Factor Analysis in Counseling Psychology Research, Training, and Practice:
Principles, Advances, and Applications

Jeffrey H. Kahn
Illinois State University

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) have contributed to test development and validation in counseling psychology, but additional applications have not been fully realized. The author presents an overview of the goals, terminology, and procedures of factor analysis; reviews best practices for extracting, retaining, and rotating factors in EFA; describes issues related to CFA, including potential hypotheses, how parameters are estimated, and current guidelines regarding sample size and the use of fit indices; and reviews more sophisticated applications of CFA, including multiple-groups analysis and testing of hierarchical models. Finally, the author recommends how to increase the application of factor analysis to counseling psychologists' work in research, training, and practice settings.

A psychologist in a counseling center is interested in understanding the similarities among the diverse symptoms of the center's clients. Of particular interest is the possibility of grouping clients' symptoms, as rated on an intake checklist, into categories to simplify assessment and treatment decisions. A psychologist in a research setting is interested in whether global and domain-specific measures of self-efficacy are distinguishable constructs. This psychologist plans to analyze data from multiple self-report measures of self-efficacy in a way that can contribute to a body of knowledge on social-cognitive theory. A faculty member in a counseling psychology program is developing a measure of graduate students' perceptions of their relationships with their graduate advisors. As part of this process, the faculty member wishes to understand how to group items on this instrument into similar dimensions to use the test scores to enhance advising and training.

These three cases share a common goal: understanding how to reduce several variables to a smaller number of groups or categories. Reducing data may solve many problems facing counseling psychologists. Such data may come

---

The author wishes to thank Margaret Nauta and Daniel Russell for their helpful comments on an earlier draft of this article. Correspondence concerning this article should be addressed to Jeffrey H. Kahn, Department of Psychology, Campus Box 4620, Illinois State University, Normal, Illinois 61790-4620; e-mail: jhkahn@ilstu.edu.

THE COUNSELING PSYCHOLOGIST, Vol. 34 No. 5, September 2006 684-718
DOI: 10.1177/0011000006286347

from clients in treatment (e.g., reducing a complex set of presenting problems into themes), students in a graduate training program (e.g., reducing a long list of training goals into core components), or participants in a research study (e.g., finding the common factors among items on a new assessment instrument). Factor analysis is a multivariate technique well suited to addressing these problems.

Factor analysis describes a set of related techniques rather than a single method. Nearly all forms of factor analysis describe how a small number of latent (i.e., not directly measurable) constructs might explain covariation among a larger number of measured variables, such as scores on test items, subtest scores, or behavior ratings (Tinsley & Tinsley, 1987). The measured variables are typically continuous scores, but they may also be dichotomous categories. The factors themselves are construed as unseen, hypothetical constructs. Put another way, psychologists frequently view factors as the underlying reasons that individuals attain the scores they do on the measured variables. Although the goal of a typical factor analysis is to find factors that explain covariation among variables, factor analysis can also identify common factors among people (*Q*-technique factor analysis) or common factors among occasions of measurement (*P*-technique factor analysis; Comrey & Lee, 1992). Thus, factor analysis is a flexible data-analytic method.

Factor analysis has a long tradition in classical measurement theory (Nunnally & Bernstein, 1994), and it has enjoyed widespread use in scale-development research and as a tool to assess test-score validity (Hoyt, Warbasse, & Chu, in press; Worthington & Whittaker, in press). Factor analysis has many other potential applications, however. Researchers may create factor scores as a means to reduce data (Tinsley & Tinsley, 1987); they may then use these factor scores as data for decisions or other statistical analyses. Researchers may also use factor analysis to develop and test theory (Russell, 2002; Thompson, 2004). It is also possible for factor analysis to aid the work of practitioners and educators, for example, by helping to identify themes and content areas that warrant further attention in counseling or in graduate training. Thus, factor analysis has the potential to advance the work of counseling psychologists across diverse settings.

Tutorials on the proper use of factor analysis have appeared in counseling psychology journals (e.g., Tinsley & Tinsley, 1987) as well as journals in other psychology disciplines (e.g., Russell, 2002). The complexity of factor analysis and the rapid developments in its methods warrant continuing attention to its appropriate application to counseling psychology. This article explains the logic, terminology, procedures, and potential applications of factor analysis to counseling psychology research, training, and practice. This article reviews basic principles and recent developments in both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The article's goals are to help the reader evaluate results of factor analyses presented in the literature and to provide the reader with a basis for conducting a factor analysis. For the reader unacquainted with factor analysis, concepts appear in simple terms when possible. For the reader with a basic knowledge of factor analysis, this article updates best practices as well as some underused applications of factor analysis in counseling psychology.

EFA

Goals of EFA

The primary goal of EFA is to identify latent factors that explain the covariation among a set of measured variables. Specifically, EFA explores how many factors exist among a set of variables and the degree to which the variables are related to the factors. For example, a 33-item measure of a graduate student's self-efficacy with respect to research tasks (e.g., Phillips & Russell, 1994) could represent 33 unique constructs, each focusing on a different domain of research self-efficacy (e.g., self-efficacy with respect to writing an introduction and literature review section of a dissertation, self-efficacy with respect to reviewing the literature in an area of interest). It would challenge many faculty members to help a student master all 33 separate research tasks, so a reduction of these data would serve a useful training function. Indeed, some of these 33 items are likely similar enough to be considered the same construct or factor (e.g., writing self-efficacy). EFA's goal in this case is to reduce 33 variables into a smaller number of factors, thereby achieving a more parsimonious explanation of the variables.

The meaningfulness of the factors that emerge in a factor analysis is dependent on the meaningfulness of the variables. This idea has led to the "garbage in, garbage out" adage of factor analysis—that is, if meaningless variables are analyzed, then meaningless factors will emerge. Thus, the variables ideally form a logical group (e.g., items on an instrument, behavior ratings within a given domain). If the variables do not have much correlation, then finding common factors will be a challenge (Nunnally & Bernstein, 1994), although artificial redundancy, such as that occurring by including both standardized and unstandardized forms of the same variable, is to be avoided (Tinsley & Tinsley, 1987). Moreover, a researcher must analyze a sufficient number of variables to generate stable factors; at least six variables need to be associated with each factor to have replicable factors

(Gorsuch, 1983). Finally, variables that do not have adequate reliability (see Helms, Henze, Sass, & Mifsud, 2006 [this issue]) are unlikely to form meaningful factors (Gorsuch, 1983; Nunnally & Bernstein, 1994).

Identification of Factors

Extraction method. The method of identifying the factors that best characterize a set of variables is known as *factor extraction*. There are several factor-extraction methods (see Gorsuch, 1983), but three are frequently used in contemporary counseling psychology research.[1] They are the principal-axis factoring (PAF) method, principal components analysis (PCA), and the maximum likelihood (ML) method.[2]

PAF analyzes the common variance among variables. Common variance, also termed *communality*, is the variance in a variable that is shared with at least one other variable in the analysis. To understand communality more concretely, consider a regression analysis in which one's self-efficacy rating for knowing which statistics to use is predicted by self-efficacy ratings for 32 other research tasks. The R^2, or squared multiple correlation, for the 32 self-efficacy ratings predicting self-efficacy for knowing which statistics to use is the communality. If there is a lot of overlap between self-efficacy for knowing which statistics to use and the other 32 tasks (e.g., avoiding the violation of statistical assumptions), then the R^2, and therefore communality, will be high. Researchers typically use this squared multiple correlation as the initial estimate of communality for PAF in statistical software programs, including those by SPSS (2004) and SAS (SAS Institute, 2004). Only common variance is analyzed in the PAF method because variance not shared with another variable cannot be due to any underlying common factor. The specification equation gives a broader view of this issue, stating that a variable's variance is due to (a) variance common to other variables (i.e., communality), (b) variance specific to the variable (i.e., reliable variance independent from other variables in the analysis), and (c) random measurement error. By using the squared multiple correlation as an estimate of how much variance is to be analyzed, the PAF method essentially ignores specific variance and measurement error. Thus, the emerging factors must explain variance common to more than one variable.

In PCA, researchers analyze all the variance among variables. In other words, PCA not only focuses on common variance but also analyzes variance unique to each variable. In operational terms, PCA analyzes correlations among variables with values of 1 on the matrix diagonal, whereas PAF analyzes a correlation matrix with (typically) squared multiple correlations on the diagonal. Tables 1 and 2 illustrate the different correlation matrices

TABLE 1: Correlations Among Eight Items With Squared Multiple Correlations on the Diagonal (Suitable for Principal-Axis Factoring Extraction)

	Variable							
Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1	.60							
2	.63	.52						
3	.40	.41	.50					
4	.31	.25	.62	.57				
5	.31	.27	.54	.72	.53			
6	.36	.28	.67	.79	.69	.70		
7	.66	.51	.40	.35	.37	.42	.57	
8	.66	.67	.42	.30	.31	.30	.55	.71

TABLE 2: Correlations Among Eight Items With Unities on the Diagonal (Suitable for Principal Components Analysis Extraction)

	Variable							
Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1	1.00							
2	.63	1.00						
3	.40	.41	1.00					
4	.31	.25	.62	1.00				
5	.31	.27	.54	.72	1.00			
6	.36	.28	.67	.79	.69	1.00		
7	.66	.51	.40	.35	.37	.42	1.00	
8	.66	.67	.42	.30	.31	.30	.55	1.00

that researchers would analyze for PAF and PCA, respectively. These tables show the same correlations among eight variables, but whereas Table 2 has unities along the main diagonal, Table 1 has squared multiple correlations. This seemingly small difference has a profound effect on the assumptions of these two methods. Because the goal of factor analysis is, by definition, identifying common factors, PCA is not a true factor analysis (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Instead, the goal of PCA is to find linear combinations (i.e., principal components) that retain as much information about the measured variables as possible (Fabrigar et al., 1999).

A less common method of extracting factors in counseling psychology research is the ML method. The ML method assesses the likelihood that the correlation matrix is derived from a population where the attained factor structure underlies the scores on the variables. One benefit of the ML method is that it provides the means to conduct significance tests and to derive confidence intervals (Fabrigar et al., 1999). Its chief drawback is the requirement of data that are multivariate normal. Perhaps because of this requirement, researchers do not use the ML method as often as PAF or PCA.

Although extraction methods differ in theory, they are often used in practice for the same purpose. This is potentially problematic because whereas PCA and PAF provide similar information in many circumstances (Velicer & Jackson, 1990), simulation studies have found PAF results to be more accurate than PCA results when communalities are low (Snook & Gorsuch, 1989; Widaman, 1993). For these reasons, when the goal is to determine latent factors that underlie one's data (which is the goal of factor analysis proper), PAF or ML are preferable to PCA (Fabrigar et al., 1999; Russell, 2002). It is important to note, however, that PCA is the default extraction procedure in SPSS and SAS software, so users need to specify PAF or ML if that is their choice.

Determining the number of factors to retain. One may extract as many factors as there are variables, but retaining a smaller number of factors than variables helps to reduce data. SPSS and SAS retain a limited number of factors on the basis of the software's default decision rules. Users will frequently want to override these defaults because the number of factors to retain should be based on principles of parsimony, and this requires human judgment. Statistics aid this judgment, indicating how closely the factors reproduce the covariation among the variables with different numbers of factors. Factor analysts attend to the percentage of variance each factor explains, which is a function of the strength of association between the variables and the factor. For example, one factor might explain 40% of the variance among variables, whereas two factors might combine to explain 50% of the variance among variables. In this case, the second factor seems important. If a third factor explains only an additional 2% of variance, then it might be best to retain two factors because, even though two factors explain slightly less variance than three, two factors are more parsimonious than three.

Determining how many factors to retain may be based strictly on theory. For example, if a new measure of the Big Five personality factors were developed, the factor analyst may retain five factors to see if they correspond to expectations. Oftentimes, however, the factor analyst lets the data guide the decision about retaining factors. Empirical decisions should be based on multiple criteria (Fabrigar et al., 1999; Tinsley & Tinsley, 1987). Favoring some criteria over others may result in divergent judgments about how many factors to retain, even for the same data set. This potential criticism of EFA is best addressed with a thorough understanding of how to apply empirical decision rules.

One intuitive method of determining the number of factors to retain is based on the percentage of variance among variables explained by each factor. Factors that explain a large enough percentage of variance to be deemed important may be retained, and those that explain only a small percentage of variance may be discarded. Of course, this involves subjectivity as to what percentage of variance is large enough to be important, and there are no agreed-upon rules to guide these decisions. One objective approach is to base the factor-retention decision on the amount of common variance among items. Specifically, it may be useful to retain the number of factors required to explain all the common variance among the variables. The common variance, which is the degree of overlap among variables, is estimated by the sum of the variables' squared multiple correlations. Indeed, this is the SAS software's default method of retaining factors when conducting a PAF factor analysis.

Other decision rules are based on the eigenvalues of the factors. A factor's eigenvalue is a measure of explained variance set on the same metric as the variables. Thus, a factor with an eigenvalue of 1 explains as much variance among variables as a single variable does (because variables are standardized to have a variance of 1). Kaiser's criterion is a decision rule to retain factors with eigenvalues of at least 1. The logic is that for a factor to be meaningful, it cannot explain less variance than a single variable does. Because Kaiser's criterion requires eigenvalues to equal or exceed the total variance for a variable, this criterion is appropriate only for PCA (Comrey & Lee, 1992; Fabrigar et al., 1999; Russell, 2002). Thus, because extraction methods such as PAF and ML assume that the variance of a variable is less than 1, it is a misapplication to apply Kaiser's criterion with a factor-extraction method other than PCA, yet this occurs alarmingly often in the literature. Moreover, Kaiser's criterion will often yield too many factors retained, particularly when analyzing a large number of variables (Fabrigar et al., 1999; Zwick & Velicer, 1986). Kaiser's criterion is SAS's default retention method when conducting a PCA, and it is SPSS's default retention method *regardless of the extraction method used*. To emphasize this point, when conducting a PAF or ML factor analysis in SPSS, the software defaults to a wholly inappropriate retention method for the method of extraction. Thus, it is critical not to let SPSS determine the number of factors to retain in this circumstance.

A better method of determining the number of factors to retain when doing a common factor analysis is to examine a plot of eigenvalues for the

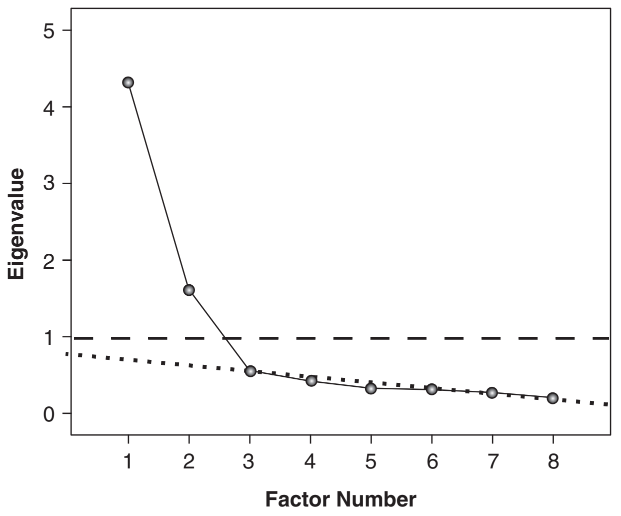

FIGURE 1. A Scree Plot for Eight Items Measuring Research Self-Efficacy
NOTE: The dashed line indicates the point at which eigenvalues are higher than 1 (Kaiser's criterion). The dotted line indicates the line through the scree.

factors, known as a scree plot. To illustrate, I conducted an EFA of eight items from the Self-Efficacy in Research Measure (Phillips & Russell, 1994). I based this analysis (n = 937) on a survey of graduate students in psychology conducted by Forester, Kahn, and Hesson-McInnis (2004). The correlation matrix in Table 1 was analyzed using PAF extraction. A scree plot from this analysis is illustrated in Figure 1. The eight factors are from left to right in order of explained variance. (The factors do not correspond to specific items.) A dashed line indicates where eigenvalues equal 1, corresponding to Kaiser's criterion. From this dashed line, only the first two factors are retained, as they are the only factors with eigenvalues of at least 1. Cattell (1966) suggested that a nonhorizontal line would provide a better indicator of factor importance. Cattell suggested drawing a scree line through the point at which the eigenvalues flatten out (the dotted line in Figure 1). Factors that are above the scree line should be retained. Note that although there is a substantial drop between the first and the second factors (as is often the case when analyzing variables that measure the same construct), using the scree test involves looking for the last substantial drop in eigenvalues. Thus, one should retain the first two factors because the last substantial drop in eigenvalues occurs after the second factor. In this example, Kaiser's criterion and Cattell's scree-plot method agree that two factors are appropriate to retain, but these methods will not always converge. The scree test does not have the same problem as Kaiser's criterion with indicating the retention of too many factors, and it is appropriate for PAF and ML extraction. One primary drawback of the scree test is its subjectivity, particularly when there is no clear scree line.

Perhaps the most effective method of deciding how many factors to retain is a parallel analysis (Horn, 1965). A parallel analysis generates eigenvalues from a random set of data based on the same number of variables and the same number of cases. The analyst plots these randomly generated eigenvalues on a scree plot along with the actual eigenvalues. The factors with actual eigenvalues higher than the random eigenvalues are retained with the logic being that a factor that explains more variance than chance is meaningful. Research has found parallel analysis to be superior to Kaiser's criterion and the scree test for identifying factors (Zwick & Velicer, 1986), and it is therefore recommended in deciding how many factors to retain. At this time, there is no straightforward way of conducting a parallel analysis in software such as SPSS or SAS, but Thompson and Daniel (1996) have provided a simple syntax program for generating random data for use in a parallel analysis in SPSS.

Given their strengths and limitations, these decision rules for retaining factors sometimes provide conflicting information. For this reason, it is important to reiterate that factor analysts should use multiple criteria (Fabrigar et al., 1999; Tinsley & Tinsley, 1987) and should retain differing numbers of factors for exploratory purposes. The factor analyst working with the Figure 1 data might retain one, two, and three factors and see which solution is more interpretable. If some variables are associated with multiple factors (as described below), it is hard to determine what makes each factor unique; extracting one additional or one fewer factor may help. Similarly, if only one or two variables are associated with a given factor, then that factor is not well-enough defined to be retained. In general, it is better to extract too many factors than too few; inconsequential factors will not hold up later in the analysis (Comrey & Lee, 1992). If many factors are not interpretable, then it is best to retain one fewer factor and to repeat sequentially until an interpretable solution is found. This is permissible because there is no correct number of factors. Rather, EFA's goal is to give meaning to covariation among variables, and this means exploring the most useful number of factors. Once the user decides how many factors to retain, he or she repeats the analysis with the specified number of factors. It is important to emphasize that the iterative process of repeating the EFA provides the best opportunity for finding interpretable results.

Rotation and Interpretation of Factors

Structure coefficients. Once the user has decided on the number of factors to extract, he or she is typically interested in examining the relationships between variables and the retained factors. Factor *structure coefficients* reflect the correlation between variables and factors.[3] Structure coefficients are typically the basis for which factors are interpreted. To have clearly interpretable factors, it is best if structure coefficients are either very large (i.e., near 1 or –1) or close to 0; this is the idea behind simple structure.

Factor rotation. The initial EFA results typically do not produce structure coefficients that provide easy factor interpretation because the goal of explaining variance drives factor extraction, not the goal of providing interpretable factors. Thus, to attain more interpretable factors, the user typically applies a rotation of the factor axes. *Factor rotation* refers to changing the reference points for the variables. To use a simple example, consider two intelligence-test examiners. One examiner is interested in two pieces of information about examinees: (a) the verbal IQ and (b) the performance IQ. The second examiner is also interested in two pieces of information: (a) the full scale IQ and (b) the difference between verbal and performance IQs. At first glance, it seems that these examiners are focusing on different data, but in fact, both examiners are considering the same information (i.e., verbal and performance IQ, either singly or in combination). Like the two test examiners, the rotated solution provides a different conceptualization of the data from the unrotated solution. More concretely, a rotated factor solution provides a different set of structure coefficients than the unrotated solution. The analyst derives these new structure coefficients by multiplying the unrotated structure coefficients by a set of constants derived via trigonometric functions (see Gorsuch, 1983).

Returning to the EFA of the eight Self-Efficacy in Research Measure items, Table 3 displays the research tasks measured by each item and two different sets of structure coefficients. The unrotated structure coefficients (i.e., correlations) for a somewhat diverse set of items are fairly large for Factor I, suggesting that this factor is a measure of global research self-efficacy. Based on the unrotated coefficients, the second factor is more difficult to interpret because the structure coefficients are not universally large. Moreover, some coefficients are positive and some are negative, and this further complicates the interpretation of the second factor. The pattern of rotated structure coefficients is easier to interpret. After considering which research tasks are highly correlated with rotated Factor I and which are highly correlated with rotated Factor II, it is clear that rotated Factor I reflects self-efficacy with respect to writing, whereas rotated Factor II reflects self-efficacy with respect to quantitative and computer activities.

TABLE 3: Two Sample Cases of Structure Coefficients for Eight Research Self-Efficacy Item Scores From the Self-Efficacy in Research Measure

	Unrotated Coefficients		*Orthogonal Rotated Coefficients*	
Self-efficacy with respect to	*Factor I*	*Factor II*	*Factor I*	*Factor II*
2. Knowing which statistics to use	.70	.46	.20	.81
6. Manipulating data to get it onto a computer system	.63	.45	.16	.76
17. Avoiding the violation of statistical assumptions	.67	.28	.30	.66
19. Using simple statistics (e.g., *t* test, ANOVA, correlation, etc.)	.68	.45	.19	.79
7. Writing a discussion section for a thesis or dissertation	.72	–.17	.64	.37
13. Writing the introduction and literature review for a dissertation	.75	–.49	.89	.15
14. Reviewing the literature in an area of research interest	.69	–.37	.76	.20
15. Writing the introduction and discussion sections for a research paper for publication	.78	–.45	.88	.21

In this example, the unrotated structure coefficients do not provide information about factors that is as interpretable as do the rotated structure coefficients. Both sets of structure coefficients provide the same information about the relationships among the eight variables, however. To understand this potentially confusing notion, it is important to understand that structure coefficients indicate the correlation between any two variables via their correlation with the factors. For example, adding the products of the structure coefficients within each factor can produce the correlation between Items 2 and 6 (see Table 3). For the unrotated case, the structure coefficients for Factor I may be multiplied ($.70 \times .63$) and added to the product of the structure coefficients for Factor II ($.46 \times .45$), yielding a correlation coefficient of .65. For the rotated case, the structure coefficients are multiplied for Factor I ($.20 \times .16$) and added to the product of the structure coefficients for Factor II ($.81 \times .76$) to again yield a correlation coefficient of .65. Factor solutions are indeterminate, meaning there are infinite combinations of

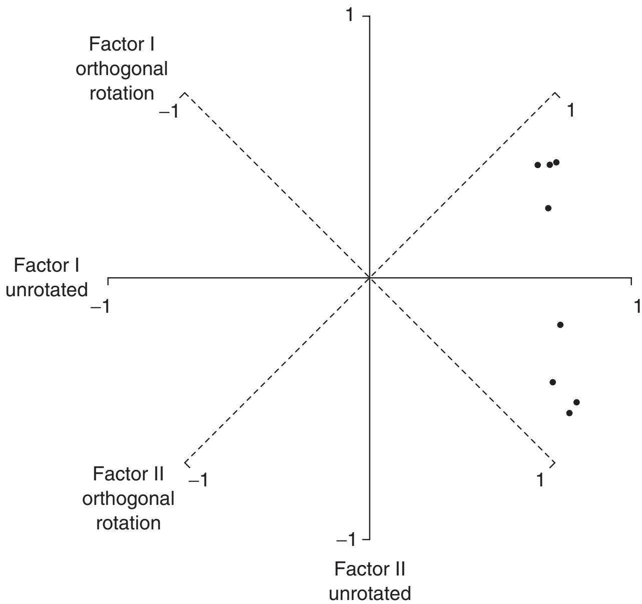

FIGURE 2. Visual Depiction of Unrotated and Orthogonal Rotated Structure Coefficients for Eight Items Measuring Research Self-Efficacy

NOTE: The solid lines are axes that correspond to the unrotated structure coefficients in Table 3. The dashed lines are axes that correspond to the orthogonal rotated structure coefficients in Table 3.

structure coefficients that explain the covariation among variables equally well from a statistical perspective. Because of this, the factor solution that yields the most interpretable factors is the desired one, which, in this instance, is the set of rotated coefficients.

Visualizing the relationship between factors and variables may facilitate understanding rotation. In Figure 2, dots represent the eight research self-efficacy items. These items are plotted along two dimensions corresponding to their unrotated structure coefficients. The horizontal solid line represents the unrotated Factor I, and all the variables have strong, positive relations with this factor. The vertical solid line represents unrotated Factor II; four of the variables have positive relations, and four of the variables have negative relations with this factor. Because these solid lines represent one of an infinite number of factor solutions that preserves the relations among variables, the dashed lines are also plausible. These dashed lines represent the rotated structure coefficients. Four of the items have strong, positive relations with rotated Factor I but weak relations with Factor II, and the other four items have strong, positive relations with rotated Factor II but weak relations with Factor I. Note that the distance between data points reflecting the items has not changed whether one adopts the unrotated or the rotated factors; thus, the relationships among items are preserved in both cases. The psychological meaning ascribed to these variables changes to constructs that more clearly differentiate the two factors, however.

Orthogonal rotation. In the example in Figure 2, the rotation of the axes preserved the orthogonality of the factors—that is, the factors remained uncorrelated. The most commonly used orthogonal rotation is the Varimax rotation. The Varimax rotation finds the pattern of structure coefficients that maximizes their variance; this is the simple-structure solution in which some structure coefficients are near zero and others are near ±1. Factors are kept orthogonal, and the solution produces the same correlations among variables as the unrotated structure coefficients. The Varimax rotation also results in eigenvalues of the rotated factors becoming more equal. For the factors in Table 3, the eigenvalues for unrotated Factors I and II are 3.97 and 1.29, respectively, but for the rotated factors, they are 2.72 and 2.54. In both cases, the total variance explained is the same (i.e., the sum of the eigenvalues is 5.26), but the variance is spread more equally across the two factors after the Varimax rotation. Thus, it is important for researchers to report eigenvalues before rotation and eigenvalues after rotation, as they will differ (Russell, 2002). It is also important to emphasize that the substantial difference in eigenvalues between the first and the second factors before rotation does not indicate that only one factor should be extracted, because after rotation, these eigenvalues even out.

Oblique rotation. Examining the dashed lines representing the orthogonal rotation in Figure 2 suggests that the axes could get nearer to the data points if they were not held at right angles. In Figure 3, the dashed axes are 60 degrees apart, not 90 degrees apart, and appear to do a better job of explaining the relations among the variables than does the orthogonal rotated solution. The rotation in Figure 3 does not preserve the orthogonality of the factors, because the closer the axes are to each other, the more highly the two factors correlate. An orthogonal rotation does not always provide simple structure, however, particularly if the two factors are correlated in the population. (For example, it is likely that writing and quantitative/computer self-efficacy are correlated among the population of psychology graduate students, as graduate training in psychology should increase both

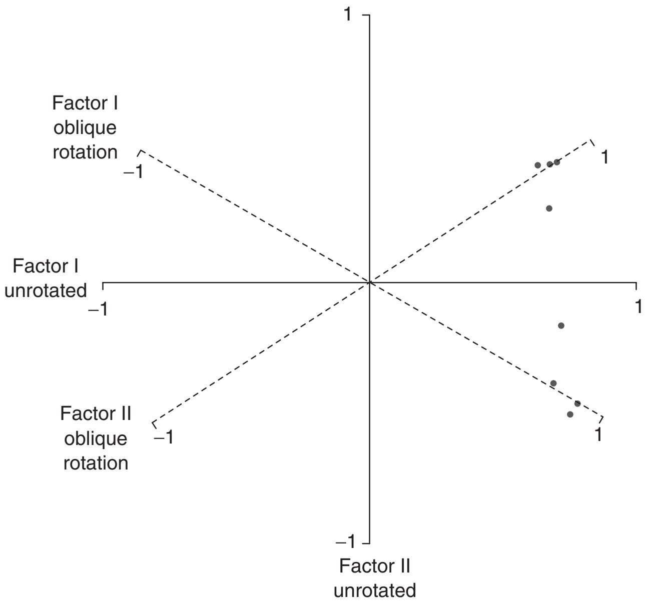

FIGURE 3. Visual Depiction of Unrotated Structure Coefficients and Oblique Rotated Pattern Coefficients for Eight Items Measuring Research Self-Efficacy
NOTE: The solid lines are axes that correspond to the unrotated structure coefficients in Table 3. The dashed lines are axes that correspond to the oblique rotated pattern coefficients in Table 4.

aspects of self-efficacy.) The type of rotation in Figure 3 that allows for factor correlation is known as an oblique rotation. In this case, the oblique rotation is appropriate to facilitate factor interpretation.

There are several methods of doing an oblique rotation. The Direct Oblimin rotation can produce either high or low correlations among factors, depending on how the analyst specifies the rotation. A Direct Oblimin rotation with the research self-efficacy items using the default SPSS method shows the same two factors (i.e., writing and quantitative/computer self-efficacy) as the orthogonal rotation, but when allowed, factors correlate at $r = .49$. Perhaps a better choice for oblique rotation is Promax (Fabrigar et al., 1999; Russell, 2002). The Promax rotation starts with an orthogonal rotation and then finesses the solution with an oblique rotation. If the factors are in fact uncorrelated, then the rotation will remain orthogonal, but if the factors have some correlation, then the rotation will be oblique.

A Promax rotation also indicated a .49 correlation between the two research self-efficacy factors.

Pattern coefficients. After an oblique rotation, two types of information indicate the relationship between the factors and the variables. Structure coefficients become universally large to the extent that the factors are highly correlated (see Table 4). It is therefore valuable to examine the relationship between variables and each factor that accounts for the factor intercorrelation; these relationships are indicated by the *pattern coefficients*. The pattern coefficient reflects the relationship between variables and factors that controls for the other factor(s). Alternatively, pattern coefficients are unique weights given to each factor in explaining an individual's score on a variable, which are similar to weights in multiple regression or discriminant analysis (see Sherry, 2006 [this issue]. Considering Item 2 in Table 4 as an example, the structure coefficient for Factor I is .40, but the pattern coefficient—which controls for the relationship between the two factors—for Factor I is only –.01. One would therefore describe the correlation coefficient between responses to Item 2 and Factor I as .40, but the relationship between Item 2 and Factor I is nearly zero when controlling for Factor II. It therefore seems that self-efficacy with respect to knowing which statistics to use correlates with Factor I at .40 only because Factor I and Factor II are correlated. This is evident because controlling for the correlation between factors makes the relationship between Item 2 and Factor I close to zero. When factors are uncorrelated, as in the orthogonal rotation, the pattern coefficient is equal to the structure coefficient.[4] When factors are correlated, however, the pattern and structure coefficients are not likely to agree. This is analogous to multiple regression where beta weights generally do not equal the correlation coefficient between a predictor and a criterion, because the beta weights reflect the relationship between a predictor and the criterion while controlling for other predictors.

Choice of rotation. Given the different types of rotations available, it can be difficult to know which is the correct rotation. Although different rotations may provide slightly different results, the differences are usually not dramatic. An analyst might base his or her decision about which rotation to use on extant theory related to the variables being analyzed. If uncorrelated factors are theoretically defensible, then the rotation should be orthogonal, and if correlated factors are expected, then an oblique rotation is indicated. Analysts may also base the rotation choice on the interpretability of the factors. Another alternative is simply to use the Promax rotation, as Russell (2002) suggests; if factors are uncorrelated, they will remain that way after rotation, but if factors are correlated, the Promax rotation will reflect that as well.

TABLE 4: Structure and Pattern Coefficients for Eight Research Self-Efficacy Item Scores From the Self-Efficacy in Research Measure After a Promax Rotation

	Structure Coefficients		*Pattern Coefficients*	
Self-efficacy with respect to	*Factor I*	*Factor II*	*Factor I*	*Factor II*
2. Knowing which statistics to use	.40	.84	–.01	.84
6. Manipulating data to get it onto a computer system	.34	.77	–.05	.80
17. Avoiding the violation of statistical assumptions	.45	.71	.13	.65
19. Using simple statistics (e.g., *t* test, ANOVA, correlation, etc.)	.38	.81	–.03	.83
7. Writing a discussion section for a thesis or dissertation	.71	.52	.61	.22
13. Writing the introduction and literature review for a dissertation	.90	.37	.94	–.09
14. Reviewing the literature in an area of research interest	.78	.39	.78	.01
15. Writing the introduction and discussion sections for a research paper for publication	.90	.42	.91	–.02

Once a rotation has yielded a factor solution that is close to simple structure, the factor analyst's goal is to name factors. Pattern coefficients are not useful in naming factors unless the factors are already known (Gorsuch, 1983); the structure coefficients are more appropriate for naming factors. The process of naming the factors involves determining which variables have strong structure coefficients (either positive or negative) for the factor and considering what those variables have in common. More weight should be given to stronger structure coefficients when naming factors. The presence of variables with high structure coefficients on more than one factor indicates that the factors are not distinct; it may therefore be useful to try to retain a different number of factors to eliminate variables with double loadings. If the factors cannot be easily named, it might be best to try a different rotation, add to or subtract from the number of factors being extracted, or reexamine the usefulness of the set of variables being analyzed.

Oftentimes, EFA results indicate that one or more variables should be omitted. For example, analysts frequently use EFA to eliminate irrelevant items during the instrument-development process. If variable elimination is a goal, then a variable with low structure coefficients on all factors is a candidate for elimination. Analysts often use a structure coefficient of .40 as a minimum for retaining a variable, but when it is of interest to retain as many variables as possible, one might retain variables with structure coefficients of at least .30. Using a higher cutoff, such as .50, would eliminate more variables and would yield purer measures of the factors.[5] This is because each factor would be marked only by those relatively few variables that are strongly correlated with the factor.

Sample-Size Recommendations for EFA

The assumption made with EFA is that sampling error is nonexistent. This means that in EFA, small samples can lead to biased estimates, so attaining a large sample is critical. The research literature has provided several rules of thumb for sample size. For example, Comrey and Lee (1992) suggested that 100 cases are poor, 200 are fair, 300 are good, and 500 or more are very good. Other sample-size recommendations are based on the number of variables being analyzed. For example, Gorsuch (1983) suggested at least 5 cases per variable. Although these are well-known rules of thumb, simulation studies do not support the importance of considering the number of variables when deciding on an appropriate sample size (Guadagnoli & Velicer, 1988; MacCallum, Widaman, Zhang, & Hong, 1999). Moreover, blanket sample-size recommendations tend to overlook important characteristics of the data, such as the communalities of the variables and the magnitude of the structure coefficients in a given data set (Fabrigar et al., 1999; Guadagnoli & Velicer, 1988; Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005; MacCallum et al., 1999).

Sample-size decisions should therefore be driven not by the number of variables being analyzed but by the degree to which (a) factors are overdetermined (i.e., at least three or four variables have high structure coefficients for each factor) and (b) communalities are high (i.e., a high percentage of common variance among variables). A small sample size (n = 100) may be appropriate if structure coefficients are uniformly high, whereas cases where structure coefficients are much lower may require in excess of 300 cases (Guadagnoli & Velicer, 1988). MacCallum et al.'s (1999) Monte Carlo study indicated that samples as low as 60 can accurately reflect the population parameters if communalities are consistently high (i.e., .60 or greater). Communalities in the range of .50 required a sample size from 100 to 200, and low communalities required samples of 300 or more to converge on the appropriate solution. Users may not be able to predict the magnitude of structure coefficients and the communalities before collecting the data. For this reason, users are encouraged to have samples of at least 300 cases to be safe, but fewer cases may be appropriate in some instances.

Reporting EFA Results

Accurately interpreting and evaluating EFA results requires knowledge of the specific procedures used. Comrey and Lee (1992) and Tinsley and Tinsley (1987) provided useful guidelines for reporting EFA results. First, the analyst needs to describe the sample in detail, as EFA results can depend on the characteristics of the sample. For example, research self-efficacy factors among doctoral students may be different from those among master's students because of different training emphases. Second, information about the variables is required, including a description of the variables as well as their means and standard deviations. The analyst must indicate the types of correlations (e.g., Pearson, biserial) analyzed. Third, the analyst must specify the factor-extraction method (e.g., PAF, PCA) along with the initial communality estimates used (e.g., squared multiple correlations). Fourth, one must report the criteria for retaining factors, such as Kaiser's criterion, the scree test, or a parallel analysis. Fifth, it is useful to report initial eigenvalues and the percentage of variance each factor explains. Sixth, the analyst must specify the types of rotation applied to the data—including those attempted but not reported as the final factor solution—and should report the factor eigenvalues and item communalities based on the rotated factors. Finally, structure coefficients—and in the case of oblique rotations, pattern coefficients—are ideally reported, including those coefficients that are less than any cutoff value.

CFA

CFA tests theoretically derived hypotheses about how many factors exist and which variables correspond to which factors. As the names indicate, analysts typically use EFA to explore possible factors that may explain covariation among variables, whereas they use CFA to confirm that a hypothesized factor structure provides a good fit to the data. Hypotheses may come from a strong theory postulating relationships among constructs and observed variables, or hypotheses may be derived from EFA results on prior samples. Regardless of the hypotheses' source, that CFA is a hypothesis-testing procedure makes it markedly different from EFA.

Several types of hypotheses are typically made in a CFA. First, the factor analyst tests whether the hypothesized number of factors underlies the data. For example, does a set of eight research tasks comprise two dimensions of research self-efficacy? Second, the factor analyst hypothesizes about which pattern coefficients are zero and which are nonzero (e.g., is self-efficacy with respect to knowing which statistics to use associated with quantitative/computer self-efficacy?). Third, analysts may also hypothesize about which correlations between factors are nonzero (e.g., is writing self-efficacy correlated with quantitative/computer self-efficacy?). Additional hypotheses are possible concerning parameters such as correlations among unique variances of the measured variables; describing these hypotheses is beyond the scope of this article (for more information, see Bollen, 1989; Kline, 1998). Along with information about the significance of individual parameters such as pattern coefficients and factor intercorrelations, CFA provides overall goodness-of-fit tests of the match between the theoretical factor structure and the data. Thus, analysts can test individual parameters and the model as a whole.

The hypotheses tested by CFA are best described visually via figures illustrating the relationships among variables and factors. Figure 4 shows hypotheses about the relationships between two research self-efficacy factors and eight research self-efficacy items. By convention, circles or ovals represent factors, and squares or rectangles represent measured variables. The two research self-efficacy factors that corresponded to the factors found in the EFA are described as quantitative/computer self-efficacy and writing self-efficacy. Additional factors representing unique variance (customarily abbreviated *E* to denote error) are associated with each measured variable. The curved, double-headed arrow between the two research self-efficacy factors indicates that these factors are hypothesized to be correlated (corresponding to the oblique rotation in EFA). In other words, it suggests that writing self-efficacy and quantitative/computer self-efficacy are related aspects of research self-efficacy. Straight arrows between factors and variables represent hypotheses that a pattern coefficient is significantly different from zero—that is, these arrows suggest that a given variable is associated with a given factor. Factors and variables not directly connected by an arrow are hypothesized to have pattern coefficients close enough to zero that they might as well be considered zero (e.g., Item 2 is hypothesized *not* to be a function of one's writing self-efficacy). Of course, it is not likely that pattern coefficients are exactly zero, but to achieve parsimony, any small pattern coefficients are not part of the model.

Hypotheses are operationalized by fixing and freeing parameters. If a parameter is specified to be free, it is assumed to be nonzero. For example, Figure 4 illustrates a free correlation between the two research self-efficacy factors (as indicated by the double-headed arrow); free pattern coefficients between the quantitative/computer self-efficacy factor and Items 2, 6, 17,

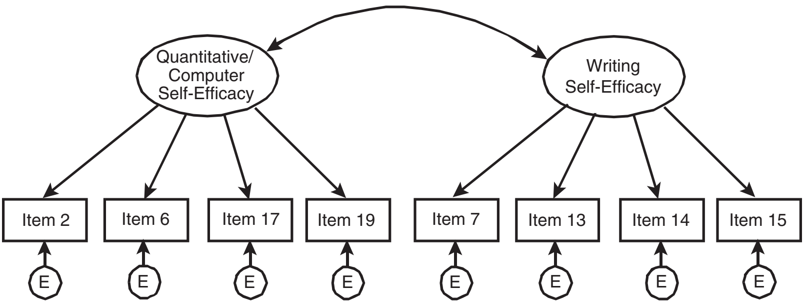

FIGURE 4. Confirmatory Factor Analysis Model of Eight Items Measuring Research Self-Efficacy

and 19; and every other place an arrow exists. On the other hand, if a parameter is specified to be fixed, the researcher specifies its value. Typically, fixed parameters are set to 0; for example, the pattern coefficients between Items 2, 6, 17, and 19 and the writing self-efficacy factor are hypothesized to be 0. This pattern of fixing and freeing pattern coefficients is akin to specifying simple structure whereby pattern coefficients are either large or 0, so CFA is a stringent test of a theory. It is important to reiterate that fixing a parameter to 0 does not imply that the factor analyst believes the parameter is 0 in the population. Rather, it simply indicates that the parameter is believed to be so small that it is more parsimonious not to consider the parameter in the model.

Planning the CFA Study

Sample size. As mentioned above, EFA analyzes a correlation matrix, but analyzing CFA with correlations can lead to improper results (Cudeck, 1989). Thus, the unit of analysis in CFA is typically the observed variance-covariance matrix among variables. The statistical theory underlying CFA assumes that elements in this sample covariance matrix are equal to the population values. For this reason, a large sample is necessary for proper estimation in CFA. Whereas early thinking was that a fixed number of cases per estimated parameter was necessary (e.g., Bentler, 1990), recent research shows that the number of parameters does not appear to be as important as overall sample size (Jackson, 2001). A sample of fewer than 100 cases may be all that is necessary to use CFA if the average pattern coefficient is high (Russell, 2002), but generally sample sizes will need to exceed 100. Most models will require well over 100 cases, so it is best to err on the side of too large a sample than too small a sample.

The necessary sample size for a CFA study can be determined a priori. MacCallum, Browne, and Sugawara (1996) provide a framework for estimating statistical power in CFA. In this framework, the necessary sample size is largely a function of the model's degrees of freedom. In other words, the number of fixed parameters in a model largely determines the necessary sample size, whereby models with more constraints require smaller samples. It is advisable for users to conduct a power analysis using MacCallum et al.'s method prior to conducting the study to ensure an adequately large sample.

Software choice. CFA requires specialized software due to the complexity of the analyses. SPSS, which is popular for EFA, is unable to run a CFA.[6] On the other hand, CFA can be done in SAS software using the PROC CALIS procedure. Nevertheless, most CFAs reported in the literature are conducted using software such as LISREL (Jöreskog & Sörbom, 1996), EQS (Bentler, 1995), and Mplus (Muthén & Muthén, 1998). There are few substantial differences among these programs; each has its strengths and limitations. LISREL historically has required complex syntax, which is imposing for most users, but a point-and-click interface has been added in recent software revisions. Moreover, much of the CFA literature uses LISREL terminology for describing parameters. EQS has an interesting interface that allows model generation via a diagram. Mplus has a more straightforward syntax language than LISREL and allows the user to conduct a host of other multivariate analyses. All three programs will yield comparable results for a given CFA, so the differences are a matter of taste and not substance.

Identification. Prior to estimating a factor model, it is necessary to confirm the identification of the model. A hypothesized model is identified if only one numerical value is possible for each estimated parameter. A lack of identification is therefore problematic because a number of different estimates are possible for each parameter. Consider, for example, a case with two test items to which an examinee responds on a Likert-type scale: "I feel sad" and "I feel blue." Each of these items is hypothesized to be associated with one latent factor representing depressed mood. Although this model seems reasonable at first glance, it is not identified because, given the amount of data, there are too many parameters being estimated to provide a solution. Specifically, four parameters are estimated: two pattern coefficients (i.e., the relationship between each item and the depressed-mood factor) and two error variances (i.e., variance in each item not overlapping with the other item).[7] Although there are four unknown parameters, there are only three unique elements in the observed covariance matrix for these two variables: two item variances and one covariance between them. Like a game of musical chairs, there is not enough information to accommodate all the values. Methods of demonstrating the identification of a model go beyond the scope of this article (see Bollen, 1989; Kline, 1998). Understanding identification issues is important for a CFA user, however, because software programs that perform CFA do not regularly test for model identification. It is therefore up to the user to confirm the identification of the hypothesized model. Estimating an unidentified model may result in the software not converging on a solution or, even worse, it converging on a false solution.

Model Estimation

The goal of estimation in CFA is to find values of parameters that reproduce the observed covariance matrix as closely as possible. Many of the factor-extraction methods common to EFA are not used for CFA. Instead, the most common method of estimating parameters is the ML method. ML's goal is to find parameter estimates that if they were the true population values would maximize the likelihood that the observed covariance matrix was drawn from that population. ML estimation is an iterative process, meaning that improvements to the final solution are based on tentative estimates. When the iterations finish, the software converges on a solution that provides estimates for each parameter.

The theory underlying the ML estimation method assumes that data distributions are multivariate normal, but this is often not the case among data sets (Micceri, 1989). For example, one would expect many measures of psychiatric symptoms to be highly skewed, particularly if administered to a nonclinical sample. Given the degree of nonnormality among a given data set, the ML method might be inappropriate. Alternative estimation procedures allow for nonnormality among the data. For example, Satorra and Bentler (1994) proposed a robust ML estimation procedure that adjusts measures of overall fit and standard errors of parameter estimates. Non-ML methods of estimating CFA models are also available, which do not assume multivariate normality (e.g., Browne, 1984).

Under the assumption of multivariate normality, the parameter estimates generated by the ML method allow for testing hypotheses about the statistical significance of parameters. CFA results provide unstandardized estimates of parameters along with standard errors. Typically, the unstandardized estimate is divided by the standard error, and quotients stronger than ± 2 (which is the approximate 95% confidence interval about 0) are considered statistically significant. To illustrate testing the significance of parameters, I used the Mplus program's ML procedure to perform a CFA on the research self-efficacy items in Figure 4. The unstandardized pattern coefficient for the research self-efficacy Item 2 was 1.07, and the standard error was .04, yielding a quotient of 25.07; this quotient indicates that the pattern coefficient was statistically significant. Bear in mind that given the large samples typically required for CFA analysis, many pattern coefficients that are only moderate in strength will be statistically significant, so it is important to examine the standardized pattern coefficient. The standardized pattern coefficient for Item 2 was .84, which is fairly large. Thus, whereas EFA provides only information regarding the standardized pattern coefficient, CFA provides information regarding the unstandardized coefficient as well as its statistical significance.

Evaluating Model Fit

After attaining a factor solution and estimating parameters, the factor analyst will also be interested in how well the overall model fits the data. The logic of model fit is as follows: When the number of estimated parameters equals the number of elements in the observed covariance matrix, the model will fit the data perfectly because as much information is known in the covariance matrix as is estimated by the model. When constraints (i.e., fixed parameters) are placed in the model, the observed covariance matrix may not be reproduced exactly, because the analyst uses a smaller number of parameters to reproduce a larger amount of information about the true relationships among variables (Bollen, 1989; Kline, 1998). This is the same idea of parsimony described with EFA; a good CFA model simplifies the complex relationships among variables without losing too much information about those relationships. CFA's goal is therefore to find a model with several constraints that still provides a good fit to the data. If there are too many constraints, the fit is likely to be poor; but if there are too few constraints, the model is overly complex.

Model fit, then, can be reflected by comparing the actual covariances among variables with the covariances implied by the parameter estimates (Bollen, 1989). If the parameter estimates can reproduce the covariances among variables with a minimum of error, then the hypothesized model provides a good fit to the data. The chi-square statistic indicates the difference between the model-reproduced covariances and the actual covariances. The larger the difference between the actual and the reproduced covariances, the higher the chi-square statistic. A chi-square high enough to be statistically significant indicates a significant difference between the hypothesized model and the data. Thus, a significant p value indicates poor model fit. Despite its intuitive appeal, the use of chi-square in measuring overall fit is problematic because the sample size is part of the chi-square computation. Specifically, the higher the sample size, the higher the chi-square statistic (Hu & Bentler, 1995; Kline, 1998). Because large samples are necessary for CFA, there is a high likelihood of the chi-square being significant even when the model provides a good fit to the data. For this reason, using chi-square as a measure of model fit is contraindicated. The same may be said about using the ratio of chi-square to degrees of freedom as an indicator of model fit. A low ratio of chi-square to degrees of freedom is sometimes viewed as an indicator of good model fit. Sample size maintains its influence on chi-square but does not affect the degrees of freedom, which are based on the number of fixed model parameters (Marsh, Balla, & McDonald, 1988). Thus, given constant degrees of freedom, there is a positive relationship between this ratio and the sample size, leading to some good-fitting model appearing to have a poor fit solely because they were tested on a large sample.

There are, however, fit indices that sample size does not adversely impact. Of the many fit indices that have been proposed, I describe four here as being useful in rejecting incorrectly specified models. Two of these indices estimate the improvement in fit in the hypothesized model over a null model in which all variables are uncorrelated: the Tucker-Lewis Index (TLI; also known as the Nonnormed Fit Index) and the Comparative Fit Index (CFI). Ideally, the hypothesized model will provide a far superior fit to the data than will a null model, and the TLI and CFI assess this improvement. The TLI and CFI are conceptually similar but make different types of corrections for sample size and model complexity (Bentler, 1990; Hu & Bentler, 1995). Both of these statistics are bound between 0 and 1, and Monte Carlo research suggests that values of .95 or higher indicate good model fit (Hu & Bentler, 1999).

Two other indices reflect the difference between the actual and the reproduced covariance matrices. The Standardized Root Mean-Squared Residual (SRMR) is the average of the standardized residuals between elements in the observed covariance matrix and those in the model-implied covariance matrix (Jöreskog & Sörbom, 1996). The Root Mean Square Error of Approximation (RMSEA) is a measure of fit between the actual covariance matrix and the model-implied covariance matrix adjusting for model complexity (Browne & Cudeck, 1993). Values of .08 or lower for the SRMR and .06 or lower for the RMSEA suggest good model fit. The TLI, CFI, SRMR, RMSEA, and their corresponding cut scores are among those Hu and Bentler (1999) recommended because they result in low rates of Type II error. The RMSEA has the added advantage of providing confidence intervals; thus, one may conduct a test of not close fit whereby a significant result indicates a good-fitting model (MacCallum et al., 1996).

Although researchers can perform significance tests for the RMSEA, the other fit statistics do not have known underlying distributions, so true significance tests are not possible. Rather, these empirically derived cut scores based on Hu and Bentler's simulation studies are applied, although it should be noted that the application of these cut scores is not universally accepted (e.g., Marsh, Hau, & Wen, 2004).

Returning to the research self-efficacy model in Figure 4, the chi-square statistic was 210.90 with 19 degrees of freedom, which is statistically significant ($p < .001$). The sample size was 937, and this likely inflated the chi-square. Turning toward the other fit statistics, the CFI was .96, and the SRMR was .06; both indicated good model fit. The TLI was .94, which is just below the cutoff for good model fit, and the RMSEA was .10, which is well above the cutoff for good model fit. Thus, there is mixed evidence about how well the model fits the data. This is often the case with CFA, and because there is no best number that reflects model fit, it is up to the factor analyst to draw conclusions about the adequacy of model fit.

Additional Benefits of CFA

CFA allows for even more complex hypotheses to be tested. CFA allows for direct comparisons between different hypothesized factor structures, and factor structures can be compared across different groups. Analysts can also use CFA to test hierarchical, or second-order, factor structures.

Comparing alternative models. Comparing alternative models is appropriate for CFA, just as it is for structural equation modeling (SEM) more generally (Martens, 2005; Martens & Hasse, in press; Quintana & Maxwell, 1999; Weston & Gore, 2006 [this issue]). Analysts often use the chi-square difference test to compare the difference in fit of two nested models. Models are nested if a more restrictive model can be obtained by imposing constraints on the more general model (Bollen, 1989). For example, if the model in Figure 4 specified that the two research self-efficacy factors were uncorrelated, this new model would be nested within the model that did specify the correlation. Likewise, if the hypothesis was that only one factor explained the covariation among the eight items, this alternative model would be nested within the hypothesized model (because specifying one factor is akin to fixing a correlation between two factors to 1).

Table 5 displays fit indices for these alternative models. Examining the chi-square statistics indicates a lower chi-square for the hypothesized model than for Alternative A, the more restrictive model specifying two uncorrelated factors. A more restrictive model will never provide a better fit than the less restrictive model—that is, the chi-square will be at least as

TABLE 5: Fit Statistics for Confirmatory Factor Analyses of Eight Research Self-Efficacy Item Scores From the Self-Efficacy in Research Measure

Model	χ^2	df	*TLI*	*CFI*	*RMSEA*	*SRMR*
Hypothesized model	210.90	19	.94	.96	.10	.06
Alternative A: Two uncorrelated factors	406.70	20	.88	.91	.14	.23
Alternative B: One factor	1440.11	20	.55	.68	.28	.17

NOTE: CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean-Squared Residual; TLI = Tucker-Lewis Index.

large in the more restrictive model as it is in the less restrictive model (Bollen, 1989). If the chi-square values are not significantly different from each other, then one opts for the more restrictive model because it is more parsimonious. The difference itself between two chi-square values is distributed as a chi-square statistic, and the difference between the degrees of freedom for two models provides a new degree-of-freedom value for the comparison test. To illustrate, the chi-square difference between the hypothesized model and the Alternative Model A, which specified two uncorrelated factors, is 195.80 and, with 1 degree of freedom, is statistically significant. Thus, a model specifying two uncorrelated factors provided a significantly worse fit to the data than did the hypothesized model specifying two correlated factors. Likewise, Alternative Model B, specifying one factor, provided a significantly worse fit to the data than did the hypothesized model, $\Delta\chi^2 = 1229.21$, $p < .001$. Thus, the hypothesized model is preferred among the three examined.

Three points about comparing models are in order. First, the chi-square difference test is applicable only to nested models. Each alternative model in the research self-efficacy example (Table 5) is nested relative to the hypothesized model, but the two alternative models are not nested relative to each other, because they are equally restrictive. The factor analyst needs to confirm that models are nested before conducting the chi-square difference test. The analyst may use alternative measures of fit, such as Akaike's (1987) information criterion, to compare nonnested models. Second, the chi-square difference test is subject to the same inflation due to large sample sizes as the chi-square statistic for a single model. For this reason, the chi-square difference test is likely to yield significant differences in model fit even when the practical difference is not very large. Third, it is tempting to modify a poor-fitting model by estimating new parameters to lower the chi-square. Although post hoc model modification can be useful to understand one's data, modifications based on the reduction in chi-square are data-based, not theory-based, changes. As such, post hoc modifications need to be theoretically defensible and clearly indicated (Kahn, 2005; Thompson, 2004).

Multiple-groups analysis. Another way of comparing factor models is to determine whether a given factor structure fits data from different groups equally well. This is known as a *multiple-groups* (or sometimes *multisample*) analysis. CFA provides a straightforward way of comparing the adequacy of a factor structure across groups defined by demographic variables (such as race, ethnicity, or sex), treatment condition, or some individual difference. Comparing the fit of a factor model between groups involves estimating a model in which some parameters are fixed to be equivalent between groups. The analyst then compares this model to a less restrictive one in which the same parameters are not fixed to be equivalent (i.e., they are free to vary between groups). If the less restrictive model provides a better fit to the data, then there is evidence that the factor structure differs between groups.

Oh and Neville (2004) provided a recent example of a multiple-groups analysis. They developed the Korean Rape Myth Acceptance Scale to measure the degree to which Koreans accepted rape myths. As part of the CFA, Oh and Neville assessed whether the factor structure provided as good a fit to men as to women. They estimated five models, including a model in which all parameters were free to vary between groups, a model in which all parameters were free to vary except for pattern coefficients, and at the most restrictive, a model in which all model parameters were held invariant. Fit indices indicated approximately equally good fit across the five models, suggesting a preference for the most restrictive model (i.e., the most parsimonious model). Oh and Neville's analyses therefore demonstrated that the factor structure of the Korean Rape Myth Acceptance Scale held for both men and women, thereby providing important validity evidence for the test scores.

Second-order factor analysis. Among some data, it is possible that three or more factors are highly correlated, for example, with correlations above .50. In these instances, it behooves researchers to consider the possibility that a second-order factor explains the high correlations among the factors. A *second-order factor* is a superordinate factor that explains covariation among first-order factors (i.e., factors directly related to the measured variables). Cutrona and Russell (1987) provided an example of how a single second-order factor reflecting global perceived social support can explain covariation among six first-order factors of perceived social support:

attachment, social integration, reassurance of worth, reliable alliance, guidance, and opportunity for nurturance. They conducted a CFA that simultaneously assessed the fit of the six first-order factors to items of the Social Provisions Scale and the fit of the second-order factor to the first-order factors. This second-order model provided an approximately equal fit to the data as did a model with only six correlated first-order factors. Because a single second-order factor (with six second-order pattern coefficients) is more parsimonious than six correlated first-order factors (with 15 factor intercorrelations), the second-order factor analysis attained a more parsimonious factor structure. This analysis provided theoretical clarification regarding how global perceived social support related to more specific provisions of social support. It is worth noting that researchers may also conduct a second-order factor analysis within the context of EFA; one merely factor analyzes correlations among factors (Gorsuch, 1983).

APPLICATIONS TO COUNSELING PSYCHOLOGY RESEARCH, TRAINING, AND PRACTICE

Both EFA and CFA have aided psychologists' work with test development and validation, and factor analysis will likely continue to be a valuable psychometric tool. Comrey (1988), Floyd and Widaman (1995), and Reise, Waller, and Comrey (2000) provided useful tutorials on using factor analysis in instrument development. Factor analysis can aid counseling psychologists in other ways, however, and these nonpsychometric applications have not been fully realized. In this section, I present some less obvious ways that counseling psychologists may apply factor analysis by focusing on how it can advance psychological theory, enhance graduate training, and inform practitioners' work.

Factor Analysis and Theory-Driven Research

A lack of theory-driven research has been a problem in counseling psychology (Karr & Larson, 2005). CFA represents a valuable method to test theory-driven hypotheses. Researchers often use CFA to test theory in the context of SEM, of which CFA is a part. Specifically, SEM is a theory-testing procedure assessing relationships among latent variables (Martens, 2005; Martens & Hasse, in press [*TCP* speical issue, part 2]; Quintana & Maxwell, 1999; Weston & Gore, 2006 [this issue]). The latent variables are formed through CFA, and these latent variables are typically regressed on one another to form the model. Given the increased use of SEM in counseling psychology research over the past 5 to 10 years

(Martens, 2005), it appears that CFA is already playing an important role in testing theory.

There are ways to use CFA to test theory without doing a full structural equation model, however. CFA can test theoretical propositions about the interrelatedness and the distinctiveness of constructs. Lent, Brown, and Gore (1997) provided an example of this application, assessing whether academic self-efficacy and self-concept are empirically distinct and whether global and specific beliefs are distinct, as a way to clarify constructs from social-cognitive theory. They used CFA to estimate five a priori models that differed in the number of factors underlying the self-efficacy and self-concept variables. Their conclusion that global and specific self-efficacy beliefs are distinguishable from one another as well as from global measures of academic self-concept advanced social-cognitive theory by supporting the distinctiveness of these theoretical constructs.

CFA can also test the theoretical structure of a construct. In the counseling psychology literature, this has been successfully applied to the study of structural models of vocational interests (Rounds, Tracey, & Hubert, 1992). For example, one may use CFA to test the structure of Holland's (1997) hexagonal model of interests. Doing so requires imposing equality constraints on correlations between interest types that are equidistant across the hexagon. If a model with equality constraints provides a good fit to the data, then the spatial relations represented by the model are supported. Researchers may also test alternative interest structures among the same data set and compare the fit of these alternative models (Rounds et al., 1992). In this way, CFA has an advantage over multidimensional scaling, a similar multivariate procedure with the goal of spatially representing dimensions underlying a set of variables (Fitzgerald & Hubert, 1987). Unlike multidimensional scaling, CFA provides several goodness-of-fit tests for the model and provides a straightforward method for comparing alternative models statistically.

Finally, a meta-analytic CFA may be conducted as a means to understand the structure of variables as determined by an aggregate of studies (see Furlow & Beretvas, 2005). This process involves two steps. The analyst first pools correlations among variables via standard meta-analytic procedures (see Henson, 2006 [this issue]; Quintana & Minami, in press [*TCP* special issue, part 2). Then a CFA is conducted on the pooled correlation matrix. Smith, Mar, and Turoff (1998) provided an example of this procedure as applied to understanding the structure of symptoms of schizophrenia. They pooled data from 28 samples and compared four theory-based factor structures, ultimately finding a three-factor model to provide the best fit to the data. Meta-analytic CFA has great potential as a means to test theories in a way that removes biases inherent in any one sample.

Factor Analysis and Graduate Training

Factor analysis can facilitate graduate training by providing focus to training goals. Many recently designed instruments can assess the research training of doctoral students, and factor analyses of these instruments can identify core issues in training. For example, Schlosser and Gelso (2001) used EFA to identify three factors that underlie doctoral students' perspectives on their relationship with their graduate advisor: rapport, apprenticeship, and identification-individuation. Kahn and Gelso (1997) used a combination of EFA and CFA to determine that aspects of the research training environment may be grouped by instructional and interpersonal elements. Forester et al. (2004) conducted an EFA on 107 items representing three measures of research self-efficacy and found that four research self-efficacy factors—data analysis, research integration, data collection, and technical writing—explain the relations among the items. These empirical investigations may be used to guide program evaluations, whereby the factors might be used as key outcome variables to assess training efficacy. Also, factor-analytic results could indicate areas where a curriculum might need to focus to enhance student training. For example, training in research methods might focus on the four areas Forester et al. identified.

The practitioner training of graduate students may also be enhanced through factor-analytic work. Factor analyzing supervisor evaluations of trainees can provide structure to practicum training by highlighting core areas of competence that need to be addressed. Trainee ratings of their own counseling competence, if factor analyzed, can provide structure for individual supervision sessions by identifying similar groups of competencies from the trainee's point of view. As an example, Sodowsky, Taffe, Gutkin, and Wise (1994) used EFA and CFA to determine that four factors underlie responses to a measure of multicultural counseling competence: multicultural counseling skills, multicultural awareness, multicultural counseling relationship, and multicultural counseling knowledge. Their findings suggest that training in multicultural competence should focus not only on knowledge, skills, and awareness but also on strengthening the interaction between counselor and client. Educators are encouraged to factor analyze the unique set of competencies addressed by their training program to develop a core set of training goals specific to that program.

Factor Analysis Applied to Practice

Factor analysis can inform practitioners' work by identifying common threads among a group of clients and by identifying themes within a given client. A treatment provider who uses a regular intake instrument, such as a symptom checklist, may learn much about his or her clientele by identifying factors that underlie the symptoms or problems endorsed by the clients. This application is illustrated by Bauman and Lenox's (2000) factor analyses on a 72-item intake symptom checklist used in a college counseling center. Both an EFA and a CFA indicated that five factors explained responses to these items: psychological disorders, interpersonal concerns, family, career/academic, and acting out. This aids practitioners' work by essentially providing subscales for the intake questionnaire, thereby facilitating the identification of themes among a client's many symptoms.

Factor analysis may also help treatment providers evaluate and enhance the effectiveness of their treatments. Counseling satisfaction surveys are often given to clients after termination, and client ratings on many of these instruments are multidimensional (Lebow, 1983). Identifying the factors underlying a client's satisfaction with treatment can provide ideas about improving treatment. Thus, treatment providers who conduct regular outcome assessments—whether they focus on satisfaction, reduction of distress, or some combination—are encouraged to factor analyze their battery to reduce the data to a form that can guide practice. Using factor analysis to identify common factors among client data that are routinely gathered by a given agency will provide information that is easily translated into practice.

Third, counseling psychologists may use multilevel CFA to analyze data from individuals nested within groups. For example, counseling psychologists working as consultants in schools may be interested in the factor structure of variables from students within classrooms, and psychologists working in organizations may be interested in the factor structure of variables from employees in different work settings. Because multiple classrooms or work settings may be represented among the data, it is inappropriate to analyze individual-level data without taking into account possible group-level differences. Multilevel CFA addresses this issue by analyzing the relations among variables at the individual as well as the group level (Mehta & Neale, 2005). In other words, multilevel CFA assesses the organization of variables between groups (i.e., between classrooms) and within groups (i.e., among students within a classroom). Johnson, Burlingame, Olsen, Davies, and Gleave (2005) recently provided an example of applying multilevel analyses to the study of therapeutic relationship factors in group psychotherapy. Understanding the organization of variables within groups and between groups may therefore benefit counseling psychologists working in therapy, consultation, and community outreach.

SUMMARY AND CONCLUSIONS

In summary, factor analysis is a way of understanding what underlying constructs explain a set of variables. Its frequent use has helped to improve the science on which counseling psychologists base their work. Nevertheless, because EFA and CFA procedures are complex, it is easy to misuse the procedures and to draw unwarranted conclusions. For the factor analyst, I hope that this article provided a useful update regarding issues that affect the implementation and reporting of factor analyses, thereby leading to more sound analyses. The goal was also to provide ideas about expanding the application of factor analysis to research, training, and practice. Conducting a factor analysis may be imposing to one with minimal training, so adopting these ideas may benefit from collaboration with those adept at conducting factor analyses. This is a worthy endeavor because many of the challenges facing counseling psychologists can be overcome by simplifying data. Factor analysis therefore has the potential to move beyond the somewhat narrow realm of psychometrics into the broader domain of counseling psychology.

NOTES

1. Based on my review of articles in *Journal of Counseling Psychology* between January 2002 and April 2005 that used exploratory factor analysis (EFA), 45% of the articles used principal-axis factoring (PAF) to extract factors, 32% used principal components analysis (PCA), and 18% used maximum likelihood (ML).

2. Some authors (e.g., Bryant & Yarnold, 1995) distinguish between *PCA* and *principal components factor analysis*. I, as well as others (e.g., Fabrigar, Wegener, MacCallum, & Strahan, 1999), prefer the term *common factor analysis* to *principal components factor analysis*.

3. In the literature, the term *factor loading* typically describes structure coefficients, but because there are multiple statistics that indicate the relationship between variables and factors, Thompson (2004; Thompson & Daniel, 1996) recommends against using this potentially confusing term.

4. Because pattern and structure coefficients are equal in an orthogonal rotation, these are sometimes referred to as *pattern/structure coefficients*.

5. Of the 14 articles in *Journal of Counseling Psychology* between January 2002 and April 2005 that reported using a cutoff for retaining variables in EFA, 4 of them used a cutoff lower than .40, 5 used a cutoff of .40, and 5 used a cutoff higher than .40.

6. SPSS Inc. distributes AMOS—software that can estimate confirmatory factor analysis models—but it is not a part of the basic SPSS software.

7. Factor variance is also typically estimated, but in this case, the factor variance would be fixed to set the metric (see Bollen, 1989).

REFERENCES

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317-322.

Bauman, S., & Lenox, R. (2000). A psychometric analysis of a college counseling center intake checklist. *Journal of Counseling Psychology*, *47*, 454-459.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.

Bentler, P. M. (1995). *EQS structural equations program manual.* Encino, CA: Multivariate Software.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structure models. *British Journal of Mathematical and Statistical Psychology*, *37*, 62-83.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99-136). Washington, DC: American Psychological Association.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245-276.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, *56*, 754-761.

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

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin*, *105*, 317-327.

Cutrona, C. E., & Russell, D. W. (1987). The provisions of social relationships and adaptation to stress. *Advances in Personal Relationships*, *1*, 37-67.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272-299.

Fitzgerald, L. F., & Hubert, L. J. (1987). Multidimensional scaling: Some possibilities for counseling psychology. *Journal of Counseling Psychology*, *34*, 469-480.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, *7*, 286-299.

Forester, M., Kahn, J. H., & Hesson-McInnis, M. (2004). Factor structures of three measures of research self-efficacy. *Journal of Career Assessment*, *12*, 3-16.

Furlow, C. F., & Beretvas, C. N. (2005). Meta-analytic methods of pooling correlation matrices for structural equation modeling under different patterns of missing data. *Psychological Methods*, *10*, 227-254.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Guadagnoli, E., & Velicer, W. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, *103*, 265-275.

Helms, J. E., Henze, K. T., Sass, T. L., & Mifsud, V. A. (2006). Treating Cronbach's alpha reliability as data in counseling research. *The Counseling Psychologist*, *34*, 630-660.

Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, *34*, 601-629.

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, *65*, 202-226.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Hoyt, W. T., Warbasse, R. E., & Chu, E. Y. (in press). Construct validation in counseling psychology research. *The Counseling Psychologist*, *34*.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179-185.

Hu, L., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76-99). Thousand Oaks, CA: Sage.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Jackson, D. (2001). Sample size and number of parameter estimates in maximum likelihood confirmatory factor analysis: A Monte Carlo investigation. *Structural Equation Modeling*, *8*, 205-223.
Johnson, J. E., Burlingame, J. A., Olsen, J. A., Davies, R., & Gleave, R. L. (2005). Group climate, cohesion, alliance, and empathy in group psychotherapy: Multilevel structural equation models. *Journal of Counseling Psychology*, *52*, 310-321.
Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide.* Chicago: Scientific Software International.
Kahn, J. H. (2005). Institutional research productivity, use of theory-driven research, and statistical application in counseling psychology: Examining the research base. *The Counseling Psychologist*, *33*, 340-348.
Kahn, J. H., & Gelso, C. J. (1997). Factor structure of the Research Training Environment Scale–Revised: Implications for research training in applied psychology. *The Counseling Psychologist*, *25*, 22-37.
Karr, C. A., & Larson, L. M. (2005). Use of theory-driven research in counseling: Investigating three counseling psychology journals from 1990 to 1999. *The Counseling Psychologist*, *33*, 299-326.
Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.
Lebow, J. L. (1983). Research assessing consumer satisfaction with mental health treatment: A review of the findings. *Evaluation and Program Planning*, *6*, 211-236.
Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology*, *44*, 307-315.
MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130-149.
MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*, 84-99.
Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*, 391-410.
Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*, 320-341.
Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, *33*, 269-298.
Martens, M. P., & Hasse, R. F. (in press [*TCP* special issue, part 2]). Advanced applications of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, *34*.
Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, *10*, 259-284.
Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*, 156-166.
Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide.* Los Angeles: Author.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
Oh, E., & Neville, H. (2004). Development and validation of the Korean Rape Myth Acceptance Scale. *The Counseling Psychologist*, *32*, 301-331.
Phillips, J. C., & Russell, R. K. (1994). Research self-efficacy, the research training environment, and research productivity among graduate students in counseling psychology. *The Counseling Psychologist*, *22*, 628-641.
Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist*, *27*, 485-527.

Quintana, S. M., & Minami, T. (in press [*TCP* special issue, part 2]). Guidelines for meta-analyses of counseling psychology research. *The Counseling Psychologist*, *34*.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287-297.

Rounds, J., Tracey, T. J., & Hubert, L. (1992). Methods for evaluating vocational interest structural hypotheses. *Journal of Vocational Behavior*, *40*, 239-259.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin*. *Personality and Social Psychology Bulletin*, *28*, 1629-1646.

SAS Institute. (2004). *Base SAS 9.1 procedures guide*. Cary, NC: Author.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.

Schlosser, L. Z., & Gelso, C. J. (2001). Measuring the working alliance in advisor-advisee relationships in graduate school. *Journal of Counseling Psychology*, *48*, 157-167.

Sherry, A. (2006). Discriminant analysis in counseling psychology research. *The Counseling Psychologist*, *34*, 661-683.

Smith, D. A., Mar, C. M., & Turoff, B. K. (1998). The structure of schizophrenic symptoms: A meta-analytic confirmatory factor analysis. *Schizophrenia Research*, *31*, 57-70.

Snook, S. C., & Gorsuch, R. L. (1989). Component analysis versus common factor analysis: A Monte Carlo study. *Psychological Bulletin*, *106*, 148-154.

Sodowsky, G. R., Taffe, R. C., Gutkin, T. B., & Wise, S. L. (1994). Development of the Multicultural Counseling Inventory: A self-report measure of multicultural competencies. *Journal of Counseling Psychology*, *41*, 137-148.

SPSS Inc. (2004). *SPSS 13.0 base user's guide*. Chicago: Author.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: An historical overview and some guidelines. *Educational and Psychological Measurement*, *56*, 197-208.

Tinsley, H. E. A., & Tinsley, D. J. (1987). Use of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*, 414-424.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, *25*, 1-28.

Weston, R., & Gore, P. A., Jr. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, *34*, 719-751.

Widaman, K. F. (1993). Common factor analysis versus principal components analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, *28*, 263-311.

Worthington, R. L., & Whittaker, T. A. (in press [*TCP* special issue, part 2]). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*.

Zwick, W. R., & Velicer, W. G. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432-442.