CHAPTER **1**

An Introduction to Multivariate Design

The Univariate and Bivariate Domain

This book is about multivariate designs. Such designs as a class can be distinguished from the univariate and bivariate designs with which readers are likely already familiar. Here is an example of a univariate design. Assume that we designed an experimental study with a single independent variable and one dependent variable. For example, perhaps we wished to study the effect of room color (white vs. light blue) on study effectiveness as measured by the number of correct responses made to a multiple-choice test administered on a computer. This is considered to be a univariate design because, to oversimplify the situation for the moment, there is only one dependent variable.

Next, consider a simple correlation design with only two measures. Sometimes these variables are referred to as the *predictor variable* and the *outcome variable* because there is no experimental treatment under the control of the researchers. However, an equally viable argument can be made to dispense with such labels altogether in correlation designs, simply calling them dependent or measured variables and referring to one as the *X* variable and referring to the other as the *Y* variable (Keppel, Saufley, & Tokunaga, 1992, p. 460).

Here is an example of a bivariate design. By administering standardized paper-and-pencil inventories to a sample of individuals, we can quantitatively assess their current levels of self-esteem and depression. The data can then be analyzed using a Pearson product-moment correlation statistical

procedure. This simplified example represents a bivariate analysis because the design consists of exactly two dependent or measured variables.

The Tricky Definition of the Multivariate Domain

Some Alternative Definitions of the Multivariate Domain

To be considered a multivariate research design, the study must have more variables than are contained in either a univariate or bivariate design. Furthermore, some subset of these variables must be analyzed together (combined in some manner). For example, let's revisit our hypothetical study of the effect of room color on study effectiveness. As we described it a moment ago, the dependent variable was the number of correct responses on a multiple-choice test. For the present illustration, let's assume that we had structured the situation with an equal emphasis placed on correct responses and speed of responding. For each participant, we would obtain two measures, number of correct responses and speed of responding. Without worrying about the details of how we would do this, it is possible to consider combining these two measures into a single composite measure that might be interpreted to reflect performance efficiency. This combining of variables (two dependent variables in this case) into a composite would bring us into the realm of a multivariate design.

To qualify for the label *multivariate design*, variables must be combined together. But which kind of variables are to be combined is far from agreed on. Some authors, such as Stevens (2002), count only dependent variables, as in our example, suggesting that a multivariate study must combine together "several" (i.e., more than two) dependent variables. But such a definition puts Stevens in the awkward situation of excluding certain designs that most authors, himself included, would incorporate in a multivariate statistics book. In his own words,

> Because we define a multivariate study as one with several dependent variables, multiple regression (where there is only one dependent variable) and principal components analysis would not be considered multivariate techniques. However, our distinction is more semantic than substantive. Therefore, because regression and components analysis are so important and frequently used in social science research, we include them in this text. (p. 2)

Other researchers, such as Tabachnick and Fidell (2001b), opt to be rather inclusive in their definition of multivariate designs. But their inclusiveness can

also get them into some difficulty. To qualify as a multivariate design, Tabachnick and Fidell require that more than one of each type of variable must be combined: "With multivariate statistics, you simultaneously analyze multiple dependent and multiple independent variables. This capability is important in both nonexperimental . . . and experimental research" (p. 2).

The problem with this definition is analogous to what Stevens (2002) faced: Multiple regression, which has only one dependent variable, and principal components analysis, where the multiple variables are traditionally not thought of as dependent variables, appear to be excluded from this definition. Because Tabachnick and Fidell (2001b) argue that you need to simultaneously analyze multiple dependent and independent variables, they, like Stevens, would ordinarily omit certain multivariate topics from their book. But analogous to Stevens's strategy, Tabachnick and Fidell do not let their definition prevent them from covering topics that are ordinarily treated in multivariate texts.

Our Characterization of the Multivariate Domain

Following the lead of Hair, Anderson, Tatham, and Black (1998) and Grimm and Yarnold (2000), we believe that a good way to think about multivariate research is to maintain that the analysis involves combining together variables to form a composite variable. The most common way to combine variables is by forming a linear composite where each variable is weighted in a manner determined by the analysis.

The general form of such a weighted composite is in the form of an equation or function. Each variable in the composite is symbolized by the letter X with subscripts used to differentiate one variable from another. A weight is assigned to each variable by multiplying the variable by this value; this weight is referred to as a *coefficient* in many multivariate applications. Thus, in the expression w_2X_2, the term w_2 is the weight that X_2 is assigned in the weighted composite; w_2 is called the coefficient associated with X_2. A weighted composite of three variables would take this general form:

$$\text{weighted composite} = w_1X_1 + w_2X_2 + w_3X_3$$

These weighted composites are given a variety of names, including *variates*, *composite variables,* and *synthetic variables* (Grimm & Yarnold, 2000). Variates are therefore not directly measured by the researchers in the process of data collection but are created or computed as part of or as the result of the multivariate data analysis. We will have quite a bit to say about composite variables (variates) throughout this book.

Variates may be composites of either independent or dependent variables, or they may be composites of variables playing neither role in the analysis. Examples where the analysis creates a variate composed of independent variables are multiple regression and logistic regression designs. In these designs, two or more independent variables are combined together to predict the value of a dependent variable. For example, the number of delinquent acts performed by teenagers might be found to be predictable from the number of hours per week they play violent video games, the number of hours per week they spend doing homework (this would be negatively weighted because more homework time would presumably predict fewer delinquent acts), and the number of hours per week they spend with other teens who have committed at least one delinquent act in the past year.

Multivariate analyses can also create composites of dependent variables. The classic example of this is multivariate analysis of variance. This general type of design can contain one or more independent variables, but there must be at least two dependent variables in the analysis. These dependent variables are combined together into a composite, and an analysis of variance is performed on this computed variate as in the case of combining number of correct responses and speed of responding in the study of room color mentioned above. The statistical significance of group differences on this variate (performance efficiency in this example) is then tested by a multivariate F statistic (in contrast to the univariate F ratio that you have studied in prior coursework).

Sometimes variables do not need to play an explicit role of being either independent or dependent and yet will be absorbed into a weighted linear composite in the statistical analysis. This occurs in principal components and in factor analysis where we attempt to identify which variables (e.g., items on an inventory) are associated with a particular underlying dimension, component, or factor. These components or factors are linear composites of the variables in the analysis.

The Importance of Multivariate Designs

The importance of multivariate designs is becoming increasingly well recognized. It also appears that the judged utility of these designs seems to be growing as well. Here are some of the advantages of multivariate research designs over univariate research designs as argued by Stevens (2002):

1. Any worthwhile treatment will affect the subjects in more than one way. . . .

2. Through the use of multiple criterion measures we can obtain a more complete and detailed description of the phenomenon under investigation. . . .

3. Treatments can be expensive to implement, while the cost of obtaining data on several dependent variables is relatively small, and maximizes information gain. (p. 2)

A similar argument is made by Harris (2001):

However, for very excellent reasons, researchers in all of the sciences—behavioral, biological, or physical—have long since abandoned sole reliance on the classic univariate design. It has become abundantly clear that a given experimental manipulation . . . will affect many somewhat different but partially correlated aspects of the organism's behavior. Similarly, many different pieces of information about an applicant . . . may be of value in predicting his or her . . . [behavior], and it is necessary to consider how to combine all of these pieces of information into a single "best" prediction. (p. 11)

In summary, there is general consensus about the value of multivariate designs for two very general reasons. First, we all seem to agree that individuals generate many behaviors and respond in many different although related ways to the situations they encounter in their lives. Univariate analyses are, by definition, able to address this level of complexity in only a piecemeal fashion because they can examine only one aspect at a time.

In the simple univariate experiment, we described earlier testing the effect of room color on learning, for example, the dependent measure was exam score. But how fast individuals responded to the questions shown on the computer screen, how many questions they answered correctly, and (to add still another dependent variable) even how confident they were in their answers to those questions could also have been evaluated. This information might have contributed to a more complete understanding of the learning experience of those individuals.

All three of these measures would most likely be correlated with each other to a certain degree and all three would most likely tap into somewhat different but related aspects of the participants' responding. Together, they may have provided a more complete picture of the learners' behavior than any one of them in isolation. This study, originally structured as a univariate design, could thus be transformed into a multivariate study with the addition of other dependent variables.

The second reason why the field appears to have reached consensus on the importance of multivariate design is that we hold the causes of behavior to be complex and multivariate. Thus, predicting behavior is best done with more rather than less information. Most of us believe that several reasons explain why we feel or act as we do. For example, the degree to which we

strive to achieve a particular goal, the amount of empathy we exhibit in our relationships, and the likelihood of following a medical regime may depend on a host of factors rather than just a single predictor variable. Only when we take into account a set of relevant variables—that is, when we take a multivariate approach—have we any realistic hope of reasonably accurately predicting the level—or understanding the nature—of a given construct. This, again, is the realm of multivariate design.

The General Organization of the Book

The domain of multivariate research design is quite large, and selecting which topics to include and which to omit is a difficult task for authors. Our choices are shown below. To facilitate presenting this material, we used two organizational tactics. First, we grouped the sets of chapters together based on the nature of the variate—the composite variable—computed in the process of performing the multivariate analysis. Second, we generated introductory univariate or bivariate chapters to lead off the first three chapter groups. This was done partly to serve as a refresher to readers and partly to serve as a way of framing certain concepts treated in the multivariate chapters in that group. We end this chapter with a more detailed description of the various parts of this book.

Part I: Foundations

The chapters in this part of the book introduce readers to the foundations or cornerstones of designing research and analyzing data. Our first chapter—the one that you are reading—discusses the idea of multivariate design and addresses the structure of this book. The second chapter on fundamental research concepts covers both some basics that you have learned about in prior courses and possibly some new concepts and terms that will be explicated in much greater detail throughout this book. The following chapters on data screening are applicable to all the procedures we cover later and so are placed as separate chapters in this beginning part. They cover ways to correct data entry mistakes, how to evaluate assumptions underlying the data analysis, and how to handle missing data and outliers.

Part II: The Independent Variable Variate

Some multivariate research designs form composites of independent variables. These designs typically have to do with predicting a value of a dependent variable. An initial chapter on bivariate correlation and simple

linear regression leads this group. This chapter is included to provide readers with an opportunity to review material that they have probably covered in previous coursework so that they have a solid foundation for the multivariate chapters that follow. Multiple regression uses quantitative variables as both predictors and as the variable being predicted (the criterion variable), whereas logistic regression can accommodate categorical variables in these roles. Finally, discriminant analysis uses quantitative variables to predict membership in groups specified in the data file. Although one can use logistic regression for this purpose, we cover discriminant analysis here because it is one of the "classic" multivariate methods.

Multiple Regression Analysis

Multiple regression analysis is used to predict a quantitatively measured variable, called the *criterion* or *dependent* variable, by using a set of either quantitatively or dichotomously measured predictor or independent variables. It is an extension of simple linear regression where only one predictor and one criterion variable are involved. Each independent variable in the set is weighted with respect to the other predictors to form a linear composite or variate that maximizes the prediction of the criterion variable. The computed value of the variate is equal to the predicted value of the dependent variable, and the weighted composite can be thought of as a specification of the prediction model for the criterion variable.

The multiple regression procedure can be employed when we can formulate the research problem in terms of predicting a quantitatively measured variable. We might use multiple regression analysis, for example, to predict the degree of success that students experience in the first year of college. Success here is the dependent variable and might be assessed by faculty ratings, grade point average, or some other quantitative measure. Predictors, or independent variables, might include high school grade point average, scores on a standardized college entrance exam, and even some attitude measures or personality characteristics that might have been assessed just prior to the start of the academic year.

Logistic Regression Analysis

Logistic regression is conceptually similar to multiple regression in that we use a set of independent variables in combination to predict the value of a dependent variable. In logistic regression, the variable being predicted is measured on a qualitative or categorical scale; in the majority of instances, this dependent variable is dichotomous; that is, it consists of two possible values. The predictors can comprise any combination of categorical and

quantitative variables. Logistic regression is more flexible than multiple regression in that it must conform to fewer statistical assumptions to be appropriately used.

One of the strengths of logistic regression is that the model it produces is not linear but instead is sigmoidal (S-shaped). This multivariate procedure therefore permits the predictors to be related to the dependent variable in a nonlinear manner. In the dichotomous dependent variable situation, the result of the procedure—what is being predicted—is the probability of the cases falling into one of the dependent variable's categories. The outcome is often expressed as an odds ratio where we may say that the odds of a case being in one category were, for example, 5.25 times greater than the chances its being in another category.

Logistic regression can be used any time we are interested in identifying the variables associated with being in one condition over another. Such conditions, which are candidates to serve as dependent variables, are created by the individuals themselves, may be imposed by the society or culture, and could be based on a personality characteristic of the individuals or whatever. Examples of such variables include students who major in the arts versus those who major in the sciences, candidates who passed versus those who failed a state license examination, and individuals who were and were not at risk for a certain disease. Predictor variables would be chosen according to the research problem and, presumably, based on the theoretical models available at the time as well as the empirical research literature.

Discriminant Analysis

Discriminant analysis (sometimes called discriminant function analysis or multiple discriminant analysis) is a technique designed to predict group membership from a set of quantitatively measured variables. There are two types of discriminant function analysis—predictive and descriptive. In predictive discriminant analysis, the goal is to formulate a rule or model that we use to predict group membership (Huberty, 1994). The other type is descriptive discriminant analysis where the focus is the interpretation of the linear combination(s) of the independent variables to describe the differences between the groups. As we shall see later in the book, this descriptive application of discriminant function analysis is often used as a follow-up analysis to significant multivariate analysis of variance.

Discriminant analysis is similar to logistic regression in that the dependent variable is measured on a categorical scale. This categorical variable ordinarily represents groups of participants in the study. The goal of discriminant analysis is to predict the group membership of the cases using a weighted linear composite of the predictor variables. Discriminant analysis

is also intimately related to multiple regression. If the same dichotomous variable is used as the criterion variable in multiple regression and as the "groups" variable in discriminant, and if the same variables are used as the predictors in both analyses, then discriminant and multiple regression analysis yield weights for the predictors that are in the same proportion to each other, thus producing exactly comparable models.

An example of predictive discriminant analysis can be found in the research of a student of one of the authors. The student was working as an intern in a facility that housed male juvenile offenders while they were extensively evaluated by a team to recommend to the court if the wards were or were not amenable to (were likely to improve as a result of undergoing) a therapeutic treatment program. The team, consisting of a cadre of health and social system professionals (e.g., psychologist, psychiatrist, social worker, case worker) administered a battery of tests, conducted numerous interviews, analyzed the history and nature of the ward's criminal behavior, and so on over a 90-day period. Files as thick as 5 or 6 inches were commonly built. At the end of the process, the team would engage in a case conference during which all the information was discussed and a "yea" or "nay" decision on treatment amenability was made.

Our student in the above example believed that there was little likelihood that all the information present in these thick files was of equal value in contributing to the amenability decision. Recognizing a research project when she saw one, the student realized that we could obtain all the information in the files for the past few years, including the decision regarding amenability, and process this data set through predictive discriminant analysis. The dependent variable that she used was the dichotomous amenability decision ("yea" or "nay"), and the predictor variables were all those in the files (intelligence score, MMPI scores, and scores on a host of variables each professional thought that it was important to assess). When the dust of the analysis settled, we identified about half a dozen variables that pretty much correctly classified the wards as well as the objective data permitted. We therefore offered the rather unpopular suggestion to the agency that it could conduct the objective part of its assessment in a couple of days rather than a couple of months. Even though the agency did not implement our suggestion, it did provide us with an interesting example of an application of discriminant analysis.

Part III: The Dependent Variable Variate

The prototypical procedure in which one forms a composite of dependent variables is multivariate analysis of variance (MANOVA). We lead off with a chapter devoted to the univariate domain of comparing means and

its companion SPSS chapter. This is followed by three pairs of MANOVA chapters covering the two-group case, the case of three or more groups, and the two-way factorial design. All the latter chapters are based on between-subjects designs.

Multivariate analysis of variance (MANOVA) is an extension of the analysis of variance (ANOVA) procedure with which you are already familiar. ANOVA can involve any number of independent variables, although researchers do not ordinarily use more than three or maybe four. No matter how many independent variables are incorporated, however, an ANOVA design contains exactly one dependent variable. For this reason, ANOVAs are classified as univariate designs.

MANOVA designs differ from ANOVA designs in terms of the number of dependent variables involved in the analysis. That is, it is not unusual in many research projects to collect data on more than one dependent variable. Through the ANOVA procedure, each of these dependent variables would be placed in a separate analysis. This is not usually an appropriate strategy to use because in many cases these dependent variables are correlated. Under these circumstances, the researchers are not really performing independent analyses. In an exaggerated sense, they are repeatedly analyzing the same data under different dependent variable labels.

Using the MANOVA procedure in the above context is considered by most researchers as the more appropriate approach to the data analysis. MANOVA will form a linear composite or variate of the dependent variables, weighting them to best differentiate the groups (as represented by the independent variables). This portion of the MANOVA analysis is a descriptive discriminant analysis akin to what we have already described (although, here, the variables making up the variate play the role of dependent variables). Each case is then assigned a value on the variate, and an analysis is run using the variate as the dependent variable. The result of this analysis is a multivariate F ratio. If it is significant, separate univariate ANOVAs are computed for each of the separate dependent variables, but the error term used to evaluate the univariate F ratios is based on the multivariate analysis.

MANOVAs can be used in most situations where we have collected data on multiple dependent variables, especially if these variables are known or are believed to be related to each other. Consider the situation where an inventory containing numerous subscales has been administered to different samples of individuals in an effort to determine if they differ on these measures. The Minnesota Multiphasic Personality Inventory-2 (MMPI-2), for example, which can be thought of as a way to detect the possibility of psychopathology, contains 10 clinical scales and numerous specialized scales. As another example, the California Personality Inventory (CPI), designed to characterize normal personality, contains about a dozen and a

half personality scales. Rather than performing 10 separate ANOVAs on the MMPI-2 scales or 17 separate analyses on the CPI, or 27 separate analyses if both inventories were used in the same study, it is not only more appropriate but also substantially more efficient to run a single MANOVA on the entire set of dependent variables.

Part IV: The Emergent Variate

The chapters in this group are all concerned with, informally speaking, factor analysis. We say "informally" because in everyday parlance, the term *factor analysis* is used by many nonstatisticians in a generic manner. Actually, it is worth distinguishing between principal components analysis and factor analysis, which we do in the first and second chapter pairs of this section.

Principle components analysis and factor analysis are often thought of as data reduction techniques. They both tap into the idea that a relatively large set of variables measured in a research study, such as items on some inventory, reflect a smaller number of underlying themes than there are items. Both techniques are used to identify the degree to which the variables are associated with these themes.

Both principle components and factor analysis fall under the auspices of exploratory approaches to the analysis of the data. Confirmatory factor analysis takes a different approach, and this is covered in the second chapter pair of this group. Exploratory factor analysis is run through SPSS, but confirmatory analysis, because it is an application of structural equation modeling, must be performed in an alternative program. We have illustrated this and the other structural equation modeling techniques covered in the next group of chapters by using the AMOS program, but other programs (LISREL, EQS, and even SAS) can run these analyses as well.

Exploratory Factor Analysis

Factor analysis begins the chapters devoted to structural analysis. Exploratory factor analysis attempts to summarize or identify the few themes or dimensions that underlie a relatively larger set of variables. These underlying dimensions are the factors that emerge from the analysis, thus revealing the structure of the set of variables. Each factor is a weighted linear composite on which all the variables in the analysis are represented. Different factors are distinguished based on the different patterns of weights assigned to the variables. The number of factors in the solution that is ultimately accepted depends, within certain limits, on the educated judgment of the researchers. More factors will account for a greater percentage of the variance and thus lose less information in the summarization process, but fewer factors will

often offer a more efficient and reasonable interpretation. This tension makes for an interesting dynamic in the interpretation of factor analytic results.

We find exploratory factor analysis extensively used in test development. When researchers perform a content analysis of the domain they are targeting for a test or inventory, they will usually generate a relatively large number of items in an attempt to represent it on paper. But if we generate, say, even 60 items, common sense tells us that we are probably not tapping into 60 separate and independent features of the domain. Rather, it is almost certainly true that one subset of items might be more concerned about one particular dimension; another subset of items might be more focused on another dimension, and so forth. Factor analysis can recognize such differential emphases by examining the relationships between the responses given to the items. These differential relationships are synthesized as factors at the end of the analysis.

Confirmatory Factor Analysis

Exploratory factor analysis is analogous to an inductive approach in that researchers employ a bottom-up strategy by developing a conclusion from specific observations. That is, they determine the interpretation of the factor by examining the variate that emerged from the analysis. Confirmatory factor analysis represents a deductive approach in that researchers are predicting an outcome from a theoretical framework, a strategy that can be thought of as a top-down approach.

Confirmatory factor analysis seeks to determine if the number of factors and their respective measured variables support what is expected from the theoretical framework—that is, if the proposed model fits the data. The measured variables (also known as *indicator* or *observed* variables) are selected on the basis of prior theory, and confirmatory factor analysis is employed to determine if the variables correlate with their respective factor(s). These factors are latent (unobservable) constructs known to us through their indicators or observable variables.

In confirmatory factor analysis, the researchers presumably already know the construct represented by the factor at the start of the analysis (just the opposite of exploratory factor analysis). Based on this knowledge, they can posit which of the indicator variables are associated with each factor. Thus, exploratory analysis is used as a theory-generating procedure, whereas confirmatory procedures are employed as a theory-testing procedure (Stevens, 2002).

As an example of confirmatory factor analysis, consider the situation where a researcher has developed an inventory based on the premise that the construct under study can be represented by three distinguishable

aspects and that these have been built into the item creation process. Each item as it was developed was tied back to one of the three aspects of the construct that it represented. The inventory is then administered to a large sample representing a population to which the construct is applicable, and it is now time for the results to be factor analyzed. Without a theoretical framework in place, an exploratory factor analysis would be in order. But here, there is a sufficiently specified theoretical framework to permit the researchers to perform a confirmatory factor analysis. The model would specify one latent variable (factor) that was driving (causing, predicting, tied to) one set of items, a second latent variable causally related to a second set of items, and a third latent variable related to a third set of items. The analysis would then indicate how well the data fit the model.

Part V: Model Fitting

In Part V, we focus on model-fitting techniques. By means of these techniques, researchers hypothesize specific causal relationships between variables and then test them to determine how well they fit the data. If the model contains only measured variables, the model-fitting procedure is labeled as *path analysis;* if the model contains latent variables, the model-fitting procedure is called *structural equation modeling.*

Once a model is proposed (i.e., relationships between the variables have been hypothesized), a correlation-covariance matrix is created. The estimates of the relationships between the variables in the model are calculated using the maximum likelihood estimation procedure. Maximum likelihood estimation attempts to estimate the values of the parameters that would result in the highest likelihood of the actual data to the proposed model. These methods often require iterative solutions.

The first pair of chapters in this section presents causal modeling from a multiple regression, path analysis, and structural equation modeling perspective. Very often, alternative models are proposed and researchers are interested in determining which are better able to represent their data—that is, which hypothesized models are better fits to the observed data. The second and final pair of chapters in this section talks about how to compare the alternative models that have been used in confirmatory factor analysis and causal modeling.

Causal Models

Causal modeling calls for researchers to hypothesize rather specifically about the causal (or predictive) relationships that exist between the variables in their research study. If the set of variables are all measured variables,

then researchers develop a path model by drawing a path diagram and performing a path analysis. If the model contains latent constructs, then structural equation modeling is used. In both cases, the object of the analysis is to estimate the strength of the relationships between variables as they are structured or arrayed in the model. It is then possible to gauge how well that model fits the data.

Research using causal modeling begins with an arrangement of the variables in diagram form showing a model in which thus and such variables are thought to produce (cause, predict) others, which, in turn, can be hypothesized as causes of still other variables. These models are then translated into a series of linear equations that allow researchers to assess direct effects and indirect effects on a variable. An indirect effect simply means that a variable influences another variable through a mediating variable. For example, one of our students conducted research on predictors of physical exercise. Her results indicated that body self-esteem did not directly influence the amount of physical exercise in which people engaged on a weekly basis; however, body self-esteem did influence exercise attitude, which in turned influenced amount of physical exercise.

Causal inferences are depicted by arrowed lines, and the intent of the analysis is to compute the various coefficients representing each variable-to-variable relationship. Path analysis, performed when all the variables are measured variables, can be accomplished through multiple regression analyses or from a program designed to do model fitting; when latent constructs are included in the model, the analysis must be done as a structural equation model. In any case, the resulting coefficients indicate the relative strength of the path connecting two variables.

Causal modeling is not a substitute for experimental design; that is, it does not involve an attempt to establish the necessary and sufficient conditions that enable researchers to infer that changes in the independent variable caused corresponding changes in the dependent variable. In fact, these sorts of analyses are typically performed on data that have been collected through a correlation methodology in which the variables covary together in the sample with no active intervention by the researchers. Nonetheless, causal modeling allows us to examine hypothesized causal relationships and evaluate alternative models.

We can apply this general approach to a wide range of research arenas. As an example, consider the construct of life satisfaction. We can hypothesize that the degree to which people feel satisfied with their lives may be a function of the depth of friendships they have made, how much love they feel toward others, the degree to which they have met their life goals, their personal wealth, their level of spirituality, their level of education, and so

forth. These variables themselves may be able to be placed into some hypothesized causal hierarchy. For example, how much love they feel toward others may predict the depth of the friendships they have made, which, in turn, may predict higher levels of life satisfaction. Alternative causal models could be constructed and compared with each other through the path analysis procedure.

Comparing Models

Model fitting is a tool that allows researchers to build their models and then test them to determine how well they fit the data. It is also possible and in many instances very reasonable for researchers to have generated a couple of different models, either in the confirmatory factor analysis or in the causal model arena, that they wish to compare. Model-fitting techniques allow the researchers not only to see how well each model fits the data but also to compare the alternative models with each other.

For example, one of us was the principal investigator developing a new instrument to assess multicultural competence of mental health practitioners. The research team reviewed the current literature and available instruments and was confronted with the conflict of whether multicultural competence would best be described as a three- or four-factor ability. Employing confirmatory factor analysis, the three- and four-factor models were compared, and the results supported the four-factor model.

In addition, it may be useful to determine whether a model developed in one context is applicable, generalizable, or exportable to another context. The applicability of a model across contexts is referred to as *model invariance.* For example, researchers might very well be interested in examining such invariance across samples representing different populations (e.g., women and men, younger and older individuals) or in determining if a given theoretical factor structure (e.g., the five-factor model of personality) is as good a fit to data collected under Inventory A as under Inventory B. We cover the topic of comparing models in the last pair of chapters.

Recommended Readings

Aiken, L. S., West, S. G., Sechrest, L., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of PhD programs in North America. *American Psychologist, 45,* 721–734.

Grimm, L. G., & Yarnold, P. R. (2000). Introduction to multivariate statistics. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 3-21). Washington, D.C: American Psychological Association.