Linear Structural Relationships (LISREL) in Family Research

Yoav Lavee


Stable URL:
http://links.jstor.org/sici?sici=0022-2445%28198811%2950%3A4%3C937%3ALSR%28IF%3E2.0.CO%3B2-R

*Journal of Marriage and the Family* is currently published by National Council on Family Relations.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/NCFR.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
Linear Structural Relationships (LISREL) in Family Research

Over the past two decades there has been a growing interest among social scientists in testing multivariate theoretical models. Most frequently, the effect of explanatory variables on a response variable, as well as their indirect effect through some intervening variables, would have been estimated by path analysis. In recent years, however, the validity of this approach for analyzing structural models has been questioned. First, it has been recognized that the underlying assumptions of measurement without error and uncorrelated residuals are unrealistic in the social sciences. Second, many of the variables of interest are unobservable, complex constructs, which are difficult to capture validly and reliably with single indicators.

The concern about the reliability and validity of empirical measurements, and the need to formulate a strategy for studying structural relationships among variables that better represent theoretical constructs, have led to the development of a new approach for studying structural models. This approach is known as latent variable structural equation models or as analysis of covariance structure models (Long, 1983). To many, the approach is known simply as LISREL (Linear Structural RELationships), after the statistical analytic technique and computer program developed by Jöreskog and his colleagues (Jöreskog, 1973, 1978; Jöreskog and Sörbom, 1984) to analyze covariance structure models. LISREL is a versatile and powerful method that combines features of factor analysis and multiple regression for studying both the measurement and the structural properties of theoretical models. It allows the estimation of causal relationships among latent (unobserved) variables, and permits for measurement errors and correlated residuals.

Since the late 1970s and early 1980s, the number of researchers using a covariance structure modeling approach (primarily LISREL) has been growing steadily in various social science fields, such as sociology, psychology, communication, education, child development, marketing, and political science. Similarly, its application to the study of marriages and families is increasing; during the past five years, more than a dozen studies using this approach have appeared in the Journal of Marriage and the Family alone.

Unfortunately, the method is complex and hard to learn. It is based on mathematical and statistical approaches (matrix algebra and maximum-likelihood function) with which many family researchers are not familiar. It is also somewhat hard to use, and researchers often struggle both with a new language and with numerous practical questions.

The purpose of this article is to address some of the common issues with which family investigators deal. It does not offer an exhaustive review of covariance structure modeling, but rather it focuses on problems that researchers often face, questions that they raise, and decisions they need to make in the process of model specification, estimation, modification, and hypothesis testing. Occasionally, these issues (and some solutions) are exemplified by empirical studies published in the family literature. But first, the core ideas behind the LISREL approach
are briefly described in order to facilitate the discussion that follows.²

**Analysis of Covariance Structures: An Overview**

The general LISREL model combines the confirmatory factor analytic model and the structural equation model. Generally speaking, factor models deal with measurement properties of constructs by estimating the common and the unique variance of sets of measured variables. In the commonly used exploratory factor analysis, for example, researchers seek the common factors that underlie a set of measured variables (or items of a measurement instrument). A common factor, then, is a latent, unobserved "entity," the meaning of which is inferred from the measured variables (indicators) it underlies. Each measured variable, in turn, has its own unique variance (something that is not shared in common with the other indicators of the latent variable), including a random measurement error. In exploratory factor analysis, researchers seek these latent variables, that is, the number of factors and the structure of relationships between each measured variable and the latent variables. Recall also that, when more than one factor underlies a set of measured variables, researchers either constrain the covariation between them to zero (in orthogonal solution) or they allow the factors to covary (oblique solution). Factor analytic models cannot, however, make a statement about the structural relationships among the unobserved factors.

Structural equation models, on the other hand, are concerned with the structural (causal) relationships among variables. In multiple regression or in path analysis, for instance, researchers attempt to estimate the effect of independent on dependent variables, either directly or through intervening variables. These models assume that the variables in a theoretical model are directly observed and that they are measured without error.

In covariance structure models, the two approaches are combined into a single model that includes (a) a measurement model, which defines the latent variables a priori in terms of their specified measured indicators, and (b) a structural model, in which the structural relationships among the latent variables are estimated. Although it is convenient to describe covariance structure models in terms of these two components (measurement and structural models), in practice they are often estimated simultaneously by using a full-information maximum-likelihood confirmatory factor analysis. Essentially, this means that the loadings of the measured variables on their respective factors, the error terms of the measured variables, the relationships between the exogenous (independent) and endogenous (dependent) variables, the relationship among the endogenous variables, and the disturbances (errors in equation) of the latent endogenous variables are all estimated simultaneously. Furthermore, these parameters are estimated in such a way that a matrix of variances and covariances reproduced from the estimated coefficients corresponds to the relationships among the measured variables (the observed variance-covariance matrix) as closely as possible. A number of goodness-of-fit tests are provided to estimate how well the model fits the data, and additional tests are provided for assessing the statistical significance of each of the estimated parameters. Additionally, the computer program (LISREL) provides diagnostic information on the model that suggests (on statistical grounds) what modifications may be needed in the specifications of the model that would improve the fit of the model to the data.

**The Role of the Researcher**

Although the LISREL technique is versatile, powerful, and highly flexible statistically, its utility for any particular research project depends upon the researcher's thoughtful use of theory at every phase of the investigation. The primary function of covariance structure analysis is to test hypothesized causal relationships among theoretical constructs (although, as will be shown later, it may also be used for exploratory purposes and as a means for refining theory). Theoretical reasoning should guide the researcher prior to the analysis, in specifying the hypothesized model, as well as after the estimation, in evaluating the results and introducing any modifications.

The process of covariance structure analysis involves a series of steps that researchers are advised to follow sequentially (see Lomax, 1982). These steps may be summarized along the phases of model specification, estimation, modification, and hypothesis testing. At each step, researchers are called upon to make some decisions: How is a model built? What and how many indicators are
needed for each latent variable? How should measurement errors of single indicators be handled? How many cases are needed to estimate a model? Should a correlation matrix be used, or an unstandardized variance-covariance matrix? Should the measurement model be estimated simultaneously with, or separate from, the structural model? How can an unidentified model be rendered testable? Should the model be specified if it doesn’t fit the data? When is the process completed? In the following sections, these and other issues are addressed as we take a journey through the steps of covariance structure analysis, from formulating a model to reporting the results.

**Model Specification**

**Step 1. Construct the hypothesized structural model.**

Specification of a model is an explicit translation of theory into mathematical equations. It is useful to begin by drawing a path diagram that shows the exogenous (independent) variables, the endogenous (dependent and intervening) variables, and the proposed causal relationships between them.

To make an adequate translation of theory into empirical model, one should make sure that (a) all the relevant constructs are being considered simultaneously (external specification), and (b) all their uni- and bidirectional interrelations are stated explicitly. Thus, one should decide, on the basis of the theory being tested, whether the relationships are only unidirectional (recursive model), or whether reciprocal relationships better represent the theory in question.

At the stage of constructing their model, researchers often struggle with issues of model complexity. Theories are multivariate and often involve complex (and sometimes ambiguous) causal processes. While one would like to capture this complexity in an empirical model, too many variables or too complex a model may render the model practically untestable. The dilemma for researchers, therefore, is that the exclusion of theoretically relevant variables may violate the requirements for valid external specifications and may bias the parameter estimates (see Godwin, 1988, in this issue). On the other hand, if one includes a large number of variables and specifies multiple reciprocal relationships, one may be unable to fit the model to the data (Kenny, 1979). An attempt should therefore be made to include plausible causal variables and specify reciprocal relationship between variables when it is so proposed by a theory, but at the same time balance the ideal of a fully comprehensive model with practical considerations (Bentler and Chou, 1987).

In family research, models ranged from relatively simple, recursive models with three constructs (e.g., Thomson and Williams, 1982, 1984; Wright and Price, 1986) to fairly complex multivariate models (Acock and Edwards, 1982; Roosa, Fitzgerald, and Carlson, 1982) and nonrecursive models (e.g., Acock and Yang, 1984; Honeycutt, 1986; Mirowsky and Ross, 1987).

**Step 2. Select indicator variables for the constructs (latent variables).**

The specification of the relations between constructs (Step 1 above) can be done without anchoring the constructs in measurement operations. Specification of the latent variables, however, should ensure that a theory’s constructs are in fact embedded in the model (Bentler, 1980). What is an adequate specification? Are multiple indicators of a latent variable better than a single indicator? Under what conditions?

Generally, the use of multiple indicators to measure a construct is preferred. Multiple indicators are more likely to be able to capture a complex theoretical construct than a single measure is. For example, specification of the construct marital adjustment (Spanier, 1976) as a latent variable with three measured indicators—consensus, satisfaction, and cohesion (see Ladewig and McGee, 1986)—provides a more accurate operational translation of this theoretical construct than any single, global measure of marital satisfaction or a simple sum of the three measures. Additionally, the error term of measured variables can be estimated only when multiple indicators are specified, and only then can a latent variable be treated as a “true,” errorless variable.

Some scholars have maintained that multiple indicators, because they reduce the error in the latent variables, tend to inflate the regression coefficients of the structural parameters (see Biddle and Martin, 1987), and therefore they are making “something out of nothing.” Although
this view is uncommon, it raises the question of whether multiple indicators should always be preferred. Specifically, researchers sometimes estimate models in which some (or all of the "constructs") have not been developed theoretically and are not well defined. To be able to estimate a "real" LISREL model, with latent variables and measurement errors, they attach a set of variables that somehow "makes sense" to a hypothetical latent variable. This may be a questionable practice, since the structural coefficients may be inflated artificially because of a latent variable that does not actually represent a theoretical construct.

Let us emphasize: since the latent variables are abstractions that presumably underlie measured variables, the specification of the measurement model should be guided by theoretical reasoning. One's theory must support the existence of a latent variable, the theoretical meaning of the latent variable must be considered, and the measured indicators must be chosen carefully.¹

How many indicators are needed for each latent variable? Bentler (1980: 425) made the observation that "choosing the right number of indicators for each latent variable is something of an art: in principle, the more the better; in practice, too many indicators make it difficult if not impossible to fit the model to data." In general, three or more indicators are often recommended because models with only two indicators of a latent variable may, in some cases, be underidentified (see Bentler and Chou, 1987: 102).

Despite the clear advantage of using multiple indicators of latent variables, many researchers specify their models with single indicators only or, more frequently, with both single and multiple indicators of latent variables. This practice may be appropriate in some applications and inadequate in others. The use of a single measure is well justified when it is used to indicate a relatively simple, measurable variable, such as age, religious preference, or annual income. In this case, the measured variable represents the "construct" sufficiently well. If the measure may be assumed errorless, the latent variable is typically made isomorphic with the indicator by fixing the loading to 1 and the measured variable's error to 0.

The use of a single measure to indicate a more complex construct (such as work commitment, role strain, or parent-child relationship) poses a doubt as to the theoretical construct being studied as well as some methodological problems. One such problem has to do with measurement error. Recall that when multiple indicators are specified, the program estimates the errors in measures, but this cannot be done when a single indicator is used. When a theoretical construct is specified by a single indicator (with no measurement error), the measurement model may be wrong (i.e., it assumes that the hypothetical construct is directly observed and that the measure is fully reliable), and the resulting estimates of the structural parameters are likely to be biased. It is therefore recommended in such cases to use an external criterion of reliability (such as Cronbach's alpha) to constrain the relation between the measured variable and the latent variable to the value of the estimated "true" variance (i.e., its known reliability), and to specify the "unreliable" portion of the measured variable as a measurement error (Jöreskog and Sörbom, 1982; see also Lavee, McCubbin, and Olson, 1987, footnote 3, for technical specifications; see Reilly, Entwistle, and Doering, 1987, for an example of a model where all the latent variables were specified with single indicators with error).

**Step 3. Assess the identification status of the model.**

After the model has been specified and prior to its estimation, the identification status of the model needs to be evaluated. Generally speaking, identification of a model's parameters has to do with the question of whether there is a unique set of parameter values consistent with the data. Under a given model structure, with certain specifications, some parameters may be uniquely estimated, others may not. If all the parameters of a model can be uniquely estimated, the whole model is identified; otherwise, the model is unidentified (Jöreskog and Sörbom, 1984).

An initial requirement for model identification is that there are at least as many, or more, known data points (variances and covariances, or correlations between measured variables) than unknown parameters. Second, the model's specifications should meet certain necessary and minimum conditions, namely, the rank and the order conditions (see Berry, 1984; Kenny, 1979; Long, 1983).

Although it may be difficult (and, in some cases, practically impossible) to assess in advance the identification status of all the parameters,
especially in complex models, some of the more common sources of unidentification can—and should—be examined before the model is estimated. An identification problem may be encountered because of misspecification of the measurement model, as well as misspecification of the structural model. In models involving multiple indicators of latent variables, for example, a failure to define the scale of a latent variable (by fixing the loading of one indicator or by fixing the variance of an exogenous latent variable) will cause an unidentification. Similarly, a model may not be identified if a covariation is allowed between measurement errors of single-indicator latent variables.

Unidentification due to structural specifications is most often encountered in nonrecursive models. Specifically, a necessary condition for identification of nonrecursive models requires that each endogenous variable in the nonrecursive system be directly influenced by an independent (exogenous) variable that is not predictive of the other endogenous variable. This condition can easily be assessed in advance, and when not met the researcher is required to impose some restrictions (preferably on a theoretical basis) to render the model testable.

MODEL ESTIMATION

Step 4. Collect the data.

Whereas sampling design is beyond the scope of this article, two practical issues related to data collection need to be raised: first, the timing of data collection in the process of model testing; and second, the size of the sample. For a discussion of additional issues of sampling adequacy, see Bentler and Chou (1987).

Ideally, the data should be collected after the model has been specified. Only then can one assure that all the relevant variables have been measured. In practice, the majority of family researchers have used extant data bases and some have found that relevant indicators for latent variables were not existent in the data, or even worse—that no measures were available for relevant constructs. The resulting models may be specified inadequately, and constructs may be indicated by only single fallible measures or by whatever “proxy” measures were available.

Researchers often deal with the question of how large should the sample be. This is an important issue, because the likelihood-ratio test of the model fit is sensitive to sample size and requires a fairly large sample to be a valid test statistic (Jöreskog and Sörbom, 1984). If the sample is too small, the chi-square test may indicate that the model fits the data even if the model is theoretically meaningless. On the other hand, if the sample size is very large, even a good model may be rejected.

In family studies, researchers have reported analyzing covariance structure models with samples as small as 58 (Wright and Price, 1986) and 62 (Roosa et al., 1982) and as large as 750 (Orthner and Pittman, 1986) and even larger. Whereas Wright and Price’s was a relatively simple model, Roosa and his colleagues estimated a complex model (10 latent with 39 measurable variables). It is no wonder that they experienced extreme difficulties in fitting their model. On the other hand, Hoelter and Harper (1987) explained their model’s lack of fit by the size of their sample (340 observations).

The complexity of the model (namely, the number of variables and degrees of freedom) is of some importance for determining the size of the sample. Bentler and Chou (1987) suggest that, as a general (though oversimplified) rule of thumb, a ratio of 10:1 between the sample size and the number of free parameters-to-be-estimated may be appropriate for the solution to be trustworthy. This ratio may be somewhat lower in some cases, or it may have to be larger to yield correct model evaluation. A more generalized guideline to sample size was provided by Bearden, Sharma, and Teel (1982). On the basis of a simulation analysis, they concluded that “a researcher who wants to reduce the risk of drawing erroneous conclusions should not use samples of less than 200” (p. 429).

It should be noted that large samples have the added advantage of allowing an investigator to test and fit a model with a random split of the sample and validate it with the second half (see, for example, Lavee et al., 1987).

Step 5. Construct the variance-covariance matrix of the measured variables.

After specifying the model and collecting the data, estimation of the model begins with the set
of known relationships among the measured variables. These relationships are arranged in a variance-covariance matrix or, if the variables are standardized, by a correlation matrix. (Actually, it is enough to input a correlation matrix and a vector of the standard deviation, from which the program computes the variance-covariance matrix. The user then has the choice of whether to analyze a standardized or unstandardized matrix).

For most practical cases, it is recommended to analyze the variance-covariance matrix (unstandardized). First, the chi-square measure of goodness-of-fit requires nonstandardization. Second, similar measures across time and different populations may be compared only when measures are unstandardized. A correlation matrix may be used, however, when the measurement units have no intrinsic meaning, the data is cross-sectional, and the model is tested within a single population (Jöreskog and Sörbom, 1984; Kenny, 1979; Long, 1976).

Step 6. Set the matrices for the LISREL program. Run the program.

The LISREL manual (Jöreskog and Sörbom, 1984) gives specific instructions for setting the program cards, including numerous examples. Normally, one would estimate the measurement model together with the structural model. In the simultaneous estimation, the solution for the confirmatory factor analysis and the solution for the structural parameters are interdependent. Thus, not only are the structural relations dependent upon the measurement operations, but these measurement properties are also dependent upon the structural relations.

Some scholars (e.g., Bielby and Houser, 1981; Burt, 1976; Kohn and Schooler, 1978) maintain that the measurement model should be analyzed first and separate from the structural model. There are some convincing advantages to this approach, namely, that simultaneous estimation may lead to interpretational confounding of the latent variables because they are assigned meaning not only on the basis of their observed indicators (epistemic criteria) but also on the basis of their indicators' relations to other latent variables in the model (structural criteria). Most scholars, however, prefer the simultaneous, "full information" estimation of all the model's parameters. In family research, most investigators have followed the simultaneous estimation approach, and only a few (e.g., Schoenbach, 1985) have estimated the measurement model prior to estimating the structural model.

Step 7. Examine the measures of overall model fit and testability. Assess whether modifications are needed.

Once the model is estimated, the program issues a number of statistics to evaluate how well the model, as a whole, fits the data and whether some specifications are fundamentally wrong. These indicators would be examined before one evaluates specific parameter coefficients (e.g., path coefficients).

First, it is necessary to examine certain estimates to assure that the model is testable. If the model is not identified, the program alerts the user to this fact and does not provide certain statistics (namely, standard errors and t values). Other indicators of major problems in the model or the data are a covariance matrix that is not positive definite, negative variances, correlations that are larger than one in magnitude, or extremely large standard errors.

Next, one examines the measures of overall model fit. The most frequently used measure is the likelihood-ratio chi-square statistic. Contrary to most applications of statistical significance tests, a statistically significant chi-square indicates that the discrepancy between the data (variance-covariance matrix) and the model (variance-covariance matrix implied from the maximum-likelihood parameter estimates) is greater than expected by chance. Conversely, a chi-square measure that is statistically insignificant indicates a good fit of the model to data.

If the model is testable but does not fit the data sufficiently well, the modification indices provide a practically useful means for assessing what changes in the model's specification would improve its fit to data. Specifically, a modification index larger than 5.0, in either the measurement or the structural model, indicates that the model's fit to the data will improve significantly if the respective path is allowed (that is, if the constraint of fixed parameter is relaxed).

Modification and Reestimation

Step 8. If modifications are needed, assess the
theoretical and/or substantive meaning of the new specifications.

The practice of modifying and reestimating a model is quite common in the analysis of covariance structures. Some modifications may be required if the model was unidentified, in order to make it testable; other modifications may be suggested by the diagnostic information of the initial analysis, in order to achieve a model that better fits the data. The decision regarding whether and what modifications to make is a critical one in the process of model testing, because the diagnostic information is based on mathematical and statistical criteria only.

Identifying restrictions. If the model was unidentified, certain restrictions need to be imposed. In the social sciences, the most commonly used strategy is zero-restriction (Berry, 1984; Long, 1983), that is, to constrain certain structural elements to zero. Alternatively, certain structural parameters may be constrained to be equal to one another, or a parameter may be fixed to a certain reasonable value (Kenny, 1979).

The various options for rendering a model identified should be evaluated in terms of their theoretical meaning, and researchers are strongly advised against the practice of specifying atheoretical restrictions to identify the structural part of a model. In practice, however, one may find that, given one's set of measures, no identifying constraint is more plausible than others, or worse—that the possible identifying assumptions cannot readily be supported by existing knowledge. Instead, it may be necessary to impose ad hoc assumptions (for example, to assume that an endogenous variable in a nonrecursive system is not related to an exogenous variable). Of course, the estimates of such a model should be interpreted with caution, because the identifying assumptions may affect the parameter estimates of the model in an unknown way.

To examine the extent to which the parameter estimates of an arbitrarily identified model depend upon the particular identifying assumptions, it is recommended that a sensitivity analysis be conducted (Land and Felson, 1981). Most frequently this is done by assigning different values to the identifying parameters in a series of repeated analyses and observing their effect on other parameter estimates. The change in the numerical values of the model parameters indicates the degree to which these parameters are sensitive to the identifying assumptions, hence the confidence one may have in the model's estimates (see Lavee, 1987, for an illustration of sensitivity analysis in family research).

Adding extra parameters. Whereas adding restrictions may be necessary to make a model testable, a greater dilemma for researchers is posed by the modification indices. These indices suggest ways to improve the model, but they are based on statistical grounds only (one can even improve one's model by automatic modifications, which will yield best fit with no theoretical considerations).

It is important to remember that post hoc specifications violate a fundamental assumption underlying statistical theory, namely, that hypotheses are formulated prior to the analysis of the data. They do, however, have important exploratory value, provided that they are not introduced carelessly for the mere purpose of achieving a better model fit. To protect against this kind of "specification abuse," it is strongly recommended to make only those modifications that make theoretical or substantive sense and to ensure that they do not affect the important parameters of interest (Bentler, 1980; Jöreskog and Sörbom, 1984; Wheaton, 1987).

Generally, family scientists have shown careful consideration in either choosing to make post hoc modifications (e.g., Acock and Yang, 1984), choosing not to make them because of lack of justification (e.g., Roosa et al., 1982), or choosing to relax some assumptions but not others (e.g., Orthner and Pittman, 1986).

Hypothesis Testing

Step 9. Evaluate your model and its parameter coefficients against your hypotheses.

How does a researcher know that his or her hypotheses were confirmed? In most statistical analyses, the answer is quite straightforward and well known: if an estimated value (e.g., a regression coefficient) is larger than would be expected by chance, the null hypothesis is rejected and the research hypothesis is said to be "confirmed."

In analysis of covariance structures, however, there is no simple and easily interpretable criterion for assessing "success" (Biddle and Marlin,
1987). Generally speaking, two sets of criteria are being used: goodness of the model as a whole, and statistical significance of specific parameters.

**Model fit.** The chi-square of model's fit to the data is most frequently cited as an indicator of success. This, however, may not be an adequate, certainly not a sufficient, criterion that the research hypotheses have been confirmed. This statistic simply shows (provided that all the statistical assumptions were met) that the model's specifications describe the structure of relationships among the observed variables against the alternative hypothesis that these relationships are of no specific structure (that is, they are random).

Additionally, as has been noted above, the chi-square statistic assumes multivariate normality (which is difficult to assure) and is sensitive to sample size. A model is more likely to be "fit" when a small sample is used than a model assessed with data from a large sample. While the ratio of chi-square to degrees-of-freedom often is cited by researchers as an alternative indicator of fit, this measure suffers from the same limitations as those of the chi-square, and there is no accepted criterion of an adequate $\chi^2/df$ ratio. Another approach for using the chi-square as a measure of fit, which adjusts for sample size, has been proposed by Hoelter (1983). As Wheaton (1987) has shown, however, Hoelter's Critical-$N$ value of acceptable fit may not be adequate in small samples.

Because of the limitations of the chi-square measure, Jöreskog and Sörbom (1984) have suggested that rather than using this measure (and its associated significance level) as a criterion of the goodness of the model, the chi-square can be used in comparative way, where a model's fit (chi-square relative to its degree of freedom) is assessed against the fit of a hierarchically nested model. Researchers often use this procedure to test the significance of the change when a model is modified or when hierarchically nested models are posited prior to the model estimation and their relative fit is tested (see, for example, Honeycutt, 1986; Mirowski and Ross, 1987; Thomson and Williams, 1984).

In assessing model fit, two additional measures often are used: the Goodness of Fit Index (GFI), and the Root Mean Square Residual (RMSR). The GFI measure is not affected by sample size and is robust against departure from normality. It may range, theoretically, from 0 to 1.0, but the exact statistical distribution of this statistic is not known. Some researchers (e.g., Hoelter and Harper, 1987) have suggested that a GFI greater than .90 indicates a good fit. Experience shows, however, that even models of relatively bad fit by other indicators may have a GFI larger than .90, and vice versa. A small difference between the GFI and the Adjusted (for degrees of freedom) GFI may also indicate that the model fits well, but no criterion of how small a difference is small enough is available.

The RMSR is a measure of the mean discrepancy between the data and the implied (reproduced) variances and covariances. The lower the index, the better the fit of the model to the data. The RMSR, however, is a valuable index only when the mean data variance-covariance is known; it is harder to evaluate when an unstandardized variance-covariance matrix is being used.

Various other statistical approaches for assessing model fit have been developed to address the limitations of the chi-square statistic, the effect of sample size, and the fit of a revised model compared to a baseline theoretically derived model (see Wheaton, 1987, for a review). However, there is no single acceptable criterion for judging the overall goodness of a model; multiple measures are always needed.

**Path coefficients.** The model being tested constitutes multiple hypotheses regarding causal relationships between constructs, as well as the relationships between constructs and their measured indicators. These hypotheses are tested by assessing the statistical significance of each parameter. Specifically, the statistical significance of each parameter is determined by a $t$ statistic, which is equal to the ratio of the coefficient and its standard error. Coefficients that are twice as large as their respective standard errors (that is, $t \geq 2.0$) are considered statistically significant.

In addition to testing structural and measurement hypotheses, a LISREL model allows for comparison of the magnitude of path coefficients. Different parameters in a model may be compared (and the difference tested) within a population, or similar parameters in a model may be compared between populations (e.g., comparing effect size
in a model estimated with husbands’ and wives’ data).

Hypothesis testing in analysis-of-covariance-structure models, as the above discussion indicates, involves both evaluation of the fit of the overall model and specific hypotheses regarding causal relations between constructs. No single measure of “success” is enough. Furthermore, the measures of overall model’s fit, which are so often emphasized in covariance structure research, can not be the only, not even the primary criterion for testing the multiple hypotheses embedded in a model (Biddle and Marlin, 1987; Wheaton, 1987).

**WRITING THE RESEARCH REPORT**

The final step in any research is reporting it in a way that other family scientists can follow its logic, procedures, analysis, and interpretation. We have become accustomed to certain standards in reporting research, including standard ways of presenting results. In reporting analysis of covariance structures, this has become especially important because the method is complex and because most readers are not yet accustomed to its standards and notations.

Certain conventional notations are often used in covariance structure model diagrams (see Bentler, 1980; Jöreskog and Sörbom, 1984). Latent variables are typically denoted by circles, and measured variables by rectangles; causal relations are shown by unidirectional arrows (both between latent variables and between each latent variable and its indicators), and noncausal relationships, either between exogenous variables or between disturbances in equation or in measures, are shown with a curved double-sided arrow. Unfortunately, these standards are not always maintained in research reports. Model diagrams are presented in a variety of ways that are sometimes hard to follow.

Additionally, certain pieces of information are needed for readers to be able to evaluate covariance structures research. For example, it is expected that a correlation matrix and the means and standard deviations of all the measured variables will be presented. Research reports should also provide information (in a table format or in their path diagram) of all the parameter estimates—measurement as well as structural—including error terms and, when applicable, covariance between errors. It is useful to include both the standardized solution and the unstandardized estimates together with their standard errors. Data regarding the model’s fit should include all relevant information (not just the chi-square statistic and its associated significance level), and when hierarchically nested models are tested, it is useful to provide a table that summarizes each model’s behavior. Finally, it is essential that all the modifications and added restrictions that were imposed to identify a model or to improve its fit be reported. In a review of research reports, I found that while some investigators have presented results of their analyses with great detail, others have excluded essential information from their report, such as a correlation matrix, estimates of the measurement model, or their identifying assumptions.

**FAMILY THEORY AND COVARIANCE STRUCTURE MODELS: PROBLEMS AND PROSPECTS**

Covariance structure analysis is a powerful tool for testing theoretical models. The combination of the measurement model with structural equation modeling enables one to test hypothesized structural relationships between “true” variables that may capture theoretical constructs better than single indicators. The treatment of correlated errors, both in measurement and in equations, and the capability of testing reciprocal causal relationships allow researchers to test more realistic models than could have been achieved by previous approaches to causal modeling.

The role of theory at every stage of the analysis has been repeatedly emphasized. Covariance structure modeling is primarily a confirmatory method, developed for testing hypotheses rather than generating them. Nonetheless, the LISREL approach includes powerful diagnostic measures that may have important serendipitous utility and may prove to be useful for reworking and revising theoretical models. As Bentler (1980: 421) has put it, “even though the goal of causal modeling is explanation rather than description, an appropriate interplay between theory and data surely involves exploration as well as confirmation.”

Review of covariance structure analysis in family studies reveals some problems as well as prospects in the interplay between theory and method. The problems come primarily from the level of development of family theories.
Specifically, theories are often too crude to permit a clear and unambiguous specification of causal relationships between constructs. Structural relationships are often described vaguely, suggesting that multiple variables affect one another. Although the LISREL model permits testing of reciprocal relationships, models that are too complex (that is, with multiple reciprocal relationships) may prove untestable (Kenny, 1979). Additionally, when identifying assumptions are needed to identify a model, they are often difficult to justify with the current state of specificity of our theories.

The problem may be even greater when it comes to the specification of the measurement model. Many of the constructs studied by family researchers are not well defined, and the gap between theory and measurement (Schumm, 1982) leads researchers to use whatever measures seem, on an intuitive basis, to be adequate proxies to their constructs (or whatever measures exist in the extant data they are using).

These theory-to-model translation difficulties bear with them some promises, however. For one, as researchers struggle with the adequate representation of theory in structural relations and in measurement operations, the weaker points of a theory surface. In the long run, I believe, this may prove to be an important phase in the process of theoretical and conceptual clarification (Lavee and McCubbin, 1985).

The second source of promise is the indicators of model misspecification. The modification indices, for example, may provide important clues for theoretical revisions. A large modification index in the measurement model may suggest ways for conceptual clarification, while a modification index in the structural model may suggest that a structural relationship between constructs exists that was not hypothesized by theory.

Of course, mucking around with the data and introducing modifications to achieve the best fit is not an appropriate way to test a theory. But as Wheaton (1987) has noted, model modification permits one to know what further models look like. If researchers carefully examine their LISREL output and take note of possible theoretical implications, the interplay between theory and method, and between confirmation and exploration, will prove fruitful even if the model does not fit.

NOTES

Work on this article was supported in part by Grant No. DA 02360 from the National Institute on Drug Abuse while the author was at the Department of Family Social Science, University of Minnesota. I am indebted to Deborah Godwin and David Johnson for their helpful comments on an earlier draft.

1. While LISREL has become a commonly used term for the approach discussed here, similar statistical analysis packages also are available, such as Bentler’s (1985) EQS.

2. In recent years, several excellent reviews of covariance structure models (particularly of the LISREL model) have been published in various professional journals and textbooks. Readers who are interested in a more complete and detailed account of the method are encouraged to refer to these reviews. I found the following particularly useful: Carmine (1986), Kerlinger (1986, chap. 36), Pedhazur (1982, chap. 15), Bentler (1980), Lomax (1982), Jöreskog and Sörbom (1982), Long (1983), and Cappella (1980), as well as the LISREL manual (Jöreskog and Sörbom, 1984) and Bentler’s (1985) EQS program manual. Additionally, Maruyama and McGarvey’s (1980) paper is a very clear presentation of LISREL application.

3. A special case of the operationalization of constructs as latent variables, somewhat unique to family research, has been the specification of a common family factor. The LISREL approach enables one to specify measures taken from individual family members as indicators of a latent variable. This “family” variable then underlies the individual measures. This approach, suggested first by Thomson and Williams (1982), seems to be a promising way for measuring family constructs (Schumm, Barnes, Bollman, Jurich, and Milliken, 1985).

4. Additionally, a moment matrix may be analyzed for some specific applications. See Jöreskog and Sörbom, 1984, chapter 5.

REFERENCES


Schumm, Walter R. 1982. “Integrating theory, measurement, and data analysis in family studies survey