processes, supplement estimate of total effects with reports of all mediated effects using effect decomposition
Moderation and latent variable relationships	How to test for moderation with multiple indicator models?	<ol style="list-style-type: none"> 1. Key issue is how to form product latent variable used to test for moderation 2. Be aware of two categories of approaches, one that uses product indicators and one that does not 3. Jöreskog and Yang (1996) approach recommended, using mean centered indicators and relaxing normality based model constraints
Latent growth models for change with latent variables	How to capture change in latent variables with longitudinal data?	<ol style="list-style-type: none"> 1. Take advantage of longitudinal data and investigate change using latent growth modeling techniques 2. Use Level-2 model to examine impact of change on outcome variables after demonstrating change using Level-1 model 3. Investigate Level-2 model with invariant measures and appropriate sample
Multilevel issues and latent variable models	How to investigate individual- and group-level relationships using latent variables?	<ol style="list-style-type: none"> 1. Be aware of advantages of SEM compared with random coefficient modeling 2. Use theory to guide measurement and to identify appropriate higher-level entity (grouping variable) 3. Note that sample size analysis should consider number of individuals and number of groups

Using control variables in latent variable models	How to optimally use control variables with latent variable models?	<ol style="list-style-type: none"> 1. Be aware that using control variables can change meaning of endogenous latent variables as they become residualized after control variables are included 2. Use control variables conservatively and only with good measures and supporting theory 3. Consider investigating extended causal role of control variables, including as mediators and moderators
Model evaluation and latent variable relationships	How to determine if model is adequate and represents latent variable relationships?	<ol style="list-style-type: none"> 1. Give priority to CFI, RMSEA, and SRMR fit measures 2. Supplement fit measures with other model diagnostics, including analysis of indicator residuals, composite reliability and variance extracted estimates, and R^2 3. Include focus on latent variable relationships with special fit indices, two-step procedure of McDonald and Ho (2002), and analysis of latent variable residuals

Note: SEM, structural equation modeling; CFI, comparative fit index; RMSEA, root mean square error approximation; SRMR, standardized root mean residual.

approaches for examining latent variables that are multidimensional in nature; and the evaluation of measurement models when data from more than one group or more than one time period are being analyzed. From there we switched to several current topics related to the structural model, including: how to approach mediation and moderation hypotheses; analyses useful for investigating change with longitudinal data; models appropriate with data that are nested by level consisting of individuals organized within groups; how to incorporate control variables in an SEM analysis; and how best to evaluate the adequacy of an SEM model of latent variable relationships.

As we covered each of these topics, we attempted to present an overview which would allow the reader with limited SEM experience to understand conceptually the central issues, we presented examples from management research relevant to the topic, and then we identified and discussed the relevant technical literature as a way of building a set of recommendations for how future research could be improved. While considerable material was covered across our 10 sections as we developed our recommendations, we would also like to present the key ones in summary form, as shown in Table 12.1.

We hope this table, and its supporting documentation in the preceding sections, reveal the wide range of difficult choices that researchers will face as they use this powerful tool for theory testing. Our recommendations are offered in the attempt to inform these choices, so as to contribute to improved management research in the future. We also hope our coverage of the issues will help researchers understand that if they want to take full advantage of this method, they will need to sustain a commitment to understanding and implementing best practices that are being recommended. Finally, we want to be very clear of our belief that in the pursuit of our science, good research emerges from the intersection of good theory, good data, and good analyses. While SEM can contribute to the latter category, it is also the case that the first two categories must also not be neglected.

Acknowledgments

The authors would like to extend thanks to Art Brief and Jim Walsh for their helpful guidance during the preparation of this chapter, as well as to Peg Williams, Sven Kepes, and Ernest O'Boyle for their comments on earlier versions.

Endnotes

1. The notion that constructs cause measures does not deny the notion that the act of administering a measurement tool can influence a construct, as when presenting respondents with an attitude survey affects the attitudes that the survey is intended to assess (Feldman & Lynch, 1988; Harrison, McLaughlin, & Coalter, 1996). In such instances, the obtained measure is still caused by the construct, such that the causal chain goes from survey administration to attitude formation to the generated score taken as a measure of the construct.

2. For each model in Figures 12.5a–12.5d, the number of asterisks on γ_i indicates the number of mediated (i.e., indirect) effects that the model includes in addition the direct effect running from ξ_1 to η_1 .
3. In some cases, tests of coefficient products might yield results that are inconsistent with tests of the coefficients that constitute the product. For example, the product of two coefficients can be significant when one coefficient differs substantially from zero while the other coefficient fails to reach significance. Based on our experience, such inconsistencies are rare. However, when they occur, we recommend drawing inferences about mediation based on tests of coefficient products, while explicitly acknowledging the nature of the inconsistencies involved. The extent of such inconsistencies and conditions under which they occur merit further research.

References

- Aguinis, H. (2003). *Regression analysis for categorical moderators*. New York: Guilford Press.
- Ahearne, M., Mathieu, J., & Rapp, A. (2005). To empower or not to empower your sales force? an empirical examination of the influence of leadership empowerment behavior on customer satisfaction and performance. *The Journal of Applied Psychology, 90*, 945–955.
- Aiken, L.A., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aldag, R.J., & Brief, A.P. (1978). Examination of alternative models of job satisfaction. *Human Relations, 31*, 91–98.
- Alge, B.J., Ballinger, G.A., Subrahmaniam, T., & Oakley, J.L. (2006). Information privacy in organizations: Empowering creative and extrarole performance. *Journal of Applied Psychology, 91*, 231–232.
- Algina, J., & Moulder, B.C. (2001). A note on estimating the Jöreskog-Yang model for latent variable interaction using LISREL 8.3. *Structural Equation Modeling, 8*, 40–52.
- Alwin, D.F., & Hauser, R.M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40*, 37–47.
- Ambrose, M.L., & Schminke, M. (2003). Organization structure as a moderator of the relationship between procedural justice, interactional justice, perceived organizational support, and supervisory trust. *Journal of Applied Psychology, 88*, 295–305.
- Ambrose, M.L., & Schminke, M. (2009). The role of overall justice judgments in organizational justice research: A test of mediation. *Journal of Applied Psychology, 94*, 491–500.
- Aquino, K., & Bommer, W.H. (2003). Preferential mistreatment: How victim status moderates the relationship between organizational citizenship behavior and workplace victimization. *Organization Science, 14*, 374–385.
- Arminger, G., & Muthén, B.O. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika, 63*, 271–300.
- Aryee, S., Chen, Z.X., Sun, L.Y., & Debrah, Y.A. (2007). Antecedents and outcomes of abusive supervision: Test of a trickle-down model. *The Journal of Applied Psychology, 92*, 191–201.

- Bagozzi, R.P., & Edwards, J.R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, 1, 45–87.
- Bagozzi, R.P., & Heatherton, T.F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling*, 1, 35–67.
- Baron, R.M., & Kenny, D.A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Bass, B.M., & Riggio, R.E. (2006). *Transformational leadership* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Bauer, T.N., Truxillo, D.M., Tucker, J.S., Weathers, V., Bertolino, M., Erdogan, B., Campion, M.A. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management*, 32, 601–621.
- Baum, J.R., & Locke, E.A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*, 89, 587–598.
- Becker, T.E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8, 274–289.
- Bell, B.S., & Kozlowski, S.W.J. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology*, 93, 296–316.
- Bentein, K., Vandenberg, R., Vandenberghe, C., & Stinglhamber, F. (2005). The role of change in the relationship between commitment and turnover: A latent growth modeling approach. *Journal of Applied Psychology*, 90, 468–482.
- Bliese, P.D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K.J. Klein & S.W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.
- Bliese, P.D., Chan, D., & Ployhart, R.E. (2007). Multilevel methods: Future directions in measurement, longitudinal analyses, and nonnormal outcomes. *Organizational Research Methods*, 10, 551–563.
- Bollen, K.A. (1987). Total, direct, and indirect effects in structural equation models. In C.C. Clogg (Ed.), *Sociological methodology* (pp. 37–70). Washington, DC: American Sociological Association.
- Bollen, K.A. (1989). *Structural equation models with latent variables*. New York: John Wiley and Sons.
- Bollen, K.A. (1995). Structural equation models that are nonlinear in latent variables: A least squares estimator. In P.V. Marsden (Ed.), *Sociological methodology* (Vol. 25, pp. 223–251). New York: Basil Blackwell.
- Bollen, K.A., & Curran, P.J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley-Interscience.
- Bollen, K.A., & Davis, W.R. (1994). *Causal indicator models: Identification, estimation, and testing*. Paper presented at the 1993 American Sociological Association Convention, Miami, FL.
- Bollen, K.A., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, 110, 305–314.

- Bollen, K.A., & Paxton, P. (1998). Interactions of latent variables in structural equation models. *Structural Equation Modeling*, 5, 267–293.
- Bollen, K.A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, 20, 115–140.
- Borsboom, D., Mellenbergh, G.J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111, 1061–1071.
- Breaugh, J.A. (2006). Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology*, 20, 429–443.
- Brown, R.L. (1997). Assessing specific mediational effects in complex theoretical models. *Structural Equation Modeling*, 4, 142–156.
- Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Brown, D.J., Ferris, D.L., Heller, D., & Keeping, L.M. (2007). Antecedents and consequences of the frequency of upward and downward social comparisons at work. *Organizational Behavior and Human Decision Processes*, 102, 59–75.
- Bryk, A.S., & Raudenbush, S.W. 1987. Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147–158.
- Burr, J.A., & Nesselroade, J.R. (1990). Change measurement. In A. von Eye (Ed.), *Statistical methods in longitudinal research* (Vol. 1, pp. 3–34) Boston, MA: Academic Press.
- Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal means and covariance structures analysis (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods*, 1, 421–483.
- Chan, D. (2002). Latent growth modeling. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 302–349). San Francisco, CA: Jossey-Bass.
- Chan, D. (2005). Multilevel research. In F.T.L. Leong & J.T. Austin (Eds.), *The psychology research handbook* (2nd ed., pp. 401–418). Thousand Oaks, CA: Sage.
- Chan, D. & Schmitt, N. (2000). Interindividual differences in intraindividual changes in proactivity during organizational entry: A latent growth modeling approach to understanding newcomer adaptation. *Journal of Applied Psychology*, 85, 190–210.
- Chen, G., Mathieu, M.J., & Bliese, P.D. (2004). A framework for conducting multilevel construct validation. In F. Dansereau & F.J. Yammarino (Eds.), *Research in multi-level issues* (Vol. 3), *Multi-level issues in organizational behavior and processes* (pp. 273–303). Oxford, UK: Elsevier Science.
- Cheung, G.W., & Lau, R.S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods*, 11, 296–325.
- Cheung, M.W.L., & Au, K. (2005). Applications of multilevel structural equation modeling to cross-cultural research. *Structural Equation Modeling*, 12, 598–619.
- Cheung, M.W.L. (2007). Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling*, 14, 227–246.
- Coffman, D.L. & MacCallum, R.C. (2005). Using parcels to convert path analysis models into latent variable models. *Multivariate Behavioral Research*, 40, 235–259.

- Cohen, J. (1982). Set correlation as a general multivariate data-analytic method. *Multivariate Behavioral Research, 17*, 301–341.
- Collins, L.M., & Sayer, A.G. (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.
- Colquitt, J.A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*, 386–400.
- Colquitt, J.A., Conlon, D.E., Wesson, M.J., Porter, C.O.L.H., & Ng, K.Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*, 425–445.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.
- Cortina, J.M., Chen, G., & Dunlap, W.P. (2001). Testing interaction effects in LISREL: Examination and illustration of available procedures. *Organizational Research Methods, 4*, 324–360.
- Costa, P.T., Jr., & McCrae, R.R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of Personality Assessment, 64*, 21–50.
- Cronbach, L.J., & Furby, L. (1970). How we should measure “change”: Or should we? *Psychological Bulletin, 74*, 68–80.
- Day, D.V., & Lance, C.E. (2004). Understanding the development of leadership complexity through latent growth modeling. In D.V. Day, S.J. Zaccaro, & S.M. Halpin (Eds.), *Leader development for transforming organizations* (pp. 41–69). Mahwah, NJ: Lawrence Erlbaum.
- Diamantopoulos, A. (2006). The error term in formative measurement models: interpretation and modeling implications. *Journal of Modeling in Management, 1*, 7–17.
- Diamantopoulos, A., & Sigauw, J.A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management, 17*, 263–282.
- Diamantopoulos, A., & Winklhofer, H.M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38*, 269–277.
- Dimitruk, P., Schermelleh-Engel, K., Kelava, A., & Moosbrugger, H. (2007). Challenges in nonlinear structural equation modeling. *Methodology, 3*, 110–114.
- Duncan, T.E., Duncan, S.E., Strycker, L.A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum.
- Edwards, J.R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. *Organizational Research Methods, 4*, 144–192.
- Edwards, J.R. (2008). To prosper, organizational psychology should... overcome methodological barriers to progress. *Journal of Organizational Behavior, 29*, 469–491.
- Edwards, J.R., & Bagozzi, R.P. (2000). On the nature and direction of the relationship between constructs and measures. *Psychological Methods, 5*, 155–174.
- Edwards, J.R., & Rothbard, N.P. (1999). Work and family stress and well-being: An examination of person–environment fit in the work and family domains. *Organizational Behavior and Human Decision Processes, 77*, 85–129.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.

- Elicker, J.D., Levy, P.E., & Hall, R.J. (2006). The role of leader-member exchange in the performance appraisal process. *Journal of Management*, 32, 531-551.
- Elovainio, M., Kivimäki, M., & Helkama, K. (2001). Organizational justice evaluations, job control, and occupational strain. *Journal of Applied Psychology*, 86, 418-424.
- Erdogan, B., Kraimer, M.L., & Liden, R.C. (2004). Work value congruence and intrinsic career success: The compensatory roles of leader-member exchange and perceived organizational support. *Personnel Psychology*, 57, 305-332.
- Facteau, J.D., & Craig, S.B. (2001). Are performance appraisal ratings from different rating sources comparable? *Journal of Applied Psychology*, 86, 215-227.
- Farmer, S.M., Tierney, P., & Kung-McIntyre, K. (2003). Employee creativity in Taiwan: An application of role identity theory. *Academy of Management Journal*, 46, 618-630.
- Feldman, J.M., & Lynch, J.G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, 73, 421-435.
- Fox, J. (1980). Effect analysis in structural equation models. *Sociological Methods and Research*, 9, 3-28.
- Fox, J. (1985). Effect analysis in structural-equation models II: Calculation of specific indirect effects. *Sociological Methods & Research*, 14, 81-95.
- Fraser, C. (1980). *COSAN user's guide*. Toronto: The Ontario Institute for Studies in Education.
- Frese, M., Krauss, S.I., Keith, N., Escher, S., Grabarkiewicz, R., Luneng, S.T., Heers, C., Unger, J., & Friedrich, C. (2007). Business owners' action planning and its relationship to business success in three African countries. *Journal of Applied Psychology*, 92, 1481-1498.
- Goerzen, A., & Beamish, P.W. (2003). Geographic scope and multinational enterprise performance. *Strategic Management Journal*, 24, 1289-1306.
- Greene, V.L. (1977). An algorithm for total and indirect causal effects. *Political Methodology*, 4, 369-381.
- Harris, M.M., & Schaubroeck, J. (1990). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43-62.
- Harrison, D.A., McLaughlin, M.E., & Coalter, T.M. (1996). Context, cognition, and common method variance: Psychometric and verbal protocol evidence. *Organizational Behavior and Human Decision Processes*, 68, 246-261.
- Heck, R.H., & Thomas, S.L. (2000). *An introduction to multilevel modeling techniques*. Mahwah, NJ: Lawrence Erlbaum.
- Henderson, N.D., Berry, M.W., & Matic, T. (2007). Field measures of strength and fitness predict firefighter performance on physically demanding tasks. *Personnel Psychology*, 60, 431-473.
- Hirst, G., Van Knippenberg, D., & Zhou, J. (2009). A cross-level perspective on employee creativity: Goal orientation, team learning behavior, and individual creativity. *Academy of Management Journal*, 52, 280-293.
- Holmbeck, G.N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65, 599-610.
- Horn, J.L., & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117-144.

- Howell, R.D., Breivik, E., & Wilcox, J.B. (2007). Reconsidering formative measurement. *Psychological Methods, 12*, 205–218.
- Hult, G.T.M., Ketchen D.J., & Nichols, E.L. (2002). An examination of cultural competitiveness and order fulfillment cycle time within supply chains. *Academy of Management Journal, 45*, 577–586.
- Hult, G.T.M., Snow, C.C., & Kandemir, D. (2003). The role of entrepreneurship in building cultural competitiveness in different organizational types. *Journal of Management, 29*, 401–426.
- Ironson, G.H., Smith, P.C., Brannick, M.T., Gibson, W.M., & Paul, K.B. (1989). Construction of a job in general scale: A comparison of global, composite, and specific measures. *Journal of Applied Psychology, 74*, 193–200.
- Jaccard, J., Turrisi, R., & Wan, C.K. (1990). *Interaction effects in multiple regression*. Newbury Park, CA: Sage.
- Jaccard, J. & Wan, C.K. (1995). *LISREL approaches to interaction effects in multiple regression*. Thousand Oaks, CA: Sage.
- James, L.R., & Brett, J.M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology, 69*, 307–321.
- James, L.R. & James, L.A. (1989). Causal modeling in organizational research. In C.L. Cooper & I. Robertson (Eds.) *International review of industrial and organizational psychology* (pp. 371–404). New York: John Wiley.
- Jarvis, C.B., Mackenzie, S.B., & Podsakoff, P.M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*, 199–218.
- Jordan, P.J., Ashkanasy, N.M., & Hartel, C.E.J. (2002). Emotional intelligence as a moderator of emotional and behavioral reactions to job insecurity. *Academy of Management Review, 27*, 361–372.
- Jöreskog, K.G. (1998). Interaction and nonlinear modeling: Issues and approaches. In R.E. Schumacker and G.A. Marcoulides (Eds.) *Interaction and nonlinear effects in structural equation modeling*. (pp. 239–250). London: Lawrence Erlbaum.
- Jöreskog, K.G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International, Inc.
- Jöreskog, K.G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International, Inc.
- Jöreskog, K.G., & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G.A. Marcoulides & R.E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 57–88). Hillsdale, NJ: Lawrence Erlbaum.
- Kalliath, T.J., Bluedorn, A.C., & Strube, M.J. (1999). A test of value congruence effects. *Journal of Organizational Behavior, 20*, 1175–1198.
- Kelloway, E.K. (1998). *Using LISREL for Structural Equation Modeling: A Researcher's Guide*. New York: Sage Publications.
- Kenny, D.A., & Judd, C.M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin, 96*, 201–210.
- Kenny, D.A., Kashy, D.A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 233–265). New York: McGraw-Hill.

- Kishton, J.M., & Widaman, K.F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement, 54*, 757–765.
- Klein, A.G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65*, 457–474.
- Klein, A.G., & Muthén, B.O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research, 42*, 647–673.
- Klein, R. (2005). *Principles and practice of structural equation modeling* (2nd edn). New York: Guilford Publications.
- Kozlowski, S.W.J., & Klein, K.J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K.J. Klein & S.W.J. Kozlowski (Eds.), *Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.
- Kraimer, M.L., Wayne, S.J., & Jaworski, R.A. (2001). Sources of support and expatriate performance: The mediating role of expatriate adjustment. *Personnel Psychology, 54*, 71–99.
- Lambert, S.J. (2000). Added benefits: The link between work–life benefits and organizational citizenship behavior. *Academy of Management Journal, 43*, 801–815.
- Lance, C.E., Vandenberg, R.J., & Self, R.M. (2000). Latent growth models of individual change: The case of newcomer adjustment. *Organizational Behavior and Human Decision Processes, 83*, 107–140.
- Law, K.S., Wong, C.S., & Mobley, W.H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review, 23*, 741–755.
- LeBreton, J.M., & Senter, J.L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815–852.
- Lee, S.Y., Song, X.Y., & Poon, W.Y. (2004). Comparison of approaches in estimating interaction and quadratic effects of latent variables. *Multivariate Behavioral Research, 39*, 37–67.
- Lee, S.Y., Song, X.Y., & Tang, N.S. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling, 14*, 404–434.
- Lee, S.Y., & Zhu, H.T. (2000). Statistical analysis of nonlinear structural equation models with continuous and polytomous data. *British Journal of Mathematical & Statistical Psychology, 53*, 209–232.
- Li, F., Harmer, P., Duncan, T.E., Duncan, S.C., Acock, A., & Boles, S. (1998). Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research, 33*, 1–39.
- Little, T.D., Cunningham, W.A., Shahar, G., & Widaman, K.F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 151–173.
- Locke, E.A. (1976). The nature and causes of job satisfaction. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 1297–1350). Chicago, IL: Rand McNally.
- Loehlin, J.C. (2004). *Latent variable models* (4th edn.). Mahwah, NJ: Lawrence Erlbaum.

- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635–694.
- MacCallum, R., & Browne, M.W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin*, 114, 533–541.
- MacKenzie, S.B., Podsakoff, P.M., & Jarvis, C.B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. *Journal of Applied Psychology*, 90, 710–730.
- MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83–104.
- MacKinnon, D.P., Lockwood, C.M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99–128.
- Marathe, S., Wan, T.T.H, Zhang, J., & Sherin, K. (2007). Factors influencing community health centers' efficiency: A latent growth curve modeling approach. *Journal of Medical Systems*, 31, 365–374.
- Marsh, H.W., Wen, Z., & Hau, K.T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, 9, 275–300.
- Martins, L.L., Eddleston, K.A., & Veiga, J.F. (2002). Moderators of the relationship between work–family conflict and career satisfaction. *Academy of Management Journal*, 45, 399–409.
- Mathieu, J.E., Gilson, L.L., & Ruddy, T.M. (2006). Empowerment and team effectiveness: An empirical test of an integrated model. *Journal of Applied Psychology*, 91, 97–108.
- Mathieu, J.E., Tannenbaum, S.I., & Salas, E. (1992). Influences of individual and situational characteristics on measures of training effectiveness. *Academy of Management Journal*, 35, 828–847.
- Mathieu, J.E. & Taylor, S.R. (2006). Clarifying conditions and decision points for mediational type inferences in organizational behavior. *Journal of Organizational Behavior*, 27, 1031–1056.
- McDonald, R.P., & Ho, M.R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7, 64–82.
- Medsker, G.J., Williams, L.J., & Holahan, P.J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*, 20, 439–464.
- Meglino, B.M., & Ravlin, E.C. (1998). Individual values in organizations: Concepts, controversies, and research. *Journal of Management*, 24, 351–389.
- Mesquita, L.F., & Lazzarini, S.G. (2008). Horizontal and vertical relationships in developing economies: Implications for SMEs' access to global markets. *The Academy of Management Journal*, 51, 359–380.
- Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. *Psychological Bulletin*, 89, 575–588.
- Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14, 5–8.
- Moulder, B.C., & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling*, 9, 1–19.

- Mulaik, S.A., James, L.R., Van Alstine, J.V., Bennett, N., Lind, S., & Stilwell, C.D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, *105*, 430–445.
- Murtha, T.C., Kanfer, R., & Ackerman, P.L. (1996). Toward an interactionist taxonomy of personality and situations: An integrative situational-dispositional representation of personality traits. *Journal of Personality and Social Psychology*, *71*, 193–207.
- Muthén, L.K., & Muthén, B.O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthen & Muthen.
- Nunnally, J.C., Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill Book Company.
- Organ, D.W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.
- Pazy, A., & Ganzach, Y. (2008). Pay contingency and the effects of perceived organizational and supervisor support on performance and commitment. *Journal of Management*, DOI: 10.1177/0149206307312505.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.
- Ping, R.A., Jr. (1995). A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research*, *32*, 336–347.
- Ping, R.A., Jr. (1996). Latent variable interaction and quadratic effect estimation: A two-step techniques using structural equation analysis. *Psychological Bulletin*, *119*, 166–175.
- Ployhart, R.E., & Oswald, F.L. (2004). Applications of mean and covariance structure analysis: Integrating correlational and experimental approaches. *Organizational Research Methods*, *7*, 27–65.
- Podsakoff, N.P., Shen, W., & Podsakoff, P.M. (2006). The role of formative measurement models in strategic management research: Review, critique and implications for future research. In D.J. Ketchen & D.D. Bergh (Eds.), *Research methodology in strategy and management* (Vol. 3, pp. 197–252). Burlington, MA: Elsevier.
- Podsakoff, P.M., MacKenzie, S.B., Podsakoff, N.P., & Lee, J.Y. (2003). The mismeasure of man(agement) and its implications for leadership research. *The Leadership Quarterly*, *14*, 615–656.
- Podsakoff, P.M., Williams, L.J., & Todor, W.D. (1986). Effects of organizational formalization on alienation among professionals and nonprofessionals. *Academy of Management Journal*, *29*, 820–831.
- Preacher, K.J., & Hayes, A.F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717–731.
- Reilly, B.J., & DiAngelo, J.A., Jr. (1990). Communication: A cultural system of meaning and value. *Human Relations*, *43*, 129–140.
- Richardson, H.A., & Vandenberg, R.J. (2005). Integrating managerial perceptions and transformational leadership into a work-unit level model of employee involvement. *Journal of Organizational Behavior*, *26*, 561–589.
- Rigdon, E.E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research*, *30*, 359–383.
- Riordan, C.M., & Vandenberg, R.J. (1994). A central question in cross cultural management research: Do employees of different cultures interpret work related measures in an equivalent manner? *Journal of Management*, *20*, 643–671.

- Roch, S.G., & Shanock, L.R. (2006). Organizational justice in an exchange framework: Clarifying organizational justice distinctions. *Journal of Management*, 32, 299–322.
- Rogers, W.M. & Schmitt, N. (2004). Parameter recovery and model fit using multidimensional composites: A comparison of four empirical parceling algorithms. *Multivariate Behavioral Research*, 39, 379–412.
- Rogosa, D.R. (1995). Myths and methods: “Myths about longitudinal research” plus supplemental questions. In J.M. Gottman (Ed.), *The analysis of change*. Mahwah, NJ: Lawrence Erlbaum.
- Rogosa, D.R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin*, 90, 726–748.
- Schaubroeck, J., Lam, S.S.K., & Cha, S.E. (2007). Embracing transformational leadership: Team values and the impact of leader behavior on team performance. *Journal of Applied Psychology*, 92, 1020–1030.
- Schermelleh-Engel, K., Klein, A., & Moosbrugger, H. (1998). Estimating nonlinear effects using a latent moderated structural equations approach. In R.E. Schumacker & G.A. Marcoulides (Eds.), *Interaction and nonlinear effects in structural equation modeling* (pp. 203–238). Mahwah, NJ: Lawrence Erlbaum.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18, 210–222.
- Schuler, R.S. (1979). A role perception transactional process model for organizational communication–outcome relationships. *Organizational Behavior and Human Performance*, 23, 268–291.
- Schwab, D.P. (2005). *Research Methods for Organizational Studies* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Scott, B.A., Colquitt, J.A., & Zapata-Phelan, C.P. (2007). Justice as a dependent variable: Subordinate charisma as a predictor of interpersonal and informational justice perceptions. *Journal of Applied Psychology*, 92, 1597–1609.
- Shook, C.L., Ketchen, D.J., Hult, G.M., & Kacmar, K.M. (2004). Assessment of the use of structural equation modeling in strategic management research. *Strategic Management Journal*, 25, 397–404.
- Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422–445.
- Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Slater, S.F., Olson, E.M., & Hult, G.T.M. (2006). The moderating influence of strategic orientation on the strategy formation capability–performance relationship. *Strategic Management Journal*, 27, 1221–1231.
- Smith, D.B., Hanges, P.J., & Dickson, M.W. (2001). Personnel selection and the five-factor model: Reexamining the effects of applicant’s frame of reference. *Journal of Applied Psychology*, 86, 304–315.
- Sobel, M.E., & Bohrnstedt, G.W. (1985). Use of null models in evaluating the fit of covariance structure models. In N.B. Tuma (Ed.), *Sociological methodology* (pp. 152–178). San Francisco, CA: Jossey-Bass.
- Sobel, M.E. (1996). An introduction to causal inference. *Sociological Methods & Research*, 24, 353–379.
- Sobel, M.E. (2002). Asymptotic confidence intervals for indirect effects in structural equation models. *Social Methodology*, 13, 290–312.

- Song, X.Y., & Lee, S.Y. (2006). Bayesian analysis of structural equation models with nonlinear covariates and latent variables. *Multivariate Behavioral Research*, 41, 337–365.
- Spreitzer, G.M. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal*, 38, 1442–1465.
- Taylor, A.B., MacKinnon, D.P., & Tein, J.Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods*, 11, 241–269.
- Thompson, M.S., & Green, S.B. (2006). Evaluating between-group differences in latent variable means. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 119–170). Greenwich, CT: Information Age Publishing, Inc.
- Tsang, E.W.K. (2002). Acquiring knowledge by foreign partners from international joint ventures in a transition economy: Learning-by-doing and learning myopia. *Strategic Management Journal*, 23, 835–854.
- Vandenberg, R.J., & Lance, C.E. (2000). A Review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70.
- Vandenberg, R.J., Richardson, H., & Eastman, L. (1999). High involvement organizations: Their antecedents and consequences. *Group & Organization Management*, 24, 300–339.
- Vandenberg, R.J., & Scarpello, V. (1990). The matching model: An examination of the processes underlying realistic job previews. *Journal of Applied Psychology*, 75, 60–67.
- Vandenberg, R.J., & Self, R.M. (1993). On assessing newcomers' changing commitments to the organization during the first six months of work. *Journal of Applied Psychology*, 78, 557–568.
- Wall, M.M., & Amemiya, Y. (2000). Estimation for polynomial structural equations. *Journal of the American Statistical Association*, 95, 929–940.
- Wall, M.M., & Amemiya, Y. (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. *Journal of Educational and Behavioral Statistics*, 26, 1–30.
- Wall, M.M., & Amemiya, Y. (2003). A method of moments technique for fitting interaction effects in structural equation models. *British Journal of Mathematical and Statistical Psychology*, 56, 47–63.
- Walters, B.A., & Bhuian, S.N. (2004). Complexity absorption and performance: A structural analysis of acute-care hospitals. *Journal of Management*, 30, 97–121.
- Wang, H., Law, K.S., Hackett, R.D., Wang, D., & Chen, Z.X. (2005). Leader member exchange as a mediator of the relationship between transformational leadership and followers' performance and organizational citizenship behavior. *Academy of Management Journal*, 48, 420–432.
- Whetten, D.A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14, 490–495.
- Williams, J., & MacKinnon, D.P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Structural Equation Modeling*, 15, 23–51.
- Williams, L., Edwards, J., & Vandenberg, R. (2003a). Recent advances in causal modeling methods for organizational and management research. *Journal of Management*, 29, 903–936.

- Williams, L., Gavin, M., & Hartman, N. (2004). Structural equation modeling methods in strategy research: Applications and issues. In Ketchen, D., & Bergh, D. (Eds.), *Research Methodology in Strategy and Management*, (pp. 303–346). Oxford, UK: Elsevier Science.
- Williams, L., & James, L. (1994). Causal models in organizational behavior research: From path analysis to LISREL and beyond. In J. Greenberg (Ed.), *Organizational behavior: The state of the science* (pp. 181–206). Lawrence Erlbaum.
- Williams, L., & O’Boyle Jr., E.H. (2008). Parceling strategies in structural equation modeling. *Human Resource and Management Review*, 18, 233–242.
- Williams, L., & O’Boyle Jr., E.H. (2009). *Goodness of fit of structural models: A review of reporting practices in organizational behavior*. Unpublished manuscript.
- Woehr, D.J., Sheehan, M.K., & Bennett, W. (2005). Assessing measurement equivalence across ratings sources: A multitrait–multirater approach. *Journal of Applied Psychology*, 90, 592–600.
- Yang-Wallentin, F., & Jöreskog, K.G. (2001). Robust standard errors and chi-squareds for interaction models. In G.A. Marcoulides & R.E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 159–171). Mahwah, NJ: Lawrence Erlbaum.
- Zedeck, S. (1971). Problems with the use of “moderator” variables. *Psychological Bulletin*, 76, 295–331.