Iacobucci, Dawn (ed.) (2001), *Journal of Consumer Psychology's Special Issue on Methodological and Statistical Concerns of the Experimental Behavioral Researcher*, 10 (1&2), Mahwah, NJ: Lawrence Erlbaum Associates, 75-82.

Factor Analysis

V.A. DIFFERENCE BETWEEN PRINCIPAL COMPONENTS AND FACTOR ANALYSIS

What is the difference between principal components analysis and factor analysis? Which is better for which situations?

Editor: The differences between principal components analysis and principal factor analysis are vast and minimal. First, the two analytical models are premised on different theoretical grounds. Factor analysis was developed to address the very real need of measurement—in our substantive theoretical work, we posit unobservable constructs that we purport to measure through the use of multiple indicator variables. Each of the directly measurable variables is thought to be measured with some error, and some "specificity" (something that is not random error but is systematically biasing scores on that particular item), but what covaries among the set of items—the covariability that is common across the measures—we attribute to one or more common underlying factors that gave rise to the variables, and, thereby, their covariances. Recall from your basic statistics class, when you were instructed that "correlation does not necessarily imply causation," because, for example, one alternative hypothesis to X causing Y or Y causing X, is that there may be a different cause common to both X and Y—this alternative model is that of factor analysis. We observe correlations among the items on an A_{ad} scale because we believe that an underlying construct or factor of attitude toward the advertisement is influencing how the respondent scores all the individual items on the attitude scale. Furthermore, we adhere to the standard scientific principle of parsimony, when we subsequently rotate extracted factors to "simple structure" to enhance our understanding of the pattern of variables as they load on factors and the interrelations among the factors themselves. The factor model requires inferences about theory and structure, not unlike the inferences researchers must draw to make progress in, say, cognitive psychology. In contrast, in principal components, there is no measurement theory driving the model; rather, the goal is simply one of data reduction, beginning with some big number of variables and ending with some much smaller number of components. (Yes, data reduction is an important by-product of factor analysis also.)

Although there exists this theoretical gulf between the two techniques, in practice solutions for factor analysis and components analysis often appear similar. The principal components model begins with the correlation matrix among p variables, R. It is factored (broken down into components) through an eigenvalue–eigenvector decomposition, $R = V\Lambda V'$, where V contains the eigenvectors in its columns, and Λ contains the eigenvalues along its main diagonal. We select r components, fewer than p and approximate the fit to our data, $\hat{R} \cong V_r\Lambda_r V_r'$. Sometimes researchers rotate components to simple structure. In factor analysis, because we have posited common versus specific factors, we begin with the correlation matrix R, but immediately make initial estimates regarding the communalities shared among the variables. Usually the squared multiple correlation (SMC) for each variable being predicted by all the other (p – 1) variables serves as that communality estimate. These R^2s replace the 1.0s in the correlation matrix R, and we call the result, R_{SMC}. We then factor R_{SMC} using the same eigenvalue–eigenvector logic, $\hat{R}_{SMC} = V\Lambda V' \cong V_r\Lambda_r V_r'$. The matrix of factor loadings is defined as $B = V_r\Lambda_r^{1/2}$. Due to the fact that we have modified the main diagonal—the area in a correlation (or covariance) matrix that contains the information regarding variances—while we speak of modeling the variance in principal components, we speak of modeling the covariance in factor analysis. That is, for components, we are seeking to maximally account for the variance among all the p original variables, with a smaller number of r components. For factors, we have already modeled the variances per se, so we now seek to maximally account for the covariance among all the p original variables, with a smaller number r of common factors. Note, of course, that were the communalities in the factor model uniformly big—that is, all approaching 1.0—then R_{SMC} will increasingly resemble R, and the factor model solution will increasingly resemble that of the theoretically simpler components model.

Given that most of the *Journal of Consumer Psychology* readers are likely to be social scientists studying human behavior and obtaining measures from fellow humans as respondents, it would seem that a model that posits the presence of measurement error (i.e., factor analysis) is going to typically dominate one that does not (principal components). (Perhaps if one worked primarily with variables like sales figures, one might convince oneself that these variables have no measurement error.) It is not good practice to abandon theory when conducting other scientific work—why do so here?

Sometimes analysts try to argue for the use of components when the purpose of the analysis is not the components per se, but the formation of composite scales to enter into a subsequent regression analysis. However, the issue on the table is actually one of multicollinearity, which dictates not so much the choice between principal components and principal factors, but rather the choice between an oblique or orthogonal rotation. The orthogonal rotation will yield factors that are uncorrelated, so as to minimize the multicollinearity problem. Nevertheless, once again, theory may dictate a preference for an oblique rotation, if you believe your factors are likely to be correlated, which implies the subsequent regression will have to be conducted in the presence of theoretically and empirically real collinearity.

In conclusion, for most of the research in our field, given its behavioral nature, likely errors when measuring human participants, and likely interrelations among the theoretical constructs being measured, a factor analysis is preferred to a components analysis, and an oblique rotation is preferred to an orthogonal one.

V.B. EXPLORATORY VERSUS CONFIRMATORY FACTOR ANALYSIS

My question concerns the difference between, and use of, confirmatory and exploratory factor analysis. In the past I have followed Stewart (1981), who suggested that the difference between confirmatory and exploratory is more the way you use different estimation techniques than the type of estimation technique you use. For example, Stewart referred to two examples of confirmatory factor analysis. One uses the familiar LISREL-type maximum likelihood, but the other uses a plain vanilla orthogonal rotation principle components-type factor analysis. The critical issue, according to Stewart, is that in confirmatory factor analysis, the researcher states what he or she expects to find before doing the analysis and then seeks to confirm this using the appropriate techniques. The flip side of this is that one may also be able to use techniques like LISREL or PROC CALIS (in Statistical Analysis System, or SAS) and Procrustes rotations in SAS in an exploratory fashion as well, despite the fact that they were developed to assist in confirming factor structures with correlated dimensions.

Stewart's (1981) idea appears to be in opposition to several colleagues and anonymous reviews who seem to insist that confirmatory factor analysis is a technique, and usually they seem to be referring to LISREL. In fact, I seem to find few users of factor analysis who appear to agree with Stewart. Is this another misapplication of factor analysis that has persisted despite Stewart's reminder? Therefore, what really is the difference between confirmatory and exploratory factor analysis?

REFERENCE

Stewart, David W. (1981). Application and misapplication of factor analysis in marketing research. *Journal of Marketing Research, 18,* 51–62.

Professor David Stewart
University of Southern California

The question is a good one. Unfortunately, the question makes all too clear how arcane we have made our methodologies and research tools. It is no wonder that our students and the managers we try to inform roll their eyes and ask why all this matters. It is also unfortunate, as the questioner implies, that the use of more complex techniques, regardless of whether necessary or even appropriate, seems to be a heuristic used by reviewers to determine the quality of research. Parsimony was once the rule, and it still seems like a useful principle even if the methodology is simple and inelegant.

The question suggests a not uncommon confusion between the type and purpose of a factor analysis and the method of estimation or methodological approach employed in a particular factor analysis. Exploratory factor analysis is a type of factor analysis, and the purpose is to identify the underlying dimensional structure, if any, of a set of measures. Confirmatory factor analysis is a type of factor analysis, and the purpose is to test whether an a priori dimensional structure is consistent with the structure obtained in a particular set of measures.

There are various methods for conducting an exploratory factor analysis: principal components, principal factor, alpha factor analysis, and maximum likelihood. These different methods for exploratory analysis differ in terms of how much information in the raw data is used in the analysis and how the information is used. Comparative research on these various methods for conducting an exploratory factor analysis suggest that, in most cases, the same structure is identified regardless of the method employed. This was the point of my 1981 article; it does not seem to matter what method of factor analysis is used in an exploratory analysis, especially if there is indeed a structure to be identified and the data are appropriate for a factor analysis. (In my 1981 article, I also addressed the issue of determining whether data are appropriate for a factor analysis.)

Similarly, there are various methods for conducting a confirmatory factor analysis. In the days before significant computing power on the desk and user friendly software (say 1981, when I wrote my article), rather simple procedures for confirmatory analysis included Procrustes rotations, whereby one could show that a factor structure obtained from a new data set could be rotated to match a second factor structure suggested by theory or prior research. Thus, the structure obtained from one data set could be shown to confirm an a priori hypothesis about structure by a demonstration that the two structures were consistent save for a simple transformation. The disadvantage of such an approach is the lack of strong statistical tests. Maximum likelihood approaches to confirmatory factor analysis use different information and use the information differently to provide a statistical test of the match of an obtained structure to a prior structure.

The problem raised by the questioner is indicative of a very general confusion about LISREL and related software. (LISREL is nothing more than a software incarnation of a particular approach to analysis using a specific methodology.) LISREL can be and is frequently used as a tool for exploratory factor analysis. Such use is appropriate and is merely a use of a maximum likelihood approach to estimate in an exploratory context. Unfortunately, in the vast majority of cases, LISREL is presented as a confirmatory technique when it is, in fact, being used as an exploratory tool (i.e., as a maximum likelihood approach to exploratory factor analysis). Many, many articles have been published as confirmatory analyses when, in fact, they were exploratory analyses. This is called shame on the editor, shame on the reviewers, and shame on the authors. In fact, I would go so far as to suggest that most publications in the marketing literature that have employed LISREL were published as a result of the use of and the obfuscation associated with LISREL rather than the substantive contribution of the article: They were exploratory factor analyses that would have been rejected but for the erroneous representation that LISREL was being used as a confirmatory tool.

Although it is the case that LISREL offers statistical measures of fit and significance, merely showing a fit of data to a hypothesis is not a reason for accepting the results of a LISREL analysis as a useful contribution. Merely suggesting a structure and showing that data fit the suggested structure is not a genuine exercise in confirmatory factor analysis. An acceptable use of LISREL as a confirmatory tool requires at least three conditions:

1. A genuine, strong theory that posits a strong and unambiguous structure of relations among constructs and the variables that represent these constructs.
2. There must be a strong and unambiguous a priori structure that serves as the basis for the test of fit.
3. The fit of the data to the a priori structure must be better (by some acceptable criterion) than the fit to structures suggested by alternative theories; alternative structures that would be consistent with the theoretical foundation; intuitively obvious alternative structures; or structures that could be readily explained on methodological grounds, such as the presence of highly correlated error terms.

Yet another misapplication of factor analysis is the unthinking use and acceptance of confirmatory factor analysis methods, like LISREL, when they are inappropriate. Given the history of its application, use of LISREL should be a heuristic for very close examination and probable rejection, rather than acceptance based on a novel, complex, and difficult to understand methodology. Therefore, yes, most applications of LISREL are misapplications of factor analysis, though not what I had in mind in 1981. On the other hand, we may need more interesting exploratory factor analyses that are labeled as such. However, since 1981, there are much better tools for identifying structural models (see Glymour, Scheines, Spirtes, & Kelly, 1989).

REFERENCE

Glymour, Clark N., Scheines, Richard, Spirtes, Peter, & Kelly, Kevin. (1987). *Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling.* San Diego, CA: Academic.

Editor: I like the tone of the question, which suggests that confirmatory factor analysis may be, sort of, a state of mind. Stewart (1981) is solid and not too dated (given advances in computing), but he actually did not have much to say about the distinction between exploratory and confirmatory factor analysis (p. 56).

My understanding is that variants of factor analysis may be placed along a continuum, with exploratory factor analysis on one end, where the researcher lets the data speak to the appropriate number of factors to extract, along with the estimation of the values for all the factor loadings (cf. Proc Factor in SAS, with a Priors = SMC statement and a Rotation = Promax option, or indeed, one can fit an exploratory factor analysis through structural equations modeling software like LISREL or EQS, by simply ignoring the structural paths and specifying only the measurement side of the model).

A more confirmatory factor analytic approach would be to specify the expected number of factors that should suffice in describing the data. An exploratory maximum likelihood factor analysis may be used to obtain a significance test of whether one, then two, then three, then four, and so on, factors fit the data at acceptable levels. Alternatively, once again, the researcher can coax this information from LISREL.

Further along the continuum is probably what most people think of when they hear confirmatory factor analysis—that is, the specification of both the number of factors to be extracted and indicators as to which variables should load on which factors. Solutions may be obtained through a procrustes rotation (e.g., in SAS), though this is a fairly clumsy, dated method; rather, at this point, most analysts rely on the LISREL class of software, specifying the number of exogenous factors and a LX (lambda X) matrix of ones (indicating parameters to be estimated) and zeros (indicating parameters to be fixed, usually to zero; i.e., theory predicts those variables would not load on those factors).

At the most confirmatory end of the continuum would be an analysis like a cross validation, where researchers would have theoretical reasons or past empirical evidence to believe they could predict the number of factors, the pattern of which variables should load on which factors, and the actual values of those loadings. In multigroup factor analyses (e.g., addressing the question of whether variables and constructs are interrelated in the same manner for a consumer database from Constantinople compared to one from Timbuktu), one might do even a wholly exploratory factor analysis in one sample and test the extent to which that solution describes the data in the other sample.

A confirmatory factor analysis can be conducted to the extent a researcher has theory and guidance regarding the expected factor structure. As the question states, the estimation procedures, even the software, are somewhat confounded with the typical use of exploratory or confirmatory approaches—for example, in using Proc Factor in SAS for an exploratory analysis and LISREL for a confirmatory one—but these computing packages are fairly sophisticated and encompassing, so the researcher can use also Proc Calis in SAS for a confirmatory analysis and LISREL for an exploratory one. The techniques, or "models," are indeed different though, with an eigenvalue–eigenvector decomposition of a correlation or covariance matrix (either adjusted by initial communality estimates, usually R_{SMC}) to obtain the principal axis orientation followed with a rotation to simpler structure for an exploratory factor analysis and a maximum likelihood estimation-based model for the confirmatory analysis (or, when Maximum Likelihood Estimation acts like the tempermental mistress she can be, an alternative unweighted or generalized least squares procedure, Unweighted Least Squares, General Least Squares). Therefore, yes, the distinction between exploratory and confirmatory factor analysis can be viewed both as an approach or logic and as a set of models or techniques. For more information on confirmatory factor analysis, see Basilevsky (1994, pp. 414–416) or McDonald (1985, chap. 3).

REFERENCES

Basilevsky, Alexander. (1994). *Statistical factor analysis and related methods.* New York: Wiley.

McDonald, Roderick P. (1985). *Factor analysis and related methods.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stewart, David W. (1981). Application and misapplication of factor analysis in marketing research. *Journal of Marketing Research, 18,* 51–62.

V.C. ORTHOGONAL OR OBLIQUE ROTATION?

Is an orthogonal rotation always the best for factor analysis?

Editor: A vehement no; in fact, I argue that an oblique rotation is almost always preferred in factor analysis. Initially, factors are extracted so as to optimize a principal axis orientation (of an ellipsoid of a cloud of data points in p dimensions). Since almost the beginning of principal factor modeling, however, researchers have held this mathematical quality to a lesser import than Thurstone's (1947) theoretically more useful criteria of "simple structure" (Comrey & Lee, 1992; Harman, 1976; Iacobucci, 1994; Kim & Mueller, 1978a, 1978b; Lawley & Maxwell, 1971; Mulaik, 1972). Hence, we rotate factors around in space, sometimes retaining the property that the axes are still orthogonal (perpendicular in geometry, uncorrelated in statistics), but often relaxing that assumption and entering into the interesting world of oblique axes. Although that new world may be less simple geometrically, it is very likely to provide a more valid representation of the phenomena under study. Factors capture the extent to which the variables are correlated, and an oblique rotation allows us to examine the extent to which the factors themselves are intercorrelated.

It is important to think about factor analysis theoretically and its intended utility in measurement and the construction of indicators of constructs. When we write items to compose a scale, we begin with some theoretical expectation regarding which items are likely to tap which constructs, which we hope is supported by the factor analysis. I suppose it could be the case that we might write a scale or survey in which each of the separable factors were expected to be unrelated with each other, but far more likely is either the scenario that (a) the scale contains items that tap multiple facets or factors of higher order factors and should thereby be moderately to fairly highly correlated or (b) items on the survey tap multiple factors that may represent different constructs but are believed to be complementary in predicting the focal attitude or behavior and hence also are probably at least somewhat correlated. If one creates a domain of items intended to tap some theoretical content area, and more than one factor is extracted from those items, it is difficult to envision a scenario under which those multiple factors would not be correlated. Either of these scenarios would be modeled better by correlated factors. Hence, an oblique rotation, one that allows for correlations among factors, would seem to describe the theoretical world better and more often than the arbitrary constraint of orthogonality.

A good choice of an oblique rotation procedure is SAS's Proc Factor's Rotation = Promax option. Promax is a routine that begins with an orthogonal varimax rotation (so you get the orthogonal rotation as a byproduct of the output if you want to compare) and then relaxes the solution to an oblique rotation, using the criteria presented in Hendrickson and White (1964). The method probably works well and robustly because it is fairly simple—after varimax loadings have been estimated, they are essentially raised to powers, so that high loadings (e.g., .8) become a little lower, but low loadings (e.g., .2) disappear to nearly zero. I know that relying on personal communication is no longer in vogue in journals, perhaps because it is unverifiable, but the simulation work that

Ledyard Tucker (himself, a former student of Thurstone's) did comparing various packages' oblique rotation procedures is not published. Tucker has his own personal program that performs best, but it is not widely available. Promax, in SAS, performed nearly as well as his optimal procedure, and both dominated SPSS's Oblimin rotation method in true factor recovery. (Tucker's procedure, *direct artificial probability factor rotation,* and Promax and its performance, are discussed in his book with Robert MacCallum on factor analysis, which may be found at MacCallum's website: http://quantrm2.psy.ohio-state.edu/maccallum/factornew.htm, pp. 398–400, chap. 11.)

One might begin the factor analysis with a presumption of correlated factors, extract the factors and perform an oblique rotation, and examine the factor intercorrelations. If the factors are all pairwise uncorrelated (e.g., .3 or lower in magnitude), then one might proceed to report the orthogonal rotation, because the model is somewhat simpler to conceptualize, and, for example, one need not report an additional matrix (of factor intercorrelations). However, if any pair of factors are reasonably correlated, it would seem to be an error in conceptualization to treat them as if they were not. (Factor intercorrelations can also be diagnostic at the high end; i.e., factors that are almost too highly correlated—e.g., .7 or more—often indicate too many factors were retained.) Therefore, it would seem there is a general preference for an oblique rotation over an orthogonal one.

Having said all that, answers in statistics also depend on your subsequent analytical needs. If the intentions are to create factor groupings of items to make composite variables that will be entered into a multiple regression, clearly it would be easier to interpret the beta weights without the complication of multicollinearity, so one might retain the orthogonal rotation, acknowledging that it might not express the construct–variables relation optimally.

REFERENCES

Comrey, Andrew L., & Lee, Howard B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Harman, Harry H. (1976). *Modern factor analysis.* Chicago: University of Chicago Press.

Hendrickson, Alan E., & White, Paul Owen. (1964). Promax: A quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology, 17,* 65–70.

Iacobucci, Dawn. (1994). Classic factor analysis. In Richard P. Bagozzi (Ed.), *Principles of marketing research* (pp. 279–316). Cambridge, MA: Blackwell.

Kim, Jae-On, & Mueller, Charles W. (1978a). *Factor analysis: Statistical methods and practical issues.* Beverly Hills, CA: Sage.

Kim, Jae-On, & Mueller, Charles W. (1978b). *Introduction to factor analysis: What it is and how to do it.* Beverly Hills, CA: Sage.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method.* New York: Elsevier.

Mulaik, Stanley A. (1972). *The foundations of factor analysis.* New York: McGraw-Hill.

Thurstone, Louis L. (1947). *Multiple factor analysis.* Chicago: University of Chicago Press.

V.D. CHECKING ASSUMPTIONS OF NORMALITY BEFORE CONDUCTING FACTOR ANALYSES

My question has to do with the lack of heuristics in using statistical analyses methods. My specific question is the following: I have a rather large data set ($N = 350$). I want to apply factor analysis, but on forehand, I am checking normality assumptions regarding the variables. Looking at normal probability plots gives us an indication, but it does not provide any guarantee that the variable is normally distributed. Then, I looked at the Kolmogorov–Schmirnof (K–S) test in SPSS, but because of the size of the sample, this test almost always rejects the null hypothesis of normality. Then, instead of looking at the significance of the test, you can also look at the value. But, what heuristic can I use?; what is a large value and what is small? I am checking normality to be able to use factor analysis with maximum likelihood estimates. My question, therefore, is, Are normal probability plots sufficient to check the assumption of normal distribution? If not, what size of the K–S is big and what is large, or is there another test I can use? (I already used the z statistic, which is defined as the skewness divided by $(\sqrt{6})/N$.) I think there is a great need for heuristics regarding, for example, SPSS statistical analyses results.

Professor James Barnes
University of Mississippi

Among the nongraphical tests for univariate normality are the chi-square goodness-of-fit, K–S, the Shapiro–Wilk test, and the use of the skewness and kurtosis coefficients. Madansky (1988) summarized a number of studies that examined four classes of normality violations: (a) symmetric, platykurtic (short tailed); (b) symmetric, leptokurtic (long tailed); (c) asymmetric, platykurtic; and (d) asymmetric, leptokurtic. The general conclusion of these studies was that as an omnibus test of normality, the Shapiro–Wilk statistic w came out best regardless of the class of violation, the K–S test was shown to be less powerful, and the chi-square suffered because of a dependence on the number of intervals used for the grouping.

In terms of power considerations, both the Shapiro–Wilk or the combination of skewness and kurtosis coefficients methods appeared equally effective. Stevens (1992) recommended the combination skewness and kurtosis coefficients because this allows for the separation of the two types of normality violations, whereas the Shapiro–Wilk test combines both into a single test.

> The reason that we wish to separate them out is because kurtosis has been shown in both the univariate and multivariate cases to have an effect on power, whereas skewness has not been shown to affect power (at least in the univariate case). (p. 253)

Tables of critical values for skewness and kurtosis for small sample sizes and formulas for large sample sizes are provided in Stevens (p. 255).

REFERENCES

Madansky, Albert. (1988). *Prescriptions for working statisticians.* New York: Springer-Verlag.

Stevens, James. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Professor Joseph Cote
Washington State University

First, let me note that many statistical procedures are fairly robust to violations of normality (which is why many researchers never test for normality). Other procedures, such as structural equation models, are sensitive to violations of the normality assumption. (I do not know if factor analysis is robust to violations of normality.) Second, I should note a preference for confirmatory factor analysis over exploratory analysis. If you use the computer program Eqs or Prelis, you can use estimation procedures for non-normal distributions. Now, I will assume that my attempts to avoid the question have not worked. I do not know of any rules of thumb to say a certain test value (irrespective of statistical significance) indicates normality. As such, I would consider several tests including Wilk–Shapiro and stem-and-leaf plots (to test for symmetry). Assuming the data are symmetric, I would consider running several different estimation procedures including principle components, least squares estimation, and so on, to check for consistency of results. If nothing looked funny, I would assume the estimates are appropriate. An interesting final note from Nunnally (1978) stated,

> Strictly speaking, test scores (for trait scales) are seldom normally distributed, even if the number of items is large. Because of the positive correlation among items, a normal distribution would not be obtained ... A precisely normal distribution of test scores ... would usually represent *dead data.* (p. 160)

Nunnally stated that test scores would have flattened distributions. Nunnally also provided an alternative test of symmetry but offered no heuristics for application. Of course, his comments only apply to summated scales.

REFERENCE

Nunnally, Jum C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Professor Robert Cudeck
University of Minnesota

This specific question deals with judgments regarding tests for normality for variables that will be used in a factor analysis, but another more general question seems to be in the background. The particular question is that standard tests such as the K–S procedure have power characteristics that lead to rejecting the null hypothesis of normality in large samples. In fact, when the test rejects the null hypothesis, it is doing what is expected. Variables rarely are normally distributed and in large samples; the K–S test should detect the situation. Probably in strict terms the question is a nonissue from the beginning: Virtually no variable follows the normal distribution.

The background question is then what to do about the non-normality if a method such as maximum likelihood factor analysis, which depends on the appropriateness of the assumption, is used? This question is difficult to answer for two reasons. It is well known in statistics in general, and certainly in the factor analysis literature in particular, that methods that assume normality run into trouble when the assumption is violated. At issue is the quality of the parameter estimates and the accuracy of the measures of fit. In general, use of data that are extremely non-normal affects the reliability of all the results. At the same time, it is equally well known that use of variables that are more or less symmetric raise no practical problems with this method at all. In fact, there is a theme in the literature dealing with robustness of factor analysis estimators showing that under some circumstances the method is robust with respect to many different kinds of non-normality. Both of these summaries are true: Data that depart markedly from normal can affect the results in drastic ways; the method is robust to violations of the normality assumption in some circumstances. What should one do for practical problems?

A tentative suggestion is the following: If the distributions of the sample variables are not wildly non-normal, use maximum likelihood and do not worry too much about the consequences. The results are probably trustworthy for most purposes. If the lack of normality is so severe as to be worrisome, consider another method that does not require the same strict distributions to be valid. The discussion in Bollen (1995, chap. 9) is quite good on these matters.

REFERENCE

Bollen, Kenneth A. (1995). *Structural equations with latent variables* (2nd ed.). New York: Wiley.

Professor Edward Malthouse
Northwestern University

"All models are wrong but some are useful" (Box, 1979, p. 202).

The goal of an exploratory factor analysis is to understand the correlation structure of some variables—how many factors are there and which variables load on them? We should pay careful attention to any characteristics of the data that might prevent us from achieving this goal. Consider two examples. By definition, data that come from 7-point scales are not normally distributed. In fact, the distributions of variables measured on such scales are often skewed toward one end of the scale, uniform, or even bimodal. This does not mean that maximum likelihood factor analysis, or other flavors that do not assume normality such as principal factor analysis (Priors = SMC in SAS), cannot be useful tools for understanding the correlation structure.

As a second example, a characteristic that can affect the usefulness of factor analysis is the presence of outliers or extreme values. Correlation matrices, the sufficient statistics for factor analysis, can be changed greatly by extreme values. (See my response to the question about outliers in this special issue; everything I said there concerning the effects of outliers on regression applies here also.) As a starting point for identifying influential cases, study the marginal distributions, for example, with normal probability plots, boxplots, extreme quantiles, and so on. Where there are outliers that cannot be attributed to coding errors, and so on, you should do something to reduce the influence of these extreme observations on the estimation. For example, count data typically have many outliers. Therefore, I suggest Winsorizing count variables before analysis. The variance of counts is a random variable that usually increases with its mean (e.g., think of the Poisson distribution). Most factor analysis methods assume that the specific variance is constant across observations. This assumption is like homoscedasticity in regression analysis. One way to address this problem is to transform the data with logarithms or square roots. See Tukey (1977) for further discussion of such transformations.

REFERENCES

Box, George. (1979). Robustness is the strategy of scientific model building. In Robert L. Launer & Graham N. Wilkinson (Eds.), *Robustness in statistics* (p. 202). New York: Academic.

Tukey, John W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.

Editor: For more information on exploratory data analysis, methods that make minimal assumptions, if any, and are often graphics-based (e.g., roughly estimating a regression coefficient off of a plot with a ruler), see Tukey (1977) or Cleveland and McGill (1984), Gentleman and Crowley (1991), Hoaglin, Mosteller, and Tukey (1985), or Velleman and Hoaglin (1981). For more information on nonparametric statistics, which make minimal distributional assumptions, see Gibbons (1993), Lehmann and D'Abrera (1975), Siegel (1956), or Mooney and Duval (1993).

REFERENCES

Cleveland, William S., & McGill, Robert. (1984). The many faces of a scatterplot. *Journal of the American Statistical Association, 79*(388), 807–822.

Gentleman, R., & Crowley, J. (1991). Graphical methods for censored data. *Journal of the American Statistical Association, 86*(415), 678–683.

Gibbons, Jean Dickinson. (1993). *Nonparametric statistics: An introduction.* Newbury Park, CA: Sage.

Hoaglin, David C., Mosteller, Frederick, & Tukey, John W. (1985). *Exploratory data tables, trends, and shapes.* New York: Wiley.

Lehmann, E. L., & D'Abrera, H. J. M. (1975). *Nonparametrics: Statistical methods based on ranks.* San Francisco: Holden-Day.

Mooney, Christopher Z., & Duval, Robert D. (1993). *Bootstrapping: A nonparametric approach to statistical inference.* Newbury Park, CA: Sage.

Siegel, Sidney. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Velleman, Paul F., & Hoaglin, David C. (1981). *Applications, basics and computing of exploratory data analysis.* Boston: Duxbury.

V.E. HOW HIGH OR LOW MUST LOADINGS BE TO KEEP OR DELETE A SCALE ITEM?

In a factor analysis, how high or low should a factor loading be for you to keep or remove a scale item?

Editor: I suppose the answer you are looking for is, "the magnitude of the factor loading must be at least .30" (for the typical rotated pattern coefficient matrices produced by SAS or SPSS on a correlation matrix, or standardized loadings in other applications). This answer is a rule of thumb, and like all subjective standards, it is simplistic in that the right answer depends on a number of issues: Is the research exploratory, or are you working with items that have fairly well-known structural properties?; Is the sample size small or large?; Are there many variables with messy structure or few variables with clean structure?; Is the loading a cross loading or is it the largest loading on any factor for that variable?; Does the loading in question appear on one of the later extracted factors or on one of the initial factors?; Are the factors themselves correlated or not?; Is there any concern that the factor analytic assumptions (e.g., multivariate normality) may not hold or are you fairly certain they do? In the first of each of these paired scenarios, you may wish to operate somewhat more conservatively, using a cutoff rule like .40 to begin to compensate for the likely noisier data quality.

In any event, individual loadings of variables on factors were never intended to be used diagnostically in isolation. If one is crafting a scale, whether the production of the scale is the end goal or it is a preprocessing stage for additional subsequent analyses, it is still best to use some good old-fashioned, but to-date not supplanted, psychometric techniques, such as examining item-total correlations, coefficient alphas, and so on.

Added to that repertoire, we can use maximum likelihood estimation in a confirmatory factor analysis to see whether a proposed factor pattern fits the data. One may fit two factor patterns—one that includes the borderline loadings and one that does not, and see if the fit is not appreciably altered (if it is, you will need to compare models nested more cleanly; i.e., dropping only one loading at a time to identify those that should be retained vs. those that may be eliminated).

A sensible test would be to fit the factor analysis via software for structural equations modeling, like LISREL (Jöreskog & Sörbom, 1996, pp. 26, 103), and examine the standard errors for the factor loadings (and factor intercorrelations) to form confidence intervals and conduct significance tests (see also Cudeck & O'Dell, 1994). Keep in mind the usual concerns with regard to large requisite samples and strict multivariate normality, and that doing so explicitly tests numerous hypotheses simultaneously (so use a conservative alpha level, etc.).

REFERENCES

Cudeck, Robert, & O'Dell, Lisa L. (1994). Applications of standard error estimates in unrestricted factor analysis: Significance tests for factor loadings and correlations. *Psychological Bulletin, 115*, 475–487.

Jöreskog, Karl, & Sörbom, Dag. (1996). *Lisrel 8: User's reference guide.* Chicago: Scientific Software International.