Structural Equation Modeling: A Primer for Health Behavior Researchers

Eric R. Buhi, MPH, PhD; Patricia Goodson, PhD; Torsten B. Neilands, PhD

***Objective*: To introduce the state of the art of structural equation modeling (SEM). *Method*: This primer is organized in a manner allowing readers to review any one of 5 freestanding sections. *Results*: SEM maintains several advantages over regression and other multivariate techniques. Through a 2-step modeling process, SEM strengthens research by allowing for the specification of complex, theory-driven models that can be tested with empirical data. Although use of SEM alone is not a magic solution, new software developments provide users with unparalleled flexibility for improving research. *Conclusion*: SEM must be thrust into the daily vocabulary and routine practice of health behavior researchers.**

***Key words*: health behavior research, structural equation models, SEM, multivariate statistics, primer**

***Am J Health Behav.* 2007;31(1):74-85**

Structural equation modeling (SEM) is a powerful multivariate statistical method being used in the social sciences with increasing frequency.[1,2] In the psychological literature, SEM citations have risen since 1979 (Figure 1); SEM now rivals analysis of variance (ANOVA) in statistical method popularity.[3] In health behavior research, however, SEM has yet to reach such popularity.

In an electronic search of articles published between 1996 and 2004 in 3 health behavior research journals, we found only 7 reports in the *American Journal of Health Behavior*, 5 in the *American Journal of Health Promotion*, and 7 in *Health Education and Behavior* which utilized SEM (Table 1). This amounted to approximately 1 in 53 data-based journal articles using SEM. Furthermore, only a fraction of the journals' reports used some other multivariate technique, such as multivariate analysis of variance (MANOVA) or canonical correlation analysis (CCA; Table 1).

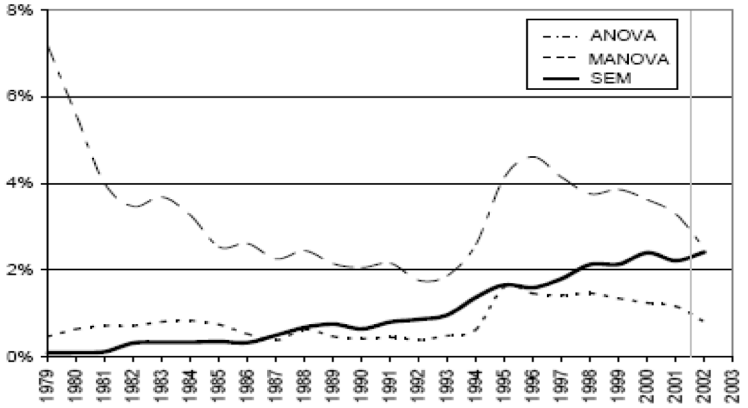
According to the American Academy of Health Behavior Work Group on Doctoral Research Training,[4] "a working knowledge of multivariate statistical procedures is crucial for generating high quality research and answering complex questions" (p. 554). Why, then, are health behavior and health promotion researchers not going beyond univariate/bivariate procedures, such as ANOVA and regression, in their research efforts? Although it is beyond the scope of the current paper to fully answer this question, we speculate there may be 3 reasons.

First, some researchers may not know *why* SEM is particularly useful or valuable. Because SEM is a relatively new tool, classes and trainings addressing

Eric R. Buhi, Assistant Professor, Department of Community and Family Health, University of South Florida, Tampa, FL. Patricia Goodson, Associate Professor, Department of Health and Kinesiology, Texas A&M University, College Station, TX. Torsten B. Neilands, Adjunct Assistant Professor, Center for AIDS Prevention Studies, University of California San Francisco, San Francisco, CA.

Address correspondence to Dr Buhi, Department of Community and Family Health, 13201 Bruce B. Downs Boulevard, MDC 56, University of South Florida, Tampa, FL 33612. E-mail: ebuhi@health.usf.edu

Figure 1
Citation Frequencies of SEM and (M)ANOVA in the PsychINFO Database Between 1979 and 2002



Note.
The numbers are standardized with respect to the total number of records per year. Figure reproduced with permission from MPR Online.

SEM's value have only recently been developed. Thus, researchers' knowledge of the assets afforded by SEM, such as the ability to account for measurement error in the modeling process, may be limited. Second, some researchers may not know *how* to conduct SEM. SEM analyses involve a distinctive nomenclature that can be intimidating for novices. Further, the complexities of models available within SEM analyses are mirrored in complex SEM software packages, which can also be intimidating for the uninitiated analyst. Third, some researchers may not be familiar with resource materials and computer programs available to aid in SEM analyses. The past 3 decades have been filled with rapid developments in SEM theory and software. For instance, the ability to model dichotomous dependent variables (eg, ever had sexual intercourse = yes/no) in SEM software packages such as M*plus* became widely available only within the last few years. Even if an individual took a graduate-level SEM course as recently as 5 years ago, he or she might not know about these developments.

The purpose of this paper, then, is to introduce—or update, depending upon the reader's familiarity—the state of the art in SEM as a multivariate analytic technique in health behavior research. This primer is organized in a manner allowing readers, both new and experienced SEM users, to review any one of 5 freestanding sections most pertinent to their needs or knowledge level. In section 1, we define the purpose of SEM; and in section 2, we present SEM's strengths as a multivariate analytic tool. These sections are most useful for individuals questioning why SEM might be valuable. For those unfamiliar with "how to do SEM," sections 3 and 4 are particularly helpful. In section 3, we provide an overview of the basic steps involved in conducting SEM analyses. In the fourth section, we present cautionary notes related to using SEM. Finally, in section 5, we review a host of available resources and provide a comparative treatment of SEM software packages. This concluding section may prove useful for beginners and seasoned "SEMers" alike. It will be most helpful for researchers seeking analytic materials or programs that can aid in SEM analyses.

Table 1
Utilization of Structural Equation Modeling and Other Multivariate Analytic Techniques in Published Articles From 3 Health Behavior Research Journals, 1996-2004

Journal Name	# of Articles Using SEM[a]	# of Articles Using Other Multivariate Techniques[b]	Total # of Articles Using Multivariate Techniques	# of Data-based Articles Published, Including in Supplements and Special Issues
American Journal of Health Behavior	7	24	31	416
American Journal of Health Promotion	5	15	20	325
Health Education and Behavior	7	14	21	275[c]

Note.
a Boolean search in MEDLINE/PsycINFO: (structural equation modeling) or (SEM) or (LISREL) or (AMOS) or (EQS) or (mplus) AND (American Journal of Health Behavior) AND (American Journal of Health Promotion) AND (Health Education and Behavior)
b Boolean search in MEDLINE/PsycINFO: (MANOVA) or (multivariate analysis of variance) or (cluster analysis) or (factor analysis) or (multidimensional scaling) or (CCA) or (canonical correlation analysis) or (discriminant analysis) AND (American Journal of Health Behavior) AND (American Journal of Health Promotion) AND (Health Education and Behavior)
c Issues from 1997 (volume 24, issue 5) to 2004. Prior to this volume/issue, journal was Health Education Quarterly, and we do not have access to these issues

1. What Is the Purpose of SEM?

Structural equation modeling includes a wide range of multivariate methods aimed at examining the underlying relationships, or structure, among variables in a model. SEM was created to test and refine theoretical models attempting to explain or predict social or behavioral phenomena.[5-7] Understanding these phenomena allows us to appreciate "why people engage in health-risk or health-compromising behavior and why (as well as how) they adopt health protective behavior."[8,p.1] These theoretical models inform the development and improvement of health-related interventions. Moreover, SEM is a useful tool in estimating these interventions' effects.[9]

Often referred to as causal, path, latent variable, or covariance structure models, SEM is similar to regression (and other correlational methods) because it belongs to the general linear model (GLM) family. For instance, SEM and regression analyses both rely on a linear combination of variables, use weights (eg, β weights) to optimize the explained variance and minimize model error variance, focus on latent (or not directly observed) variables, and yield variance-accounted-for effect sizes (eg, R^2, eta^2).[10] In short, SEM subsumes a range of other analytic methods[11,12] and may be utilized to conduct both simple analyses (including t-tests, ANOVA, and regression) as well as complex ones, such as multilevel modeling[13,14] (eg, examining youth-within-classrooms-within-schools) and latent curve modeling (which examines change as a continuous process over time).[14]

2. Why Use SEM?

Although SEM's increased application in the social sciences partially stems from improvements in software packages,[15] we argue such increasing use has been driven by 4 factors. First, multivariate methods such as SEM best honor the reality to which investigators are attempting to generalize.[16] In health behavior research, most outcomes (ie, behaviors) have multiple causes (ie, predictors), and most causes have multiple outcomes, all interacting dynamically. Health behavior researchers investigate multivariate, not univariate/bivariate or isolated, phenomena with only one or 2 determinants. It is impossible to assess how multiple variables behave in each other's company when a researcher limits an analysis to a univariate/bivariate examination. Instead, SEM allows all variables—multiple independent and dependent variables—to be examined simultaneously.

Second, multivariate methods such as SEM control for inflation of experimentwise (EW or Type I) error. Type I error is defined by alpha (α), usually set at .05, and is "the

probability of getting a result … that leads to an incorrect decision to reject the null hypothesis."[17,p.38] Inflated EW error may occur when a researcher conducts multiple univariate/bivariate tests (ie, with a single dependent variable or hypothesis, such as in ANOVA) with a single sample's data. These analyses can lead a researcher to falsely reject too many of the hypotheses being tested.[18] For example, assume a researcher conducts c number of tests in a single study (eg, 20 statistical significance tests), each at α = .05. Using the following formula, the risk of making a Type I error across the entire set of tests is 64%.[17]

$$^{\alpha}EW = 1 - (1 - \alpha)^{c}$$

$$^{\alpha}EW = 1 - (1 - .05)^{20} = .64$$

Further, out of 100 statistical tests conducted, the researcher could be rejecting the null hypothesis, incorrectly, 64 times. In these cases, the probability of making one or more Type I errors can be very serious (see Fish[19] for examples), and the implications for health behavior research can be sobering. If a researcher conducts multiple univariate/bivariate statistical tests to investigate determinants of an individual's involvement in a health-risk behavior, the researcher may erroneously (due to Type I error) conclude that a predictor is associated with the outcome when, in reality, it is *not*. Employing multivariate methods such as SEM, however, can correct "up front" for this analytic limitation by avoiding the use of multiple univariate/bivariate tests and, instead, testing hypotheses/research questions across several variables *at once*.

Third, researchers have begun to realize the utility of SEM over other multivariate analytic methods. SEM gives health behavior researchers unparalleled flexibility in specifying theory-driven models that can be tested with empirical data. SEM goes further than older multivariate techniques, such as MANOVA and CCA, by allowing users to automatically and efficiently compute indirect, direct, and total effects in complex models, including models that evaluate statistical mediation in which an exogenous predictor variable X impacts an intermediary variable Y that in turn exerts influence on a distal outcome Z. Additionally, unlike these older techniques in which the researcher basically "dumps in" all the variables, SEM allows researchers to test theories and assumptions directly by specifying which variables are related to other variables. That is, the researcher can test some paths (or relationships) but not others in the analysis. Some SEM programs even allow researchers to draw these hypothesized relationships visually and fit the drawn model to the underlying data, an intuitive and user-friendly process. Finally, SEM allows researchers to examine relationships among latent variables with multiple observed measures. The relationships among latent variables, thus, are purged of measurement error, leading to more accurate and often stronger relationships between latent variables than what would be observed using multivariate methods that consider observed variables only (eg, MANOVA or even regression). In short, although these older techniques assume zero measurement error in sample data (which is *never* the case), SEM controls for measurement error.

Lastly, SEM is useful because it enables the advanced treatment of incomplete data. Missing data in health behavior research can represent an important problem during analyses. SEM software developers have dealt well with the problem of missing data by incorporating sophisticated missing data techniques—such as optimal full information maximum likelihood (FIML)[20,21]—ahead of the general purpose software vendors (eg, SPSS, SAS, and Stata). Thus, ANOVA, regression, MANOVA, and ANCOVA can be conducted using SEM programs with or without incomplete data, and a researcher can thereby capitalize on the more sophisticated missing-data-handling capabilities (Table 2 lists SEM software programs capable of conducting FIML or multiple imputation). These capabilities allow researchers to proceed with SEM or other analyses as if there were no missing data, if certain missing data assumptions are met.[22,23]

3. What Are the Basic Steps of SEM?

In SEM, the researcher utilizes the theoretical literature to specify a health behavior model for testing. The researcher subsequently determines how to measure the variables pertinent to the theory and collects data, for instance, using a survey instrument. Next, he or she passes

Table 2
A Comparison of Various SEM Features by Program Package

Key Features	Program Name: Amos 5 www.spss.com/amos/	EQS 6 www.mvsoft.com/	M*plus* 3 www.statmodel.com/	LISREL 8 www.ssicentral.com/	PROC CALIS www.sas.com/	Mx http://www.vcu.edu/mx/
Has the ability to handle missing data via FIML or MI	×	A	×	×	B	×
Offers accessible/online technical support	×	×	×	×	×	×
Allows for graphical forming of models	×	×		×		×
Can conduct analyses with categorical outcome variables		×	×	×		×
Has the ability to fit multilevel or hierarchical SEM models		×	×	C	×	×
Has the ability to model non-normal continuous data	×	×	×	×	×	×
Has the ability to model non-normal continuous incomplete data		×	×			×
Has the ability to model categorical incomplete data			×			×
Computes direct, indirect, and total effects and associated asymmetric confidence intervals	D	E	×	×	E	×
Offers tests for multivariate normality	×	×	F	×	×	
Pricing[G] in dollars	599(Academic), 549(Government), 999(Commercial)G	595(Academic), 695(Government/ Corporate)G	595(Academic), 695(Commercial/ Government/ Non-profit)G	495(Academic and Commercial)G	Proc Calis is included with SAS. Contact local sales office through SAS website.	This program is freely available through the Mx website.
Comments	* Easy to use, quick learning curve * Provides models and model results in a graphical form, using a drawing tool, which may be useful for more "visual" analysts * Reads a wide variety of data file formats, including SPSS and Excel	* Provides models in a graphical form, using a drawing tool * Comes in Windows or Macintosh versions	* Very flexible overall * Excellent online technical support (Linda Muthén, one of the software authors, handles all online queries) * Can only read data from a text (.dat) file.	* Program syntax is sufficiently complex to require an unusually careful attention to detail on the part of the analyst * Detractors refer to its user interface and the level of sophistication. It is more complicated than other programs	* Can manage missing values with PROC MI[B] * Non-graphical format, requires knowledge of syntax	* Does many of the usual SEM analyses (and then some) and is free to try * Offers model fit indexes found in the major commercial programs, such as LISREL, EQS, and Amos * Can only read data from a text (.dat) file

Note.
A Implements ML-EM procedure, which is similar to Arbuckle's FIML.20
B Multiple imputation available through SAS PROC MI (available in version 9.x)
C LISREL 8.72 for Windows includes a multilevel SEM module which allows general 2-level structural equation models.
D Can compute as long as there are no missing data.
E Can compute direct, indirect, and total effects, but not asymmetric CIs.
F Available for mixture models only.
G Pricing is current as of summer 2005, and is listed for the basic-level package only. Add-ons, user's guides, and technical support may not be included.

data to an SEM software package, which fits the data to the specified model and produces results, including model fit statistics and parameter estimates.[24] SEM analyses entail essentially a 2-step modeling process[25] of building and testing (a) a measurement model and (b) a structural model. Although these 2 are the *funda-*

mental analytic steps, there is an additional "up front" step in the analytic process: examining the critical assumption of multivariate normality of the data.

Examining Multivariate Normality

Prior to any analyses, the researcher should test a critical assumption underlying both SEM and other multivariate techniques: the assumption regarding normality of the distribution of multivariate data. Before testing this assumption, however, analysts must first assess univariate normality because "normality on each of the variables [in a model] separately is a necessary, but not sufficient, condition for univariate normality to hold."[26,p.262] Testing multivariate normality can be accomplished (in both SPSS and SAS) graphically—by examining normal probability plots or Q-Q Plots—or nongraphically—by assessing skewness and kurtosis coefficients, or through statistical testing with the Shapiro-Wilk test.

Once the univariate normality assumption has been evaluated and met, the multivariate normality assumption can be assessed. Because ordinary least squares, generalized least squares, and maximum likelihood statistical estimation theories all presume a multivariate normal distribution, not meeting this assumption can be problematic, particularly when assessing statistical significance.[27] In SEM, nonnormality can result in an underestimation of overall model fit, downwardly biased parameter estimates, and underestimated standard errors.[28] Many SEM software packages such as EQS and Amos offer multivariate normality tests, such as Mardia's measure of multivariate kurtosis,[29,30] which can be carried out with a single mouse click. Recent advances in many packages even make it possible for researchers to analyze nonnormal continuous and categorical data (Table 2).

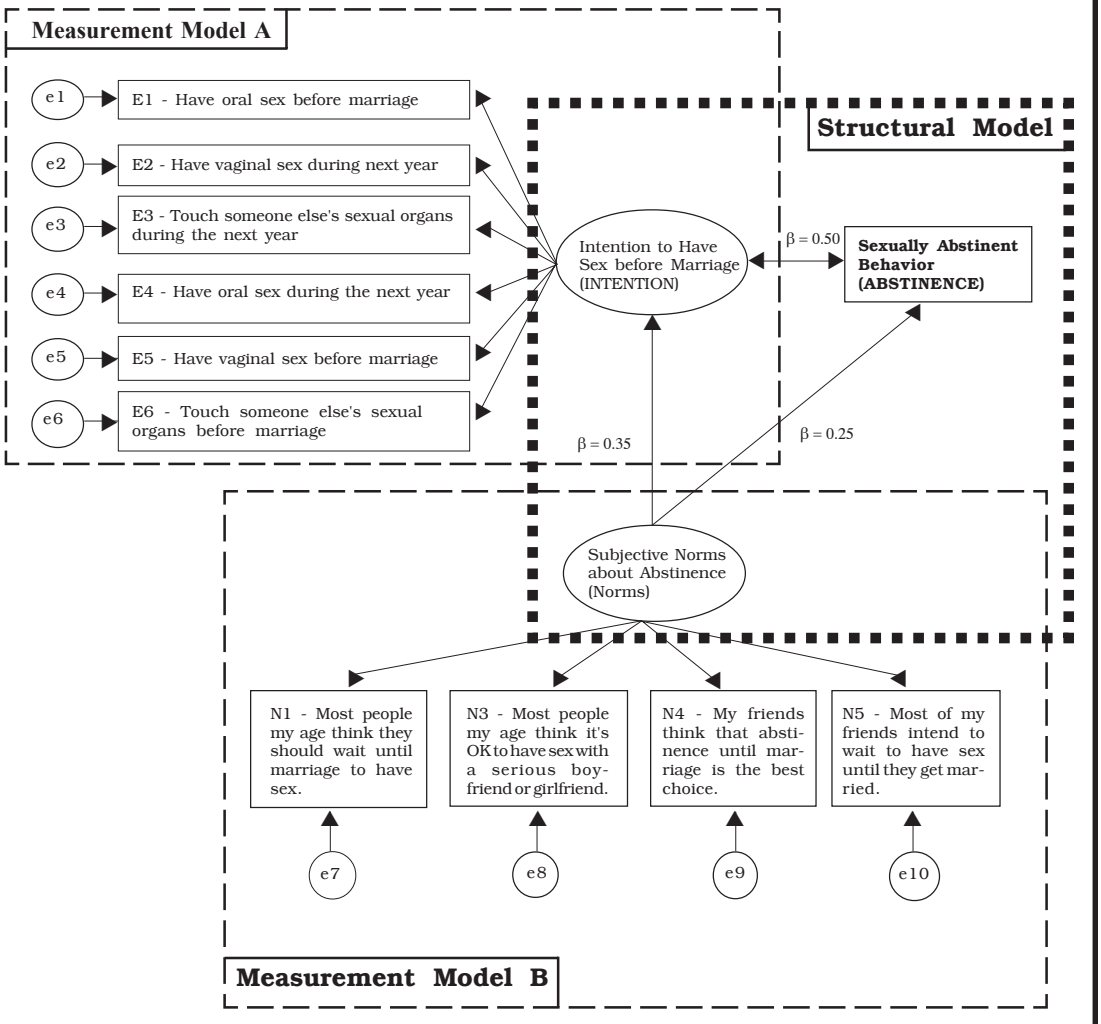
Using the 2-Step Modeling Approach

Once assumptions are evaluated and met, the researcher may commence with the 2-step modeling process of building and testing a (a) measurement model and (b) structural model. First, the researcher begins by building and testing the measurement model within the SEM software package. The purpose of the measurement model step is to test indicator/construct relationships. In a psychometric sense, this step is imperative in ascertaining the validity of the constructs. In Figure 2, for example, 2 sample measurement models (within the narrow dotted boxes) have been formulated using a drawing tool found in one SEM package, Amos. This measurement model's purpose is to take into account measurement error in all variables which are not *directly* observable (eg, the constructs of behavioral intentions, perceived norms, attitudes, or motivation).[7] In short, the shared variance derived from the correlations/covariances among multiple observed variables, such as survey items (the boxes in Figure 2), is used to infer the presence of a common latent factor (the ovals in the model). The software package converts survey data from Excel, SPSS, or raw data (in ASCII text format) into covariances and means which are then used in subsequent analyses.[24] In the Figure 2 sample model, the boxes—items E1-E6, N1-N5, and ABSTINENCE—are observed measurement items, or indicators (E1-E6 and N1-N5 are scales). The ovals in the model—INTENTION and NORMS—are the latent factors, or constructs being represented by their respective scales.

The measurement model step is equivalent to performing a confirmatory factor analysis. In this approach, the numbers of hypothesized underlying constructs (ie, factors) are specified by the literature (the theory under examination) and researcher *a priori*, and the model is fitted to sample data to assess its convergent and discriminant validity (which together provide evidence for construct validity). Ideally, after an acceptable fit is achieved, the measurement model is then cross-validated using a second set of sample data. In other words, in this step the researcher assesses how well the scales measure the latent constructs which will be included in the structural model.

Once the measurement model has been formulated and tested, a structural model is estimated as the second step. In a structural model, the goal is to examine the underlying relationship, or structure, between the latent constructs tested in the measurement model and other (observed) variables proposed by the theory (see Figure 2, within the wide-dotted box). This structure accounts for the direct, indirect, and total effects among factors.[31] A *direct effect* is the directional relation-

Figure 2
Sample Measurement and Structural Model



ship between 2 variables and is the type of relationship usually examined through ANOVA and regression. An *indirect effect* is an independent variable's influence on a dependent variable, through a single or possibly multiple mediating variables.[32] The *standardized indirect effect* is the product of the standardized direct effects. Using the model in Figure 2 as an example, if NORMS has a direct effect on INTENTION (β=0.35), and INTENTION has a direct effect on ABSTINENCE (β=0.50), then NORMS can be said to have an indirect effect on ABSTINENCE (NORMS→INTENTION*INTENTION→ABSTINENCE = 0.175). To compute the total effect of NORMS on ABSTINENCE, one would take the sum of the NORMS→ABSTINENCE direct effect and the NORMS→INTENTION*INTENTION→ABSTINENCE indirect effect (0.25+0.175=0.425).

To test the fit of sample data to the structural model, SEM software examines covariances rather than individual cases (as happens in regression techniques). To examine relationships in SEM,

matrix algebra is used to account for variances of each variable and covariances of each pair of variables.[26] This covariation makes SEM a more applicable and generalizable technique than regression, allowing for the simultaneous examination of multiple independent and dependent variables. It is important to bear in mind that these variables are known in SEM, respectively, as exogenous and endogenous variables. The name *exogenous variable* means that the cause of the variable is determined outside of the specified model. *Endogenous variables*, on the other hand, are determined within the model (that is, endogenous variables are hypothesized to be predicted by other variables in the model).

Assessing Model Fit

A strength of SEM is that the analyst obtains both a global assessment of model fit and tests of individual parameters. The researcher begins by evaluating global model fit. Quantifying the correspondence between the predicted covariances and the observed covariances (which is the analytic focus of SEM) generates a *goodness-of-fit* value or *index*. In classical statistics, effect sizes characterize the fit of a model to data (eg, R^2 for a regression model). Similarly, in SEM, fit indexes may be thought of as effect sizes. Consulting these indexes and checking for model fit can lead to important model improvements. For instance, specific paths can be re-drawn to hypothesize new relationships or the entire model can be respecified to exclude factors with weak explanatory power.

Although there are a number of fit indexes available, unfortunately there is not one index appropriate for all analytic conditions. There are, nevertheless, general fit index "rules of thumb" to consider with recommended cutoff values. Hu and Bentler,[33] for example, suggest that researchers always examine and report chi-square (χ^2) for *exact fit*, which tests whether there is a statistically significant difference between the model and the sample data and degrees of freedom (df) for each model estimated. Because χ^2 can be heavily influenced by sample size, however, the χ^2/df ratio may be reported. According to Bollen,[31] there is little consensus, for the χ^2/df ratio, on what represents a "good fit," with recommendations as high as 5 and as low as 2 (or less). In health behavior research studies, ratios between 2 and 5 have often been employed.

Hu and Bentler[33] further recommend assessing and reporting results from several *approximate fit* indexes because one or more are insensitive to sample size and/or impervious to estimation methods (eg, maximum likelihood or generalized least squares). The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI), for example, indicate how much better the model fits the sample data than a null model, which stipulates that there are no common factors. Unlike the CFI, however, the TLI moderately corrects for model parsimony. The Root Mean Square Error of Approximation (RMSEA) is similar to the chi-square test in that it is more or less a "badness of fit" test. Possible values for these fit indexes range between 0 and 1, although TLI can exceed 1. With regards to what a good fit actually means, there is little consensus. Cutoff values of 0.95 for TLI/CFI (the higher the better) and .06 for RMSEA (the lower the better) have been suggested.[34] However, Browne and Cudeck[35] suggested that fair fitting models have RMSEA values in the range of 0.05 to 0.08, even as Marsh et al[36] cautioned against broadly applying these "golden rules of fit" without first considering their limitations. In brief, then, values further away from the recommended cutoff points indicate potential inconsistency between the model and sample data, whereas values near the recommendations suggest that the model might be useful. Once global fit is met, then the researcher examines individual parameter estimates and confidence intervals to learn which paths denote the strongest relationships or explain the greatest amount of model variance.

In sum, the theory the researcher is examining guides the measurement and structural modeling process by specifying (a) the number of underlying constructs in the model and (b) the interrelationships of these constructs. Results produced by the software package, such as model fit statistics and parameter estimates, are used to test and improve overall model fit. For good examples of this modeling process, see Blue et al,[37] Park et al,[38] and Saunders et al.[39]

4. What Cautions Should SEM Users Exercise?

Although SEM is a sophisticated ana-

lytic tool for testing theoretical models with multiple endogenous/exogenous variables, its application alone does not resolve (or even address) the limitations of behavioral and social science research. When considering SEM, researchers must exercise the following 4 cautions concerning its utilization.

SEM Does Not Compensate for Poorly Conceived Ideas or Weak Theoretical Grounding

Statistics is an important tool in data analysis, but it represents only one of its components. Logical reasoning is another vital data analysis component. Critical as it is, the use of statistical techniques is becoming separated from the sound manipulation of ideas,[40] in part due to the rapid development of powerful computer programs. Progress in science, however, is significantly hampered when researchers use, as their guides, implausible theoretical reasoning, frameworks, or models. Although SEM is a more elegant analytic technique than univariate/bivariate methods, as a statistical tool it will never compensate for bad logic and poor ideas/models. As Kenny stated in his seminal contribution, *Correlation and Causality*,[6] "Causal modeling provides no certain path to knowledge. In fact, models are maximally helpful only when good ideas are tested. Good ideas do not come out of computer packages, but from people's heads" (p. 8). In short, SEM should be used simply as a tool for testing carefully thought-out ideas that are empirically grounded and/or theoretically generated.

SEM Analyses Are Correlational

SEM is misleadingly called *causal* modeling, but as part of the GLM family, it is still a correlational method. Inferring causation requires more than simply *employing* SEM; instead, a number of conditions must first be met. The most basic condition is that an *association* must exist between the variables postulated to have the cause-and-effect relationship. Second, temporal priority, temporal ordering, or directionality[31] of the causal relationship must be established (ie, the cause must *precede* the effect). Rather than being an analytic issue, temporal priority is primarily a methodological matter. That is, to ensure the cause precedes the effect, data must be collected over time. Lastly, a single variable Y must be *isolated* from all influences, with exception of a second variable X. If a change in X accompanies a change in Y, then it can be said that Y *causes* X. In reality (eg, in health behavior research), however, the isolation of Y from other variables of influence is virtually impossible. Thus, according to Bollen,[31] all models must be looked upon as estimations of reality. So, SEM may give an *indication* of causal relations but, by itself, SEM cannot ensure that association, temporal priority, and isolation have been met.[41] Rather, rigorous methodological planning and implementation in research must accompany use of SEM to increase confidence that causality is being observed (or established).

Model Parsimony Should Be a Top Priority

When building a model, parsimony should be a top priority for researchers. A parsimonious model consists of the fewest number of variables, explaining the greatest possible amount of variance in the outcome(s) of interest. In short, the *simplest* model is the best model. In the second section, however, we noted that health behavior researchers should stay true to the examination of the complexities (ie, multivariate nature) of human behavior. These complexities force researchers to measure all possible causes and consequences of the health behavior under study. Thus, the resulting model may be one of enormous proportion (ie, nonparsimonious), with a massive number of variables accounting for these various influences and outcomes. A number of fit statistics, however, penalize the researcher for model complexity, or this lack of parsimony. The challenge for researchers, therefore, is to capture the complexity of human behavior using the fewest number of variables possible. One solution to the model parsimony issue is to include large numbers of observed variables, measured by the smallest possible number of latent factors with the fewest possible number of structural paths. See Sivo and Willson[42] for a more detailed discussion of model parsimony.

Sample-size Issues

Having an adequate study sample size can be a major concern in SEM utilization. In general, small samples are more likely to result in unreliable and untrust-

worthy parameter estimates and fit statistics, yielding models which are nonreplicable. West et al[28] noted that decreasing sample size leads to an increase in the probability that analyses will fail to converge or result in inappropriate solutions. What, then, constitutes an *adequate* sample size? Several authors[43,44] have suggested that at least 200 cases are necessary for adequate model specification. Stevens[26] noted that 15 cases per predictor in standard ordinary least squares regression is a good rule of thumb. Because SEM and regression are similar in many respects, 15 cases per measured variable in SEM is not unreasonable.[24] In reality, however, there is no ideal sample size for all situations. Adequate sample size may depend on the complexity of the model being tested and the statistical estimator used. For instance, more complex models and those with multiple indirect effects may need a greater number of cases. For confirmatory factor analyses, Flora and Curran[45] provided evidence that one can use weighted least-squares estimators with sample sizes as small as 100 cases. The interested reader should refer to other sources[34,43] for a more detailed discussion of SEM and sample-size issues. See also Muthén and Muthén[46] for a demonstration of how researchers can use a Monte Carlo (or simulation) study to decide on sample size and determine power for SEMs, using M*plus*.

5. What Resources are Available to SEM Users?

Aids and resources abound for both the novice and experienced SEM user, including books, journals, e-mail discussion lists, and statistical software packages. For beginners, there are a number of textbooks[7,32] and book chapters[47,48] providing useful overviews. Some texts even present user-friendly introductions to specific SEM software packages, such as Amos,[49] EQS,[50] and LISERL.[51] For seasoned SEMers, Bollen's[31] encyclopedic reference may serve as a key resource.

In 1994, *Structural Equation Modeling: A Multidisciplinary Journal* began quarterly publication. This journal has, since, served as the flagship peer-reviewed periodical for researchers utilizing SEM analyses in various disciplines, including health/medicine, psychology, education, economics, sociology, business, and political science. Contents of the journal include theoretical, methodological, and applied pieces; book reviews; software package reviews; and a teacher's corner with instructional modules.

Supplementing the various text and journal resources is an electronic mail network for SEMers called SEMNET. Begun in 1993, and owned by Dr Carl E. Ferguson Jr, professor of marketing at the University of Alabama, SEMNET serves as an open forum for ideas and questions regarding analysis of covariance structures, path analysis, and confirmatory factor analysis. SEMNET archives can be searched at http://bama.ua.edu/archives/semnet.html, and additional information can be retrieved from the SEMNET information site (http://www2.gsu.edu/~mkteer/semnet.html).

Finally, numerous software packages exist for SEM analyses, including Amos (Analysis of Moment Structures), EQS (Equations), M*plus*, LISREL (Linear Structural Relationships), CALIS (Covariance Analysis and Linear Structural Equations, in SAS), and Mx. These programs, historically known for their complex code commands and large computer memory/space requirements, are now much more user-friendly and accessible. Most programs still allow users to write code containing matrix algebra commands. However, many packages such as Amos and EQS have graphical interface options, allowing analysts to draw their measurement/structural models on the computer screen and tap into a data set to generate output. Most range in cost from $500-600 for the basic-level package, which may exclude add-on features, user's guides, and technical support. Mx, however, is freely available for download through the WWW, and the package can do the usual SEM analyses and more. For a fee (usually between $900 and 1200), individuals can enroll in training courses on conducting SEM with various software packages, which are held in major cities and on college campuses. These trainings are regularly offered through professional associations (see the APA Advanced Training Institutes online at http://www.apa.org/science/ati.html), universities (see the University of Kansas Continuing Education website at http://www.continuinged.ku.edu/programs/rda/index.php), and software manufacturers (see the SPSS/Amos website at http://www.spss.com/training/ or the

M*plus* training site at http://www.statmodel.com/courses.html). Table 2 summarizes the more commonly used SEM applications, including resource and pricing information as well as specific strengths and weaknesses regarding each package.

CONCLUSION

We have organized this primer to allow readers, both new and experienced users of SEM, to review any one of the 5 free-standing sections most pertinent to their needs or knowledge level. It is our hope that this paper has familiarized health behavior researchers with the purpose of SEM, *why* SEM is valuable as an analytic technique, *how* SEM is conducted, and cautionary notes related to using SEM. We also hope that we have equipped the interested reader with the necessary resources and information regarding SEM analyses and available software packages. Finally, we hope that this primer will thrust SEM into the daily vocabulary and, most importantly, into the routine practice of health behavior and health promotion researchers. The generation of high-quality research depends upon it. ■

REFERENCES

1. Tremblay PF, Gardner RC. On the growth of structural equation modeling in psychological journals. *Structural Equation Modeling*. 1996;3:93-104.
2. Hershberger SL. The growth of structural equation modeling: 1994-2001. *Structural Equation Modeling*. 2003;10:35-46.
3. Nachtigall C, Kroehne U, Funke F, et al. (Why) should we use SEM? Pros and cons of structural equation modeling. *Methods of Psychological Research*. 2003;8:1-22.
4. American Academy of Health Behavior Work Group on Doctoral Research Training. A vision for doctoral research training in health behavior: a position paper from the American Academy of Health Behavior. *Am J Health Behav*. 2005;29:542-556.
5. Bentler PM. Causal modeling via structural equation systems. In: Nesselroade JR, Cattell RB, eds. Handbook of Multivariate Experimental Psychology (2nd ed.). New York: Plenum Press; 1988:317-335.
6. Kenny DA. Correlation and Causality. New York: John Wiley & Sons; 1979.
7. Raykov T, Marcoulides GA. A First Course in Structural Equation Modeling. Mahwah, NJ: Lawrence Erlbaum; 2000.
8. Crosby RA, Kegler MC, DiClemente RJ. Understanding and applying theory in health promotion practice and research. In: DiClemente RJ, Crosby RA, Kegler MC, eds. Emerging Theories in Health Promotion Practice and Research: Strategies for Improving Public Health. San Francisco: Jossey-Bass; 2002:1-15.
9. Short LM, Hennessy M. Using structural equations to estimate effects of behavioral interventions. *Structural Equation Modeling*. 1994;1:68-81.
10. Thompson B. Five methodology errors in educational research: the pantheon of statistical significance and other faux pas. Invited address at the annual meeting of the American Educational Research Association, San Diego, CA. 1998, April. (ERIC Document Reproduction Service No. ED 419 023)
11. Bagozzi RP, Fornell C, Larcker DF. Canonical correlation analysis as a special case of a structural relations model. *Multivariate Behav Res*. 1981;16:437-454.
12. Fan X. Canonical correlation analysis and structural equation modeling: what do they have in common? *Structural Equation Modeling*. 1997;4:65-79.
13. Bauer DJ. Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*. 2003;28:135-167.
14. Hox J. Multilevel Analysis: Techniques and Applications. Mahwah, NJ: Lawrence Erlbaum; 2002:251-274.
15. Mueller RO. Structural equation modeling: back to basics. *Structural Equation Modeling*. 1997;4:353-369.
16. Thompson B. Why multivariate methods are usually vital in research: some basic concepts. Paper presented at the biennial meeting of the Southwestern Society for Research in Human Development, Austin, TX. 1994, February. (ERIC Document Reproduction Service No. ED 367 687)
17. Kline RB. Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research. Washington, DC: American Psychological Association; 2004.
18. Huberty CJ, Morris JD. Multivariate analysis versus multiple univariate analyses. *Psychol Bull*. 1989;105:302-308.
19. Fish LJ. Why multivariate methods are usually vital. *Measurement and Evaluation in Counseling and Development*. 1988;21:130-137.
20. Arbuckle JL. Full information estimation in the presence of incomplete data. In: Marcoulides GA, Schumacker RE, eds. Advanced Structural Equation Modeling: Issues and Techniques. Mahwah, NJ: Lawrence Erlbaum; 1996:243-277.
21. Wothke W. Longitudinal and multigroup modeling with missing data. In: Little TD, Schnabel KU, eds. Modeling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches, and Specific Examples. Mahwah, NJ: Lawrence Erlbaum Associates; 2000:219-240,269-281.
22. Little RJA, Rubin DB. Statistical Analysis

with Missing Data (2nd Ed.). Hoboken, NJ: John Wiley & Sons, Inc; 2002.
23. Allison PD. Missing Data. Thousand Oaks, CA: Sage; 2002.
24. Research Consulting, ITS, UT. Structural Equation Modeling using AMOS: An Introduction. Available at: http://ww.utexas.edu/its/rc/tutorials/stat/amos/. Accessed October 25, 2005.
25. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. *Psychol Methods*. 1988;103:411-423.
26. Stevens JP. Applied Multivariate Statistics for the Social Sciences. Mahwah, NJ: Lawrence Erlbaum; 2002:262.
27. Henson RK. Multivariate normality: what is it and how is it assessed? *Advances in Social Science Methodology*. 1999;5:193-211.
28. West SG, Finch JF, Curran PJ. Structural equation models with nonnormal variables: problems and remedies. In: Hoyle RH, ed. Structural Equation Modeling: Concepts, Issues, and Applications. London: Sage; 1995:56-75.
29. Mardia KV, Kanazawa M. The null distribution of multivariate kurtosis. *Communications in Statistics: Simulation and Computation*. 1983;12:569-576.
30. Yuan K, Lambert PL, Fouladi RT. Mardia's multivariate kurtosis with missing data. *Multivariate Behav Res*. 2004;39:413-437.
31. Bollen KA. Structural Equations with Latent Variables. New York: John Wiley & Sons, Inc; 1989.
32. Hoyle RH, Panter AT. Writing about structural equation models. In: Hoyle RH, ed. Structural Equation Modeling: Concepts, Issues, and Applications. London: Sage; 1995:158-176.
33. Hu L, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparamerterized model misspecification. *Psychol Methods*. 1998;3:424-253.
34. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999;6:1-55.
35. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. Testing Structural Models. Newbury Park, CA: Sage Publications, Inc; 1993:136-162.
36. Marsh HW, Hau K, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*. 2004;11:320-341.
37. Blue CL, Black DR, Conrad K, et al. Beliefs of blue-collar workers: stage of readiness for exercise. *Am J Health Behav*. 2003;27:408-420.
38. Park K, Wilson MG, Lee MS. Effects of social support at work on depression and organizational productivity. *Am J Health Behav*. 2004;28:444-455.
39. Saunders RP, Motl RW, Dowda M, et al. Comparison of social variables for understanding physical activity in adolescent girls. *Am J Health Behav*. 2004;28:426-436.
40. Aneshensel CS. Theory-Based Data Analysis for the Social Sciences. Thousand Oaks, CA: Pine Forge Press; 2002.
41. Bullock HE, Harlow LL, Mulaik SA. Causation issues in structural equation modeling research. *Structural Equation Modeling*. 1994;1:253-267.
42. Sivo SA, Willson VL. Is parsimony always desirable? Identifying the correct model for a longitudinal panel data set. *Journal of Experimental Education*. 1998;66:249-255.
43. Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*. 1999;6:56-83.
44. Chou C, Bentler PM. Estimates and tests in structural equation modeling. In: Hoyle RH, ed. Structural Equation Modeling: Concepts, Issues, and Applications. London: Sage; 1995:37-55.
45. Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol Methods*. 2004;9:466-491.
46. Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*. 2002;9:599-620.
47. Klem L. Structural equation modeling. In: Grimm LG, Yarnold PR, eds. Reading and Understanding MORE Multivariate Statistics. Washington, DC: American Psychological Association; 2000:227-260.
48. Thompson B. Ten commandments of structural equation modeling. In: Grimm LG, Yarnold PR, eds. Reading and Understanding MORE Multivariate Statistics. Washington, DC: American Psychological Association; 2000:261-283.
49. Byrne BM. Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
50. Byrne SM. Structural Equation Modeling with EQS and EQS/Windows: Basic Concepts, Applications, and Programming. Thousand Oaks, CA: Sage; 1994.
51. Byrne BM. Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming. Mahwah, NJ: Lawrence Erlbaum Associates; 1998.