DOCUMENT RESUME

ED 418 127	TM 028 221
AUTHOR	Kieffer, Kevin M.
TITLE	Some Comments on Analytic Traditions in EFA As against CFA: An Analysis of Selected Research Reports.
PUB DATE	1998-04-13
NOTE	70p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 1998).
PUB TYPE	Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Construct Validity; Factor Analysis; *Factor Structure; Heuristics; Literature Reviews; *Research Reports; Tables (Data)
IDENTIFIERS	*Confirmatory Factor Analysis; Data Reduction Methods; *Exploratory Factor Analysis

ABSTRACT

Factor analysis has historically been used for myriad purposes in the social and behavioral sciences, but an especially important application of this technique has been to evaluate construct validity. Since in the present milieu both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are readily available to the researcher, it is interesting to note several differences in the analytic traditions between the two techniques, even though they are both part of a single general linear model. This paper presents heuristic data to compare these two analytic methods, in addition to discussing recent published exemplars of both practices. The methods are compared with a dataset from K. Holzinger and F. Swineford (1939). EFA is a data reduction technique that permits the reduction of a large number of variables into constituent components by examining the amount of variance that can be reproduced by the latent or synthetic variables underlying the observed or measured variables. The development of true CFA has allowed researchers to test directly the fit between theories and data structure with a sample of data in hand rather than allowing a factor structure to emerge from data without regard to theoretical expectations as in exploratory factor analysis. Both techniques allow researchers to develop and critically examine theories regarding the structure of data sets. (Contains 16 tables and 49 references.) (Author/SLD)

Running head: ANALYTIC TRADITIONS

Some Comments on Analytic Traditions in EFA as Against CFA: An Analysis of Selected Research Reports

Kevin M. Kieffer

Texas A&M University 77843-4225

Paper presented at the annual meeting of the American Educational Research Association, San Diego, April 13, 1998.

Abstract

Factor analysis has historically been used for myriad purposes in the social and behavioral sciences, but an especially important application of this technique has been to evaluate construct validity. Since in the present milieu both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are readily available to the researcher, it is interesting to note several differences in the analytic traditions between the two techniques, even though they are both part of a single general linear model. The present paper presents heuristic data to compare these two analytic methods, in addition to discussing recent published exemplars of both practices.

Some Comments on Analytic Traditions in EFA as Against CFA: An Analysis of Selected Research Reports

The utilization of factor analytic techniques in the social sciences has been indelibly intertwined with developing theories and evaluating the construct validity of measures. As stated by Gorsuch (1983, p. 350), "A prime use of factor analysis has been in the development of both the theoretical constructs for an area and the operational representatives for the theoretical constructs." Since factor analysis has been deemed the "reigning queen of the correlational methods" (Cattell, 1978, p. 4), it is not surprising that Pedhazur and Schmelkin (1991, p. 66), stated that, "Of the various approaches to studying the internal structure of a set or indicators, probably the most useful is some variant of factor analysis."

The majority of researchers utilizing factor analytic techniques have employed what are contemporaneously termed "exploratory" factor analytic techniques (EFA). In this application of factor analysis, researchers are primarily concerned with the development of theories or the generation of alternative explanations for commonly accepted theories about a phenomenon of interest. More recently, however, a growing number of researchers have been employing "confirmatory" factor analytic techniques (CFA) that directly permit the testing of extant theories and/or the evaluation of instrument structure based on theoretical expectations by generating maximum likelihood parameter estimates. A hybrid of these two techniques utilizes exploratory factor extraction in combination with confirmatory factor rotation (Thompson, 1992); this hybrid, however, is only considered here briefly.

The purpose of the present paper was to briefly explicate both analytic techniques using a small heuristic data set that permits readers to grasp the conceptual underpinnings of both methods. The two techniques are directly compared using the Holzinger and Swineford (1939) data set that has been utilized extensively in illustrating factor analytic principles (cf. Gorsuch, 1983). The data set consists of the scores of 301 junior high school students on 24 different psychological inventories. For the purpose of the present heuristic example, however, only 13 of the original 24 tests were utilized. In addition to the comparison of EFA and CFA analytic techniques using the Holzinger and Swineford (1939) data set, various published exemplars of both approaches from the counseling psychology literature were examined as well.

Exploratory Factor Analysis

Historical Overview

The genesis of the analytic technique commonly referred to as exploratory factor analysis (as well as its confirmatory variant) can be traced back to the seminal work of Pearson (1901) and Spearman (1904). Pearson (1901) first conceptualized a general method of examining and extracting latent variables underlying data structures. Spearman (1904), through his work on personality theory, provided the conceptual and theoretical rationale for both exploratory and confirmatory factor analysis. Despite the fact that the conceptual basis for these methods have been available for many decades, it was not until the genesis and widespread availability of the computer that these analytic techniques were employed with any regularity.

Spearman (1904) was interested in providing support for his theory that there was one singular and universal personality factor, the G factor, that undergirded the larger sense of the self. He developed a mathematical and statistical method of evaluating the tenability of his hypothesis by subtracting a matrix of cross-products of structure coefficients from a matrix of correlation coefficients. The resultant residual matrix was examined to discern how much of the original variance remained after the variance associated with the G factor was extracted. Unfortunately, Spearman (1904) soon realized (as did others) that large coefficients remained in the residual matrix and that other "factors" could also be removed to better explain the variance in the data. Although the G factor theory of personality was not supported, an important statistical procedure was developed to help researchers evaluate the construct validity of a theory.

The complexity of the mathematical and statistical manipulations necessary to complete a factor analysis prevented many researchers from employing them prior to the advent of computers. A renewed interest in factor analytic techniques was evidenced, however, in the middle of the 20th century following a schism within the American Psychological Association (APA). Many practitioners abandoned the flagship organization of psychologists during the late 1930's and early 1940's due to an overemphasis on research interests and a paucity of practice related pursuits (Thompson & Daniel, 1996). Following the reunification of the organization, new standards of practice and research evolved out of the separation. Practitioners and researchers alike understood the need for standards to guide practice, and the first test standards were then developed concurrently with codes of ethical practice.

The test standards emphasized demonstrating the construct validity of test scores which, historically, had even been referred to as factorial validity. As stated by Guildford (1946, p. 428),

> The factorial validity of a test is given by its loadings in meaningful, common reference factors. This is the kind of validity that is really meant when the question is asked "Does this test measure what it is supposed to measure?"

Following the test standards developed by APA, many researchers and test constructors began utilizing factor analytic techniques to demonstrate the validity of scores generated by their instruments. The true power of factor analytic techniques, however, was not realized until the latter half of the 20th century when the computer facilitated the development of better exploratory techniques as well as the derivation of confirmatory analytic techniques that then became widely available to most researchers.

Theoretical Underpinnings

Contemporaneous exploratory factor analysis (EFA) is an analytic technique in which the primary concern is to reduce a set of larger variables into a smaller and more manageable set based on the consistency of the data. As indicated by Gorsuch (1983, p. 2),

> The purpose of factor analysis is . . . to summarize the interrelationships among variables in a concise but accurate manner as an aid in conceptualization. This is often achieved by including the maximum amount of information from the original variables in as few derived variables, or factors, as possible to keep the solution understandable.

Gorsuch (1983, p. 90) further stated, "Reducing the number of variables to a more reasonable subset is often a prime goal of factor analysis. A reduced variable set is sought that will contain as much of the information in the initial set of variables as possible."

Thus, EFA is a data reduction technique that permits the reduction of a large number of variables (e.g., test items, individuals) into constituent components by examining the amount of variance that can be reproduced by the latent or synthetic variables underlying the observed or measured variables. Tinsley and Tinsley (1987, p. 414) summarized the purpose of factor analysis by stating, "The goal of factor analysis is to achieve parsimony by using the smallest number of explanatory concepts to explain the maximum amount of common variance in a correlation matrix."

EFA has historically been utilized for two general purposes in the social sciences (Pedhazur & Schmelkin, 1991). The first general purpose has been to better understand the internal structure of an instrument or a data set when no previous information on the data structure is available. This is an example of utilizing EFA as a tool in the generation of theories about phenomena of interest, and is an example of an appropriate application of this analytic technique. The second general application of EFA has been to re-examine patterns in data sets when researchers have questioned the tenability of the emergent factors in previous research. Many researchers have utilized EFA to determine if the data structure in a current study resembles the factor structure of previous research on the same phenomenon of interest or when utilizing the same instrument. In this second application, many researchers have erroneously employed EFA to evaluate whether a given theory adequately fits a set of data by examining if the same general factors reemerge in subsequent studies. This use of EFA is often more appropriately explored through confirmatory factor analysis, as model-to-data fit can be directly evaluated in the confirmatory case. The more common and appropriate of these two applications of EFA has been in exploring the factor structure of a set of indicators (e.g., variables, test items, individuals, occasions) when no previous research is available.

EFA is not conceptually different from other techniques, as all analytic techniques are correlational (Thompson, 1997a, Cohen, 1968) and are part of one general linear model (GLM) subsumed by canonical correlation analysis (Knapp, 1978). Like other analyses, the goal of conducting an EFA is to explain the maximum amount of shared variance with the fewest number of explanatory concepts. The conceptual unity of factor analysis with canonical correlation analysis (and thus all other analyses) was illustrated by Hetzel (1996, p. 178):

> [S]ince all parametric methods are actually special cases of canonical correlation analysis . . . and since canonical correlation analysis itself invokes a principal components analysis [as factor analysis does] . . . [a]ll parametric methods actually invoke principal components mathematics.

Thus, EFA is not conceptually dissimilar from other statistical analyses as all analyses attempt to explain shared variance through principal components mathematics. The manner in which EFA permits the examination of shared variance among a set of indicators, however, is somewhat different from other analyses. Thus, to understand the conceptual and practical distinctions between exploratory and confirmatory factor analysis, it is critical to understand the terminology and mechanics of the both techniques.

Logic and Mechanics of EFA

Performing an EFA can be conceptualized as a series of steps which require that certain decisions be addressed at each individual stage. Consequently, there are many different ways in which to conduct a factor analysis, and each different approach may render distinct results when certain conditions are satisfied (cf. Gorsuch, 1983). The one consistent element in conducting an EFA, however, is that the results of the analysis are based solely on the mechanics and mathematics of the method and not on the a priori theoretical considerations of the researcher (Daniel, 1989).

Matrices of Association. One of the first decisions to be made in performing an EFA is to determine the manner in which the data matrix will be represented in the analysis. Since all statistical analyses are correlational, the focus of every statistical analysis is on the relationship among a set of variables or other entities (e.g., people) that may be factored. Matrices of association (e.g., correlation matrices, variance-covariance matrices) are arrays of numbers that are utilized to concisely express the linear relationships between a larger set of variables and an even larger set of scores on the variables. The most common matrices of association utilized in EFA are correlation matrices (in which values of 1.0 are on the main diagonal and bivariate correlation coefficients between the variables are on the off-diagonals) and variance-covariance matrices (in which the variance for a given variables is on the main diagonal and the covariances between pairs of variables are on the off-diagonals). The common analytic tradition in EFA, however, is to utilize the correlation matrix in computational analyses, perhaps partly because this is the default in most statistical software packages.

Factor Extraction. After the researcher has chosen which matrix of association will be utilized in the analysis, the researcher must then determine which extraction method to employ in conducting the analysis. Factor extraction refers to removing the common variance that is shared among a set of variables. There are currently several different techniques available for the extraction of common variance, and the results generated by the analysis can differ based on the particular method of extraction utilized.

Of the techniques available, principal components analysis and principal factors analysis are the two most widely used extraction methods in EFA. Although some researchers have argued that the difference between these extraction methods is negligible (cf. Thompson, 1992), other researchers have contended that the difference is substantial enough to warrant careful consideration of the extraction method utilized (cf. Gorsuch, 1983). Although the relative advantages and disadvantages of the two extraction methods have been discussed elsewhere (cf. Gorsuch, 1983; Stevens, 1996; Tinsley & Tinsley, 1987), factor extraction methods demand brief consideration in the present paper.

Principal components analysis (PCA) uses the total variance of each variable in examining the shared variance between variables (Hetzel, 1996). This is accomplished by placing values of 1.0 on the main diagonal of the correlation matrix (as each variable is expected to correlate perfectly with itself) and leaving the bivariate correlation coefficients on the off-diagonals. One limitation in utilizing PCA is that as the number of factored variables decreases or as the factored variables become less reliable, "... some of the factors represent correlated error variance and as such would be unlikely to be replicated in subsequent studies" (Hetzel, 1996, p. 186). Conversely, (a) as the number of factored variables increases and (b) as the factored variables become more reliable, the differences associated with utilizing PCA versus other extraction methods becomes negligible

(Thompson & Daniel, 1996).

Snook and Gorsuch (1989, p. 149) explained this first influence (i.e., the number of factored entities), noting that "As the number of variables decreases, the ratio of diagonal to off-diagonal elements also decreases, and therefore the value of the communality has an increasing effect on the analysis." For example, with 5 factored variables the 5 diagonal entries in the correlation matrix represent 20% (5 / 25) of the 25 entries in the matrix, but with 50 variables the diagonal entries represent only 2% (50 / 2,500) of the 2,500 matrix entries. Thus, Gorsuch (1983) suggested that with 30 or more variables the differences between solutions from these two methods are likely to be small and lead to similar interpretations. Of course, researchers rarely factor data involving only a small number of variables.

In principal factors analysis (PFA), an estimate of the reliability of each variable is placed on the main diagonal of the correlation matrix rather than a value of unity (i.e., 1.0). The reason for restructuring the correlation matrix is that it is believed that only the common or reliable variance indigenous to a variable will correlate with other variables in the matrix (Gorsuch, 1983). Thus, since placing a value of unity on the main diagonal would introduce error variance into the factor extraction procedure, proponents of PFA have contended that the most appropriate value to place on the main diagonal of the correlation matrix is some index of the common or reliable variance of a variable. Several indices of common or reliable variance can be utilized on the main diagonal of the correlation matrix, and the interested reader is referred to Gorsuch (1983), Stevens (1996) or Tinsley and Tinsley (1987).

Factor and Coefficient Generation. After the researcher has decided which factor extraction method to employ, the analysis can be conducted. One advantage in employing a factor analysis is that each latent or synthetic variable (factor) extracted from the analysis is perfectly uncorrelated with all of the other factors. This is often advantageous when the purpose of the EFA is theory generation as the interpretation of the extracted factors is thereby greatly simplified.

The extracted factors represent an attempt by the researcher to mathematically re-express the relationships between a set of variables with the fewest explanatory concepts possible. By doing this only a certain portion of the variance for any given variable will be reproduced by the factors (although it is theoretically possible for all of the variance associated with a variable to be reproduced by the factors, this rarely occurs in practice). The resultant mathematical manipulations required to extract the factors result in the formation of two matrices, the factor pattern matrix and the factor structure matrix. Regardless of the type of extraction method, the rows of the factor pattern and structure matrices are composed of the variables in the study and the columns of the matrices are composed of the latent constructs, or factors.

The factor pattern matrix is comprised of a series of weights (identical to ß weights in multiple regression analysis) that indicate the relative importance of a given variable to the extracted factors with the influence of the other variables removed (Stevens, 1996). The factor structure matrix, however, is simply the bivariate correlation of a measured/observed variable with scores on the extracted latent/synthetic factor. Factor structure coefficients are identical to structure coefficients in other analyses as they are simply the correlation between observed and latent variables (Thompson, 1997b; Thompson & Borrello, 1985). The extracted factors are perfectly uncorrelated, which always results in the equality of the factor pattern matrix and the factor structure matrix. Since the two matrices are equivalent, both matrices can be accounted for by employing the term, "factor pattern/structure matrix," to describe both matrices.

After the factor pattern/structure matrix is contrived, it is possible to generate two variance-accounted-for statistics that help the researcher determine the amount of variable variance that is reproduced by the latent constructs. The first of these is the communality coefficient, $\underline{h}^2$, which can be defined as the amount of variance on a variable that is reproduced by the factors. This value is calculated by summing the squared pattern/structure coefficients across the row for each variable. The resultant coefficient is an index of the proportion of total variance for a given variable that is reproduced by the extracted factors. Since it is a squared, variance-accounted-for statistic, it can range from 0 to 1.0.

Another variance-accounted-for statistic that is generated from the factor pattern/structure matrix is the eigenvalue. An eigenvalue represents the amount of variance in the original data set that is reproduced by a given factor. For the principal components case, eigenvalues can be computed by summing the squared factor pattern/structure coefficients down the columns of the matrix. Eigenvalues represent the amount of factor-reproduced variance and their values can range from 0.0 to the total number of variables in the analysis. Eigenvalues can also serve as an effect size measure, as each eigenvalue can be divided by the number of total variables in the analysis and a percentage of the total variance for a given factor can be computed. Due to the mathematics of EFA, factors with the largest eigenvalues are always extracted first, and each additional factor extracted will have a smaller eigenvalue than the first factor that was removed. The eigenvalues sum to the number of factored entities (e.g., variables), so if the first two eigenvalues for an EFA of 10 variables were 6 and 4, the remaining 8 eigenvalues would all be zero (and the pattern/structure coefficients on the last 8 factors would each also be zeros).

Factor Retention. After variance-accounted-for statistics and factor pattern/structure matrices have been computed, the researcher must decide the number of factors to retain in the analysis. Since different retention methods can often generate divergent results, two of the most popular retention methods will be briefly discussed.

The most popular method of determining the number of factors to retain is the eigenvalue greater than 1.0 rule. This decision rule was initially developed by Kaiser (1960), based on the work of Guttman (1954) and is often the default option on statistical software packages (Hetzel, 1996). Since eigenvalues are variance-accounted-for estimates that can range from 0.0 to the number of variables in the analysis, only the most salient factors are usually retained when this retention method is employed. It is important to carefully examine all of the eigenvalues, however, as previous research has reported that in certain situations the eigenvalues greater than 1.0 rule can underestimate or overestimate the number of factors that should be retained (cf. Hetzel, 1996).

A second popular decision rule utilized to determine the number of factors to retain is the scree test developed by Cattell (1966). The scree test is a graphical technique in which the eigenvalues are listed along the X-axis and their magnitude is plotted on the Y-axis. The resultant chart is visually inspected to ascertain the point at which the slope of the line connecting the eigenvalues becomes zero (horizontal). All of the factors above the point at which the slope of the line becomes horizontal are retained in the analysis. This technique has been considered by some researchers as too subjective to be considered noteworthy, but other researchers have contended that the scree test is one of the most accurate retention methods currently available (Zwick & Velicer, 1986).

Interpretation of Results. After the appropriate number of factors are retained in the analysis, it is necessary to interpret the results. It is often difficult to interpret the initial factor/pattern structure matrix as many of the variables typically manifest noteworthy coefficient magnitudes on many of the retained factors (coefficients greater than |0.60| are often considered large and coefficients of |0.30| are often considered moderate) and especially on the first factor. Thompson (184, pp. 31-34) demonstrates how the unrotated pattern/structure matrix actually misrepresents the true nature of the factors, and how rotation resolves this misrepresentation. Interpretation of the factor analytic results is therefore almost always aided by the rotation of the factor solution, as it is possible to redistribute the common variance across the factors to achieve a more parsimonious solution. It is important to note, however, that factor rotation is not "cheating" and does not generate or discover more common variance; rather, factor rotation merely redistributes the variance that has been previously explained by the extracted factors.

After a factor solution is rotated, the first factor may not account for the largest portion of the variance and thus may not have the largest variance-accounted-for value. Since the variance has been redistributed throughout the factors, any of the factors could account for the largest proportion of the total variance. Additionally, after the rotation is conducted, eigenvalues are no longer termed as such; rather, after rotation, the variance-accounted-for statistic for the factors (columns of the factor pattern/structure matrix) is termed "trace." One of the most common mistakes that researchers frequently commit is believing that the eigenvalue for a given factor after extraction will equal the trace after the factor solution is rotated (Hetzel, 1996).

The objective of factor rotation is to achieve simple structure (Thurstone, 1947) through the manipulation of the factor pattern matrix. The most parsimonious solution, or simple structure, has been explained by Gorsuch (1983, pp. 178-179) in terms of five principles of factor rotation:

1. Each variable should have at least one zero loading.
2. Each factor should have a set of linearly independent variables whose factor loadings are zero.
3. For every pair of factors, there should be several variables whose loadings are zero for one factor but not the other.
4. For every pair of factors, a large proportion of variables should have zero loadings on both factors whenever more than about four factors are extracted.
5. For every pair of factors, there should only be a small number of variables with nonzero loadings on both.

Thus, factor rotation is a technique devised to shift the factors in their factor space so that each variable in the analysis has a large factor pattern coefficient on only one factor and has very small or zero factor pattern coefficients on the other extracted latent constructs.

Two types of factor rotation are available: orthogonal and oblique. Orthogonal rotation shifts the factors in the factor space at 90 degree angles to one another to achieve the best simple structure. This rotation strategy maintains the perfectly uncorrelated nature of the factors after the solution is rotated, and often aids in the interpretation process since uncorrelated factors are easier to interpret. There are several orthogonal rotation strategies available, but one of the most popular orthogonal rotation technique is the varimax criterion developed by Kaiser (1960). In this technique the factors are "cleaned up" so that every observed variable has a large factor pattern/structure coefficient on only one of the factors. Varimax rotation produces factors that have large pattern/structure coefficients for a small number of variables and near-zero or very low pattern/structure coefficients with the other group of variables.

There are several advantages to employing orthogonal rotation strategies. First, the factors remain perfectly uncorrelated with one another and are inherently easier to interpret. Secondly, the factor pattern matrix and the factor structure matrix are equivalent and thus, only one matrix of association must be interpreted. This means that the solution is more parsimonious (i.e., fewer parameters are estimated) and thus, in theory, is more replicable.

Orthogonal rotation strategies do, however, have limitations. Orthogonal rotations often do not honor a given researcher's view of reality as the researcher may believe that two or more of the extracted and retained factors are correlated. Secondly, orthogonal rotation of factor solutions may oversimplify the relationships between the variables and the factors and may not always accurately represent these relationships. Consequently, some researchers have challenged the utility of orthogonal rotation strategies. Thurstone (1947, p. 139) contended that the use of orthogonal rotation indicates "...our ignorance of the nature of the underlying structure...The reason for using uncorrelated [factors] can be understood, but it cannot be justified." Similarly, Cattell (1978,

p. 128) argued in regard to researchers performing orthogonal rotation, "...in half of [the] cases it is done in ignorance of the issue rather than by deliberate intent." Consequently, even though orthogonal rotation eases the interpretability of the factor solution, it may not accurately portray the relationships between the variables and the emergent factors.

The second type of factor rotation is termed oblique rotation. This method of rotation provides for correlations among the latent constructs. This rotation strategy is termed oblique because the angle between the factors becomes greater or less than the 90 degree angle that is utilized to perform an orthogonal rotation.

One of the most popular oblique rotation strategies is promax (see Hetzel, 1996). In this technique, the researcher is attempting to achieve the most parsimonious simple structure given that the factors are allowed to be correlated with one another. However, an oblique factor solution inherently tends to be less parsimonious. For example, if 5 factors for 100 factored entities (e.g., variables) are extracted and orthogonally rotated, only 500 factor pattern/structure coefficients are estimated (the 5 x 5 factor correlation matrix is not estimated, since it is constrained to have 1's on the diagonal and 0's everywhere else). If the same EFA factors are rotated obliquely, *1,010* coefficients (500 factor pattern coefficients, plus 500 factor structure coefficients, plus 10 factor correlation coefficients (the 10 non-redundant off-diagonal entries in the 5 x 5 factor correlation matrix)) are estimated. [It might be argued, however, that only *510* coefficients are estimated in this case, since with either the 10 unique factor correlation coefficients, and either the 500 pattern or the 500 structure coefficients, the remaining 500 pattern or structure coefficients are fully determined.]

Oblique rotation strategies can be useful to researchers for a variety of reasons. One advantage of using an oblique rotation strategy is that the solution more closely honors the researcher's view of reality. Secondly, oblique rotations often demand more careful consideration of factor results by the researcher which can increase the generalizability of the results to other situations and which may increase the likelihood that the results will be replicated. Unfortunately, oblique rotations may be difficult to interpret, especially if there is a high degree of correlation among the factors. Since the factor pattern and factor structure matrices are not equal, both have to be interpreted in conjunction with the other.

However, a simpler structure for the pattern matrix resulting from an oblique rotation will always be the benefit of paying the price for this interpretation difficulty. But as the degree of correlation between the factors decreases, both orthogonal and oblique solutions will tend to provide increasingly similar results. Given that oblique solutions are less parsimonious and therefore less replicable, an oblique rotation would therefore only be employed when the benefits of simpler, more interpretable structure outweigh the costs of less replicability (i.e., when the orthogonal factors are not readily interpretable, and the oblique factors are fairly highly correlated but more interpretable).

Heuristic Example Using EFA

Since the basic conceptual tenets of EFA have been described in the preceding sections, of the present paper, a heuristic example can be utilized to illustrate in concrete terms the manner in which an analysis would be conducted. For the purposes of the heuristic example in the present paper, only tests T1, T3, T4, T7, T8, T10, T13, T15, T16, T17, T19, T22, T23 from the original Holzinger and Swineford (1939) data set will be utilized. The means, standard deviations and labels for the raw data are presented in Table 1.

Insert Table 1 About Here.

As described previously, the first step in completing a factor analysis is determining which matrix of association will be utilized in the analysis (e.g., the correlation matrix or the variance-covariance matrix). Since the analytic tradition in EFA has been to utilize the correlation matrix as the analyzed matrix of association, the correlation matrix for the 13 measured variables will be utilized in the present computational solution as well. The correlation matrix for the example data is presented in Table 2.

Insert Table 2 About Here.

After the matrix of association is chosen, it is possible to determine the extraction method. Since principal components analysis (PCA) yields results similar to principal factors analysis as the number of factored entities increases and since PCA requires no further manipulation of the data array, PCA was chosen to extract the latent constructs. The results of the PCA factor extraction are presented in Tables 3 and 4.

Insert Table 3 and 4 About Here.

Notice that two matrices are contrived as a result of the PCA: one is a factor pattern matrix (Table 3) and the other is a factor structure matrix (Table 4). Since PCA was utilized to extract the factors, the extracted factors are perfectly uncorrelated and, consequently, the factor pattern and structure matrices are exactly equal. Thus, the factor pattern matrix and the factor structure matrix can be combined into one factor pattern/structure matrix since all of the values are identical and no information will be lost.

The next step in completing the EFA is to determine how many factors to retain. To conserve space, only the retained factors are presented in Tables 3 and 4. The eigenvalue greater than one rule was utilized to determine an appropriate number of factors to retain and resulted in the retention of four factors. In practice, it is often helpful to utilize several methods in determining the final set of factors to retain since each method tends to produce different results in some instances (cf. Zwick & Velicer, 1986). Notice that the first factor had an eigenvalue of 3.99744 and accounted for 30.7% of the total variance (total variance is calculated as (3.99744 / 13) * 100). The eigenvalues of the subsequent factors decrease as the factor pattern/structure matrix is scanned from left to right. Correspondingly, the percentage of variance-accounted-for by each factor decreases in magnitude as the matrix is scanned from left to right.

The factor correlation matrix is presented in Table 5. The table indicates that each of the factors correlates perfectly with itself but does not correlate with any of the other factors (i.e., the factors are perfectly uncorrelated). This is the universal result of factor extraction, and if any other result where attained it would alert the researcher to either a serious psychometric or data entry/analysis problem.

Insert Table 5 About Here.

An examination of Table 4 will reveal that the factor saturation (which observed variables have large coefficients on which latent constructs) is so complex that it is difficult to interpret the factor pattern/structure matrix in its present form. That is, as reported in Tables 4 and 5, 12 of the 13 variables have pattern/structure coefficients greater than |.30|, and the thirteenth coefficient is almost |.30| (i.e., .29739 for variable T10). Thus, to more easily interpret the results, the four factor solution was first rotated to the varimax criterion. The results of the varimax rotated solution are presented in Table 6.

Insert Table 6 About Here.

An examination of the results presented for the four factor solution rotated to the varimax criterion reveals that ascertaining which variables are associated with which factors has been greatly facilitated by the rotation procedure. Factor I is most highly saturated with tests T7, T8, T19, T22, and T23. By examining the Table 1 information, it is possible to see that the five tests which were associated with Factor I seem to have a strong verbal component. Thus, it might be possible to name this factor the *verbal comprehension* factor, as all of the tests associated with this factor may measure verbal ability.

Factor II is most highly saturated with tests T1, T3, and T4, each of which seem to be assessing spatial ability. Consequently, this factor might be termed the *spatial ability* factor, as all the tests associated with it seem to be measuring perceptual ability. The third factor is most highly saturated with tests T15, T16, and T17. All of these tests appear to be measuring a *memory* component. Finally, Factor IV is most highly saturated with tests T10 and T13, which seem to be strongly associated with the *speeded execution* of tasks.

It is important to notice that the communality coefficients for the varimax rotated solution are identical to the communality coefficients in the unrotated four factor solution. The reason for this is that the variable variance reproduced by a given factor is redistributed in the rotated solution as no new variance is ever generated through a rotation procedure.

As stated previously, after the solution is rotated, the first extracted factor often does not account for the preponderance of the total variance. By examining the trace in Table 6, it is possible to view this dynamic in the present data set. Factor I accounted for 30.7% of the total variance in the unrotated factor solution and is still the strongest overall factor, but after rotation this factor accounts for only 19.7% of the total variance. However, the total variance-accounted-for the by the four factor solution before rotation (60.8%) is exactly equal to the total variance-accounted-for after rotation. As previously stated, no new variance is generated when the factors are rotated; rather the variance is merely distributed differently among the factors.

For the sake of comparing the parsimony of the results generated by orthogonal and oblique rotations, the unrotated factor pattern and factor structure matrices in Tables 3 and 4 were rotated to the direct oblimin criterion with delta equal to zero (cf. Gorsuch, 1983 for a more detailed explanation of the effects of varying the value of delta). The results of the direct oblimin rotation are presented in Tables 7 and 8.

Insert Table 7 and 8 About Here.

When interpreting the results of an oblique rotation, it is necessary to interpret two separate factor association matrices since the factor pattern matrix is no longer identical to the factor structure matrix. Similarly to multiple linear regression, however, it is *critically* important to interpret *both* pattern coefficients (standardized weights) and structure coefficients as each can provide only one piece of information regarding the larger relationship (Thompson, 1997b).

By examining the variance-accounted-for by each factor (trace), the results of the direct oblimin rotation appear to closely resemble the results generated by the varimax rotation. The communality coefficients in the oblique case are exactly equal to the results attained in the unrotated and varimax rotated factor solutions, thus again illustrating only the redistribution of common variance. The sum of the communality coefficients is still equal to 60.8%, as in all of the prior analyses. The trace and communality coefficients are computed slightly differently, however, as it is necessary to multiple a given factor pattern coefficient by the corresponding factor structure coefficient and then to sum down the columns or across the rows to derive the various variance-accounted-for estimate (see the note on Table 6 and Table 7 for a more detailed explanation).

After examining both the factor pattern and structure coefficients, the saturation of the factors can be determined. Factor I is most highly saturated with tests T7, T8, T19, T22, and T23 which is identical to the results generated by the orthogonal rotation. Factors II is most highly saturated with T15, T16, and T17 (similarly to Factor III in the varimax rotation), Factor III is most highly saturated with tests T1, T3, and T4 (similarly to Factor II in the varimax rotation) and Factor IV is most highly saturated with tests T10 and T13 (identical to Factor IV in the varimax rotation). Thus, based on the present example, both the orthogonal and oblique rotations provide generally the same conclusions in terms of factor-variable saturation. In the oblique case, however, it is necessary to consult two different matrices of association whereas in the orthogonal rotation is it necessary to examine only the single factor pattern/structure matrix.

The primary difference between these two rotation strategies is the correlation between the factors. The factor correlation matrix for the oblique rotation is presented in Table 9. As is evidenced by the table, all of the factors are correlated with one another. One of the difficulties incurred in interpreting oblique rotations is how to explicate a high degree of correlation among the factors. Most of the correlations in the present example are relatively small (with the possible exception of Factor I with Factor II, where $\underline{r}^2 = -.30059^2 = 9.0\%$) and pose little difficulty in the present analysis. Since both solutions are similar and the orthogonal rotation is interpretable and more parsimonious, most researcher would select this as the preferred solution.

Insert Table 9 About Here.

<u>Summary of Exploratory Factor Analysis</u>

The present heuristic example of EFA has demonstrated the usefulness of EFA as a tool in theory development and construct validation. EFA, however, does not invoke an inherent theoretical rationale for the determination of the number of factors to retain or for which rotation strategy to employ. Consequently, researchers must use subjective judgment in determining the extraction methods and rotation strategies that will be utilized in completing the analysis. EFA, therefore, is very useful in assessing the construct validity of scores and for examining the structure of data for which there is either a paucity of research or for which no research has previously been conducted.

Since exploratory methods do not directly address the issue of theory validation, their application is this realm is extremely limited. A second class of factor analytic techniques, confirmatory maximum likelihood techniques, are more appropriate to utilize when evaluating the fit of data to previously delineated theories. Thus, the limited value of EFA have predisposed some to characterize exploratory methods as neither ". . . a royal road to truth as some apparently feel, nor necessarily an adjunct to shotgun empiricism, as others claim" Nunnally (1978, p. 371).

Confirmatory Factor Analysis

As mentioned previously, the conceptual base upon which both contemporaneous exploratory and confirmatory factor analysis rest was initially developed as a method of providing statistical confirmation for a psychological construct (Spearman, 1904). Thus, factor analysis has historically been associated with issues of construct validation as illustrated in the statement, "factor analysis is intimately involved with questions of validity . . . Factor analysis is at the heart of the measurement of psychological constructs" (Nunnally, 1978, p. 112).

It was not until the development of the "true" confirmatory factor analytic techniques that directly permit the testing of hypotheses and theories in the latter half of the 20th century that the true power of factor analytic techniques has been realized. As stated by Thompson (1992, p. 83), "Without question, confirmatory factor analysis is more important to evaluating theory than is exploratory analysis." Similarly, Gorsuch (1983, p. 134) has stated, "...confirmatory factor analysis is the more theoretically important--and should be the much more widely used--of the two major factor analytic approaches."

Since one variant of confirmatory factor analysis (confirmatory rotation) invokes an exploratory analysis in its execution, the demarcation between EFA and CFA has been blurred in the minds of some researchers. Dickey (1996, p. 221) conceptualized the distinction between the appropriate use of EFA and CFA with the following statement:

> The decision over the legitimate use of either exploratory or confirmatory factor analysis depends, to a large degree, on what the presenting question is that the researcher wishes to address. In situations where the researcher has a specific detailed hypothesis and theoretical support, the confirmatory process makes more sense and provides more information than does the exploratory analysis. In situations where the researcher is charting new territory, or does not have a specifically outlined theory of underlying constructs, the use of exploratory analysis is a necessary first step.

The utilization of exploratory techniques to assess the tenability of specific research hypotheses, therefore, is unacceptable in light of the ready availability of true confirmatory techniques that directly permit theory evaluation.

Confirmatory techniques have long been needed in social science research to provide support for the theories that have been generated through exploratory analyses. Although other researchers had attempted to develop CFA prior to the mid 1960's, the seminal work of Jöreskog (1966, 1969) forever redefined the landscape of factor analysis. Jöreskog (1966, 1969) developed a group of maximum likelihood estimation techniques that has since been loosely termed confirmatory factor analysis (CFA). As stated by Thompson (1992, p. 83), the CFA methods are,

> a truly confirmatory class of maximum likelihood methods; that is, the methods assume a sample and focus use of the sample data on best estimating population parameters. The methods are truly confirmatory because the methods allow the testing of any of the many possible factor analytic hypotheses as a model representing a complete *omnibus system* of hypotheses [emphasis in original].

Thus, the development of true confirmatory factor analytic methods has allowed researchers to directly test the fit between theories and data structure with a sample of data in hand rather than allowing a factor structure to emerge from a set of data without regard to theoretical expectations as in exploratory analysis.

Confirmatory Rotation Through Exploratory Factor Extraction

Of the methods of CFA currently available, two methods, confirmatory rotation through exploratory factor extraction (hence referred to as confirmatory rotation) and confirmatory factor extraction procedures (hence referred to as confirmatory extraction), are discussed in the present paper. The first method, confirmatory rotation, refers to a procedure of extracting factors through an EFA and then rotating only the expected factors (which are based on past research or theoretical assumptions) to a best fit "Procrustean" criterion consistent with previous research findings. Since the purpose of CFA is to determine if a set of data support previous research, this approach can be used to determine the degree of fit between the structure of present data and past research findings.

Confirmatory rotation consists of three general steps. First, the researcher must extract the latent constructs from the data array using an exploratory analysis (typically principal components analysis) and then rotate the factor solution to an orthogonal criterion. The resultant factor pattern/structure matrix is then examined to determine which factors are saturated with which observed variables. A saliency level of |0.40| is typically employed to evaluate which variables are associated with which factors. The researcher can then determine if the extracted and rotated structure resembles past research, and, subsequently, interpret the emergent factors.

The second step in the confirmatory rotation procedure is to rotate the varimax solution to a best fit Procrustean fit with a theoretical target or a target defined using the factor solution from a previous study. Procrustean rotation

involves projecting the observed and expected solutions into

> the same factor space by rotating actual results to the best fit position with the expected factors. The cosines of the paired factors across the observed and "target" models are actually correlation coefficients, and hence provide estimates of the degree of goodness (or badness) of fit between the two factor solutions. (Daniel, 1989, p. 18)

The Procrustean rotation procedure is utilized to determine the adequacy of fit between the attained (observed) model and the expected (hypothesized) model. The observed factor structure is said to fit the hypothesized model if the cosines among paired factor axes approach unity (1.0), and if the observed factors correlate weakly with all other factors than the one they are hypothesized to embody. The computer program RELATE developed by Veldman (1967) can be utilized to evaluate the adequacy of the fit between the observed data and the hypothesized model. Thompson (1992) provides a test distribution.

Confirmatory Factor Extraction

A second type of CFA procedure is the confirmatory extraction procedure. Confirmatory extraction is considered a "pure" form of CFA and does not require the rotation of factor results, because the maximum likelihood estimates generated through the analysis are final estimates. The purpose of confirmatory extraction is to determine the true fit of the observed variables with the theoretical assumptions of the researcher. In this pure application of CFA, no consideration is given to exploratory techniques as they are not useful in determining the fit of the data with the theoretical model.

It is necessary in confirmatory extraction to utilize a statistical software package such as LISREL 7 (Jöreskog & Sorbom, 1989). Several competing models are input into the computer program, rather than only the single expected model, as more than one model can adequately represent a given data structure. When inputing the models, the researcher must decide which parameters of the model are "fixed" and which are "free." Free parameters refer to those specific parameters that will be estimated by the analysis. Conversely, a fixed parameter is a parameter that the researcher does not wish to estimate during the analysis. Fixed parameters are denoted as zeros in the target matrix, whereas freed parameters are generally demarcated as 1.0. Factor correlation coefficients, factor coefficients (pattern/structure coefficients) and the variance/covariance of the error of measurement are typically freed or fixed during a given analysis based on the a priori theoretical assumptions of the researcher.

The purpose in conducting the analysis is to determine which model or models best fit the data. Consequently, the analysis generates several statistics which help the researcher determine if the fit of the model with the data is supported. These "fit" statistics include the Bentler Comparative Fit Index (CFI; Bentler, 1990), the chi square--degrees of freedom ratio, the parsimony ratio, The Goodness of Fit Index and the Adjusted Goodness of Fit Index (GFI, AGFI; Joreskog & Sorbom, 1989), the Root Mean Square Residual (RMSR) and the Root Mean Square Error of Approximation.

More complete explanations of these fit statistics are presented elsewhere (cf. Daniel, 1989; Gillaspy, 1997), but each of the fit statistics is not an individual indicator of the model fit. Only when the fit statistics are used in conjunction with one another can the model-to-data fit be accurately assessed (Campbell, Gillaspy & Thompson, 1995).

If the analysis is successful in identifying the freed parameters, then the model is said to demonstrate a fit with the data. This does not mean that the results have been confirmed, however, as more than one model can fit the data (even models that were not hypothesized in the original analysis). Fit is assessed through the consultation of a variety of the aforementioned fit statistics. If the model is not identified by the data, it is possible to use the analysis to generate better fitting models. For instance, LISREL provides diagnostic information as to the fit of the model if certain parameters would have been freed (termed modification indices) or fixed in the analysis (often evaluating by considering "fixing" parameters when the ratio of a given parameter to its standard error is less than 2.0). Although it is not always appropriate to change the model based on these statistics, in certain situations it might be appropriate to alter some of the fixed and freed parameters to generate a better data to model fit (see Byrne, 1989; 1994).

Heuristic Example Using CFA

Since confirmatory factor extraction is considered a pure form of CFA, it is the only technique that is examined in some detail in the present paper. However, confirmatory rotation is a viable methodology and in some instances may provide important explanatory information. When using the confirmatory rotation procedure, the retained factors and the expected factors are projected into the same factor space. Using a Procrustean or best fit rotation strategy, the retained exploratory factors are rotated to a position that best honors the expected factor structure. The cosines between the expected factors and the retained factors are then computed. If there is a high degree of fit between the expected and retained structure, then the cosines should approach unity (Daniel, 1989). Since this is the lesser utilized of the two types of CFA analyses, it is not considered any further in the present paper. Interested readers are referred to Daniel (1989), Gorsuch (1983), and Thompson (1992).

The same heuristic data set utilized in the example application of EFA (Holzinger & Swineford, 1939) was again employed to demonstrate the completion of a confirmatory extraction procedure. As noted previously, only tests T1, T3, T4, T7, T8, T10, T13, T15, T16, T17, T19, T22, and T23 of the 24 original tests were utilized in the analysis.

Model Specification

The first step in completing a CFA is typically to specify several competing models. Since several models can fit a given data set, the finding that a particular model best fits the data given testing of several models, provides stronger support for the fitting model. Both of the tested models fixed factor variance to 1.0 and freed selected factor pattern and factor correlation coefficients. The covariance matrix among the 13 variables was utilized to complete the two analyses. In addition to the two posited models, each of the models was compared to a null model indicating that no factors would be identified by the analysis. Further, Model #2 was also compared to a CFA analysis which utilized the same fixed and freed parameters but which used the correlation matrix rather than the covariance matrix in completing the analysis.

Model #1 (v=13, n=301). This model indicated that the original five factors delineated by the Holzinger and Swineford (1939) study would be identified in the analysis. The expectation is that each variable will have large factor pattern coefficients on only the one factor with which the variable is expected to be associated: (a) Verbal Factor: T7 and T8; (b) Visual Factor: T1, T3, and T4; (c) Speed Factor: T10 and T13; (d) Memory Factor T15, T16, T17, and T19; and (e) Math Factor: T22 and T23. Holzinger and Swineford (1939) indicated that the five expected factors were uncorrelated, but for the purpose of the present analysis, the factor correlation coefficients were estimated. Furthermore, in all of these analyses involving factors presumed to be measured by only two variables, the pair of pattern coefficients were constrained to be equal.

Model #2 (v=13, n=301). This model indicated that the four factors retained in the EFA completed in the present paper would be identified by the analysis. The expectation is that each variable will have large factor pattern coefficients on only the one factor with which the variable is expected to be associated: (a) Verbal

Factor: T7, T8, T19, T22, and T23; (b) Visual Factor: T1, T3, and T4; (c) Speed Factor: T10 and T13; (d) Memory Factor T15, T16, and T17. Based on the results from the oblique rotation in the present EFA, it was expected that all of the factors would be correlated.

It very is important to emphasize a distinction between the heuristic rationale for the models specified in the present CFA analysis and the rationale that would be invoked in actual practice. In actual practice it would be considered inappropriate to utilize the same data set in both an EFA and CFA analysis, since the CFA application would no longer be truly confirmatory. Thus, if the present CFA were being conducted for anything other than heuristic value, two separate samples of data would be required to conduct the EFA and CFA analyses. Similarly, if the CFA were being conducted in practice, it would be necessary to test several other models as well. The more models that are included in the analysis, the more support that can be demonstrated for the model being supported (or conversely not being supported) by the data. It is important to remember, however, that CFA does not ever prove that a model is true; rather, a CFA only indicates which of the tested models best fits the data.

Interpretation of the CFA Results

The next step in completing the CFA is to run the analysis and interpret the generated fit statistics. The current analysis was completed using LISREL 7. Selected fit statistics from the CFA analysis are presented in Table 10. It is important to note that the presented fit statistics are only a few of the several dozen fit statistics currently available (Fan, Wang & Thompson, 1996). Fan, Wang and Thompson (1996) indicated that since many of the fit statistics have been developed with different rationales, it is important to first appreciate the rationale and the differential interpretation of the statistics as well as to understand that these statistics may vary across research situations.

Insert Table 10 About Here.

There are three general types of fit statistics currently available: covariance matrix reproduction indices, comparative model fit indices, and the parsimony weighted indices. The three types of fit statistics are explained more fully by Fan, Wang and Thompson (1996, pp. 6-7):

> *covariance matrix reproduction* indices . . . attemp[t] to assess the degree to which the reproduced covariance matrix based on the specified model has accounted for the original sample covariance matrix . . . . [These include] the Goodness-of-Fit Index (GFI) and the Adjusted Goodness-of-Fit Index (AGFI). . . . The second type of fit index--*comparative model fit* indices--assesses model fit by evaluating the comparative fit of a given model with that of a more restricted null model. . . . [These include] Bentler and Bollen's normed and non-normed fit indices (NFI and N_NFI) and Bollen's incremental fit index (DELTA2). . . . The third type of fit index--*parsimony weighted* indices--specifically takes the model parsimony

into consideration by imposing penalties for specifying more elaborate models. More specifically, these fit indices consider both model fit and the degrees of freedom used for specifying the model.

Since there is not currently one accepted fit statistic that accurately determines the model-to-data fit, it is necessary to examine several of these different types of indexes to determine if a given model fits (Thompson & Daniel, 1996). The researcher hopes that a variety of fit statistics jointly corroborate the fit of a model.

The maximum likelihood estimates and factor correlation matrix for model #1 are presented in Tables 11 and 12, respectively. The maximum likelihood estimates and the factor correlation matrix for Model #2 are presented in Tables 13 and 14, respectively. Both of the factor correlation matrices indicated a moderate to large degree of correlation between the factors for each model (ranging from .328 to .786). The chi-square statistic for Model #1 (χ^2 = 165.16, df = 58) was smaller in relation to the number of degrees of freedom utilized than in Model #2 (the ratio of chi-square to df should be 2 to 1 or less to indicate a good model-to-data fit), indicating that Model #1 better fit the data on this statistic. The GFI index indicated a better fit for Model #1 (GFI = 0.960) than for Model #2 (GFI = 0.899) as well. After sample size was taken into account, the AGFI still indicated that Model #1 better fit the data than Model #2.

Insert Tables 11, 12, 13 and 14 About Here.

In terms of the parsimony of the two models, both parsimony ratios indicated that Model #2 was the less complex of two models. When the parsimony ratio was utilized to weight the GFI and the Comparative Fit Index (CFI), a better model-to-data fit was indicated by Model #1. Similarly, the Root Mean Square Residual (RMSR) and the Root Mean Square Error of Approximation (RMSEA) both indicated that Model #1 was a better fit to the data than Model #2. It is generally accepted that a good fit on the RMSR and RMSEA statistic is indicated if the value attained by the model is less than 0.05 (a better fit is indicated as the statistics approach 0). Thus, the RMSEA statistic indicates a good fit for both models whereas the RMSR statistic indicates a very poor model-to-data fit for both models. Consequently, it is apparent from the present analysis that it is critical to interpret several fit statistics when assessing model fit, as each statistic can provide a divergent conclusion.

After examining the fit statistics, it is important to examine the modification indices and the parameter-to-standard-error ratios, respectively, to determine if a better model-to-data fit would be generated if some of the parameters were freed or fixed, respectively. It is important to note that utilizing the modification indexes and parameter-to-standard-error ratios to alter the expected model constitutes using the CFA in an exploratory manner, as the only theoretical rationale invoked in altering which parameters are fixed or freed results from the mathematics of the technique. Based on the information generated by the modification indexes, a noteworthy reduction in the chi-square would be achieved in Model #1 if the T10 and T13 parameters were freed on the Visual factor (39.939 for both variables). Similarly for Model #2, a noteworthy reduction in the chi-square would be evidenced if the T10, T13, and T23 parameters were freed on the Visual factor (estimated reduction of 40.758 for T10 and T13 and a reduction of 30.331 for T23). Freeing these parameters, however, would decrease the parsimony of the two models and therefore effect the parsimony weighted fit statistics.

An important consideration in completing a CFA is the effect that utilizing the correlation matrix versus the variance-covariance matrix will have on the analysis. As stated previously, the correlation matrix is a symmetric matrix with 1's on the main diagonal and correlation coefficients on the off-diagonals. The variance-covariance matrix has the variance of each variable on the main diagonal and covariances on the off-diagonal. Using the correlation matrix has the effect of standardizing the solution, as here both the factors and the variables would be constrained to unit variance (1.0). It can be argued, therefore, that using the correlation matrix can be likened to utilizing a variance-covariance matrix where the measured variables have been standardized to a variance of 1.0. If the fit statistics generated by utilizing the correlation matrix are equivalent to those generated by using the covariance matrix, then the correlation matrix solution can be interpreted. If the results are divergent, however, the variance-covariance matrix should be used in calculating the fit statistics, although a standardized or completely standardized solution would still be used to interpret the results (see Cudeck (1989) or Byrne (1989, 1994) for more detail).

Comparatively, utilizing the correlation matrix versus the covariance matrix in the analysis of Model #2 made very little difference in the present analysis as the generated GFI, AGFI, and chi-square were very similar. In terms of interpretation of the results, the same conclusions would be rendered in this instance by using either matrix of association.

Based on the results presented in the present paper, Model #1 best fit the data presented in Holzinger and Swineford (1939). Model #1 posited that each variable was associated with only one of five expected factors and that the factors were correlated.

Summary of Confirmatory Factor Analysis

The present heuristic example of CFA has demonstrated the usefulness of CFA as a tool in providing support for the theoretical expectations of data structures. CFA allows researchers to directly test the theoretical expectations generated either by previous research or through intrinsic value assessment. As stated by Daniel (1989, p. 24), "Used alone or in tandem with exploratory methods, confirmatory methods can help the researcher avoid erroneous conclusions about factor structures which might emerge using exploratory methods alone."

Analysis of Select Research Articles in the Counseling Literature

Since exploratory methods should be reserved for true conditions in which there is no prior research on a given area of inquiry, and since some have stated that CFA should be the more widely utilized of the factor analytic approaches (cf. Gorsuch, 1983), the present author randomly reviewed the factor analytic practices in the counseling psychology literature. Since the Journal of Counseling Psychology is the flagship journal of the American Psychological Association (the primary association of psychologists), this resource was deemed the best and most accurate indicator of the practices throughout the field of counseling psychology. Ten journal articles (five articles each that reported to be either exploratory or confirmatory factor analyses) were randomly selected from all of the articles utilizing CFA or EFA techniques over the past 10 years (1988-1997). The chosen articles are presented in Table 11.

Insert Table 15 About Here.

As with many disciplines in the social sciences, factor analytic techniques have been employed in counseling psychology to help confirm the findings of past research as well as to provide initial evidence for the construct validity of test scores and psychological theories. Tinsley and Tinsley (1987) noted that most studies which employed factor analytic techniques in counseling psychology did not provide adequate information for independent and post hoc analysis of the study results. Consequently, these authors found that interested and even skeptical researchers were simply forced to rely on the interpretation posited by the authors of the study rather than to have the necessary statistics to independently evaluate the analysis themselves. It is not surprising, therefore, that several authors have recommended that researchers routinely include pertinent information or make such information available to interested readers (Gorsuch, 1983; Hetzel, 1996; Thompson & Daniel, 1996).

In a recent review of select issues of JCP, Hetzel (1996) revealed that the remarks posited by Tinsley and Tinsley (1987) nearly 10 years before had seemingly passed unnoticed by researchers publishing in the journal. Hetzel (1996, p. 198-199) noted the information necessary to permit other researchers to replicate a given analysis:

> (a) background information, such as sample size, sample composition, method of selecting the subjects, and that method of selecting the variables; (b) matrix of association used; (c) method of factor extraction; (d) initial communality estimates used; (e) the criteria used for determining the number of factors to retain; and (f) the method of rotation used. In addition, the following . . . should be included . . . (a) the means and variances of the items; (b) the matrix of associations among the items; (c) the rotated factor pattern and structure matrices; and (d) the final communality coefficients, eigenvalues, and the proportion of variance explained by

each rotated factor.

Regrettably, Hetzel (1996) found that despite the ardent efforts of other authors to promote good reporting practice, none of the 13 articles reviewed reported all of the necessary components. An even more disparaging result indicated by Hetzel (1996) was that none of the reviewed articles reported the matrix of association utilized in the analysis (which is necessary to replicate the analysis), and only 7% of the articles reported the means and variances of the variables. Hetzel (1996, p. 202) concluded his review by somberly stating that,

> As . . . noted over two decades ago, complete reporting of a factor analysis can permit re-analysis of obtained results by other researchers or use of other methods that may lead to new and valuable insights. If such replication is the cornerstone of the science, then the field of counseling psychology stands upon a less-than-sturdy foundation.

In a recent editorial, Thompson and Daniel (1996) delineated several guidelines for authors to follow when reporting the results of factor analytic studies. In reference to reporting exploratory analyses, authors were (a) admonished on the use of the term "loading" as it is ambiguous and does not clearly indicate whether the specified coefficient is a factor pattern or structure coefficient; (b) encouraged the use of multiple criteria to determine the number of EFA factors to retain; (c) demanded distinctions between prerotation eigenvalues and postrotation trace;

and (d) required reporting of the factor extraction method utilized. Similarly, in regard to confirmatory analyses, authors were (a) encouraged to utilize several competing models in providing support for the confirmation of the theory under consideration; (b) requested to provide and interpret several fit statistics since the "characteristics of the fit statistics are not yet entirely clear" (p. 204); and (c) asked to base the analysis on the variance-covariance matrix, as recommended by other authors in the literature. Thompson and Daniel (1996) concluded their editorial by stating that the inclusion of appropriate information in reporting factor analytic studies "facilitate[s] further tests of rival models and the more rapid accumulation of insight" (p. 206).

Analysis of Selected Articles

To facilitate comparison of the present review with previous conjecture on the subject, the articles included in the present review were evaluated on the basis of the guidelines presented by Hetzel (1996) and Thompson and Daniel (1996). All of the articles were examined on the same dimensions explored in the Hetzel (1996) review and additionally on the recommendations provided by Thompson and Daniel (1996). These data are presented in Table 12.

Insert Table 16 About Here.

Consistent with the findings in the Hetzel (1996) review, all of the studies in the present analysis presented sufficient background information that included the size of the sample utilized and the composition of the sample. In regard to the exploratory analyses, all of the studies reported the criteria for factor retention, the method of factor rotation, and the eigenvalues of the retained unrotated factors. Unfortunately in several cases it was not possible to determine if the term "eigenvalue" truly referred to prerotation eigenvalues or if the authors intended this term to more precisely indicate the postrotation trace. This finding is in contrast with results reported in Hetzel (1996), as the previous review indicated the reporting rates of these variables to be 69% 92% and 53%, respectively.

Other exploratory findings in the present study were consistent with results reported in Hetzel (1996). For instance, Hetzel (1996) noted that communality coefficients, method of factor extraction, item means and variances, rotated factor pattern coefficients, rotated factor structure coefficients and final communality coefficients were reported in 23%, 92%, 7%, 38%, 33%, and 23% of the studies, respectively. The present study indicated that the same information was reported in 20%, 80%, 20%, 60%, 60%, and 60% of the studies examined. The most striking difference between the results of the present study and the prior review is that the matrix of association utilized in the analysis was reported in 40% of the studies examined in the present review compared to none of the studies reviewed by Hetzel (1996). Even though the articles included in the present review did not fare well on many of the evaluated dimensions, the trend of reporting the matrix of association is an improvement over prior reporting practice.

In regard to the recommendations posited by Thompson and Daniel

(1996), many of the articles again failed to report the critical information necessary to replicate the analysis and/or to provide readers with sufficient information to generate independent interpretations. For instance, in the exploratory analyses all of the articles simply termed the factor pattern or structure coefficients, "loadings," rather than clearly indicating whether the coefficients were indeed pattern or structure coefficients. Many of the articles labeled the factor variance-accounted-for statistics "eigenvalues" even after indicating that the factor solution had been rotated to various criterion. Additionally, only 20% of the articles included the method of factor extraction utilized.

The confirmatory analyses did not fare much better than the exploratory analyses in regard to the evaluated dimensions. Only 60% of the CFA studies reported the matrix of association utilized and only 60% of the studies indicated that other models were used in the falsification process. Of the studies that reported using other models, only 40% actually indicated the structure of these models (e.g., one general factor, three uncorrelated first-order factors). Additionally, only 60% of the confirmatory analyses interpreted several different fit statistics (i.e., more than three different statistics). Regrettably, the most common fit statistic utilized was the chi-square goodness of fit statistic which is subject to inflation by the influence of large sample sizes, as all statistic significance tests are (Thompson & Daniel, 1996). Based on the data compiled in the present paper, therefore, the results reported in Hetzel (1996) were corroborated, and present evidence suggests that many studies are failing to provide readers with necessary information for independent evaluation of results.

Brief Comments on Analytic Traditions in EFA and CFA

Analytic traditions in EFA and CFA have differed to some degree. For instance, most EFA analyses have historically utilized orthogonal rotation strategies that allow the latent constructs to remain perfectly uncorrelated after the factor solution is rotated for the purpose of interpretation. One reason for the emergence of this tradition is the degree of difficulty in interpreting the results of the solution. When factors are removed from the underlying data structure, it is generally easier to interpret the results if the factors are perfectly uncorrelated.

Conversely, most researchers estimate the correlation among the factors when performing a confirmatory analysis. One reason the factors are allowed to be correlated in CFA is to evaluate the independence of the factors and to ensure that two latent constructs are not perfectly correlated and thus evaluating the same construct. Another reason that factors are typically allowed to correlate in CFA is so that a given researcher can evaluate the fit of the data with the delineated model, especially if the expected factors were correlated in previous research studies.

In the present study, 60% of the confirmatory studies and 80% of the exploratory studies allowed the factors to be correlated. This finding appears contrary to the analytic traditions present in both analyses. A possible explication of this phenomenon is that in several of the exploratory studies oblique rotation strategies were employed because previous research indicated that the factors could be correlated. Since many of these researchers had theoretical expectations about the structure of the data prior to analysis, confirmatory analysis may have been more appropriate to utilize in these instances. Other authors in the exploratory case indicated that both orthogonal and oblique rotations were utilized initially, but since there were small correlations between the factors, the orthogonal results were interpreted. Similarly, several of the authors that reported utilizing confirmatory analyses contended that the correlations among the factors were indicated in previous research. Based on the results of the present study, however, there is little support provided for the notion that exploratory analyses typically use orthogonal rotations.

Another tradition that is often evidenced in factor analytic studies is a propensity to utilize one matrix of association more often than another. For instance, most exploratory analyses typically utilize correlation matrices in computations whereas most confirmatory analyses use variance-covariance matrices. Several authors have indicated that variance-covariance matrices are often more appropriate to utilize with confirmatory techniques (Bollen, 1989; Thompson & Daniel, 1996) whereas either matrix is often appropriate to use in exploratory analyses.

In the present review, it was difficult if not impossible to evaluate this dynamic as only 40% of the articles reported the matrix of association utilized in the analysis. Of those CFA studies reporting which matrix was used, 30% employed the correlation matrix in the analysis compared with only 10% that utilized the variance-covariance matrix. Similarly, 20% of the exploratory analyses reported using the correlation matrix in computing the factor analysis. While the trend manifested in the exploratory analyses examined in the present study is consistent with analytic tradition, the trend for the confirmatory analyses is contrary to both recommendations in the literature and analytic traditions.

Support for two analytic traditions evidenced in exploratory and confirmatory analyses (i.e., degree of correlation between the factors and propensity for one type of matrix of association) was not evidenced in the present review. It is possible that these traditions are changing to some degree in the contemporary milieu of factor analysis. Due to the small number of studies reviewed, however, it is impossible to render definitive conclusions. The present review, however, indicates that these traditions have not been supported in the current counseling psychology literature.

Summary and Conclusions

The present paper has provided an introductory treatment of both exploratory and confirmatory factor analytic techniques. The pertinent aspects of each of the analyses have been illustrated and the conceptual explanations have been illuminated by concrete heuristic examples. In addition, exemplars of both techniques published in recent issues of a prominent counseling psychology journal have been examined in regard to analytic traditions and practices of reporting indigenous to each technique.

Factor analysis remains a useful and viable analytic tool in social science research. Exploratory and confirmatory techniques allow researchers to develop and critically examine theories regarding the structure of data sets. Several critical decisions must be rendered throughout the analytic process and these analyses require careful and thoughtful consideration on the part of the researcher. When utilized properly, however, factor analysis can be a powerful analytic tool for both the purposes of data exploration and theory confirmation.

References

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

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

Byrne, B.M. (1989). A primer of LISREL. New York: Springer-Verlag.

Byrne, B.M. (1994). Structural equation modeling with EQS and EQS/Windows. Thousand Oaks, CA: Sage.

Campbell, T.C., Gillaspy, J.A., & Thompson, B. (1995). The factor structure of the Bem Sex-Role Inventory (BSRI): A confirmatory factor analysis. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas. (ERIC Document Reproduction Service No. ED 380 491)

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

Cattell, R.B. (1978). The scientific use of factor analysis in behavioral and life sciences. New York: Plenum.

Cohen, J. (1968). Multiple regression as a general data-analytic system. Psychological Bulletin, 70, 426-433.

Cole, D.A., & Jordan, A.E. (1989). Assessment of cohesion and adaptability in component family dyads: A question of convergent and discriminant validity. Journal of Counseling Psychology, 36, 456-463.

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

Daniel, L.G. (1989, November). Comparisons of exploratory and confirmatory factor analysis. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock. (ERIC Document Reproduction Service No. ED 314 447)

Dickey, D. (1996). Testing the fit of our models of psychological dynamics using confirmatory methods: An introductory primer. In B. Thompson (Ed.), Advances in social science methodology (Vol. 4, pp. 219-227). Greenwich, CT: JAI Press.

Fan, X., Wang, L., & Thompson, B. (1996, April). The effects of sample size, estimation methods, and model specification on SEM fit indices. Paper presented at the annual meeting of the American Educational Research Association, New York. (ERIC Document Reproduction Service No. ED 400 336).

Gillaspy, J.A. (1997). The factor structure of the working alliance inventory: Confirmatory and exploratory factor analysis. Unpublished doctoral dissertation, Texas A&M University, College Station, TX.

Good, G.E., Robertson, J.M., O'Neil, J.M., Fitzgerald, L.F., Stevens, M., DeBord, K.A., & Bartels, K.M. (1995). Male gender role conflict: Psychometric issues and relations to distress. Journal of Counseling Psychology, 42, 3-10.

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

Guttman, L. (1954). Some necessary conditions for common-factor analysis. Psychometrika, 19, 149-161.

Hayes, T.J., & Tinsley, H.E.A. (1989). Identification of the latent dimensions of instruments that measure perceptions of an expectations about counseling. Journal of Counseling Psychology, 36, 492-500.

Hetzel, R.D. (1996). A primer on factor analysis with comments on patterns of practice and reporting. In B. Thompson (Ed.), Advances in social science methodology (Vol. 4, pp. 175-206). Greenwich, CT: JAI Press.

Holzinger, K.J., & Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution (No. 48). Chicago: University of Chicago.

Jöreskog, K.G. (1966). Testing a simple structure hypothesis on factor analysis. Psychometrika, 31, 165-178.

Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. Psychometrika, 34, 183-202.

Jöreskog, K.G. & Sorbom, D. (1986). LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods (4th ed.). Uppsula, Sweden: University of Uppsula Department of Statistics.
maximum likelihood factor analysis. Psychometrika, 34, 183-202.

Jöreskog, K.G. & Sorbom, D. (1989). LISREL 7: A guide to the program and applications (2nd ed.). Chicago: SPSS.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. Educational and Psychological Measurement, 20, 141-151.

Knapp, T.R. (1978). Canonical correlation analysis: A general parametric significance testing system. Psychological Bulletin, 85, 410-416.

Larson, L.M., Suzuki, L.A., Gillespie, K.N., Potenza, M.A., Bechtel, M.A., & Toulouse, A.L. (1992). Development and validation of the counseling self-estimate inventory. Journal of Counseling Psychology, 39, 105-120.

Lee, R.M. & Robbins, S.B. (1995). Measuring belongingness: The social connectedness and the social assurance scales. Journal of Counseling Psychology, 42, 232-241.

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

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 6, 559-572.

Pedhazur, E.J., & Schmelkin, L.P. (1991). Measurement, design, and analysis: An integrated approach. Hillsdale, NJ: Erlbaum.

Rice, K.G., Cole, D.A., & Lapsley, D.K. (1990). Separation-individuation, family cohesion, and adjustment to college: Measurement validation and test of a Theoretical Model. Journal of Counseling Psychology, 37, 195-202.

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.

Spearman, C. (1904). "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 201-293.

Stevens, J. (1996). Applied multivariate statistics for the behavioral sciences (3rd ed.). Mahwah, NJ: Erlbaum.

Thompson, B. (1984). Canonical correlation analysis: Uses and interpretation. Thousand Oaks, CA: Sage.

Thompson, B. (1992). A partial test distribution for cosines among factors across samples. In B. Thompson (Ed.), Advances in social science methodology (Vol. 2, pp. 81-97). Greenwich, CT: JAI Press.

Thompson, B. (1997a, August). If statistical significance tests are broken/misused, what practices should supplement or replace them? Invited paper presented at the annual meeting of the American Psychological Association, Chicago.

Thompson, B. (1997b). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. Educational and Psychological Measurement, 57, 5-19.

Thompson, B., & Borrello, G.M. (1985). The importance of structure coefficients in regression research. Educational and Psychological Measurement, 45, 203-209.

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

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

Tinsley, H.E.A., Bowman, S.L., & York, D.C. (1989). Career decision scale, my vocational situation, vocational rating scale and decisional rating scale: Do they measure the same constructs? Journal of Counseling Psychology, 36, 115-120.

Tinsley, H.E.A., Roth,. J.A., & Lease, S.H. (1989). Dimensions of leadership and leadership style among group intervention specialists. Journal of Counseling Psychology, 36, 48-53.

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

Tracey, T.J., Glidden, C.E., & Kokotovic, A.M. (1988). Factor structure of the Counselor Rating Form-Short. Journal of Counseling Psychology, 35, 330-335.

Veldman, D.J. (1967). FORTRAN programming for the behavioral sciences. New York: Holt, Rinehart & Winston.

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

Table 1

Means and Standard Deviation of 13 Variables From Holzinger and Swineford, (1939)

Variable	Mean	SD	Label
T1	29.61462	7.00459	Visual Perception Test
T3	14.22924	2.83030	Paper Form Board (Spatial)
T4	18.00332	9.04784	Thorndike Lozenges (Spatial)
T7	17.36213	5.16189	Sentence Completion Test
T8	26.12625	5.67544	Word Classification
T10	96.27575	25.05927	Speeded Addition Test
T13	193.46844	36.32946	Speeded Discrimination
T15	90.00997	7.72937	Memory of Target Numbers
T16	102.52492	7.63306	Memory of Target Shapes
T17	8.23256	4.91587	Memory of Number-Object Assn.
T19	14.03654	4.07701	Memory of Figure-Word Assn.
T22	26.23920	9.19724	Math Word Problem Reasoning
T23	18.13621	9.13992	Completion of Number Series

Table 2

Correlation Matrix for Example Data

	T16	T7	T19	T10	T22	T13	T4
T16	1.00000						
T7	.16554	1.00000					
T19	.27723	.21636	1.00000				
T10	.11697	.10204	.06949	1.00000			
T22	.27768	.46978	.32246	.06934	1.00000		
T13	.27755	.22747	.19590	.34065	.24529	1.00000	
T4	.30530	.07720	.15858	.07193	.30618	.32865	1.00000
T15	.33817	-.01889	.11645	.10896	.07494	.07205	.21191
T17	.25924	.09230	.21645	.33090	.14452	.19811	.14747
T1	.36458	.29344	.18538	.06686	.39850	.39034	.44067
T23	.36969	.38528	.28307	.19040	.53497	.33216	.39663
T3	.18388	.17340	.09142	.03962	.18817	.22682	.30508
T8	.29240	.67441	.29742	.13157	.40326	.22283	.17149

	T15	T17	T1	T23	T3	T8
T15	1.00000					
T17	.30523	1.00000				
T1	.18447	.10842	1.00000			
T23	.18636	.20212	.48077	1.00000		
T3	.03616	.02634	.36529	.27570	1.00000	
T8	.05271	.10671	.33100	.42436	.21152	1.00000

Table 3

Factor Pattern Matrix

Variable	I	II	III	IV	h-squared
T1	.69068	-.02816	-.38058	.05191	.62537
T3	.43788	-.10197	-.48157	.24793	.49552
T4	.56234	.21117	-.49011	.03743	.60243
T7	.58618	-.53658	.35086	.01402	.75481
T8	.64814	-.42985	.29512	-.06185	.69577
T10	.29739	.39051	.45153	.58433	.78626
T13	.56813	.17673	-.02777	.51382	.61879
T15	.30666	.58832	-.01760	-.44019	.63425
T16	.59425	.29379	-.06148	-.29873	.53246
T17	.36849	.55471	.41229	-.02875	.61430
T19	.47415	-.00210	.25841	-.33303	.40251
T23	.75701	-.06298	-.04720	-.02451	.57986
T22	.67579	-.27936	.05491	-.14756	.55953
					(Sum =)
Eigenvalues	3.99744	1.53436	1.27791	1.09216	7.90187
% of Variance	30.7	11.8	9.8	8.4	60.8

Table 4

Factor Structure Matrix

Variable	I	II	III	IV	h-squared
T1	.69068	-.02816	-.38058	.05191	.62537
T3	.43788	-.10197	-.48157	.24793	.49552
T4	.56234	.21117	-.49011	.03743	.60243
T7	.58618	-.53658	.35086	.01402	.75481
T8	.64814	-.42985	.29512	-.06185	.69577
T10	.29739	.39051	.45153	.58433	.78626
T15	.30666	.58832	-.01760	-.44019	.63425
T16	.59425	.29379	-.06148	-.29873	.53246
T17	.36849	.55471	.41229	-.02875	.61430
T19	.47415	-.00210	.25841	-.33303	.40251
T22	.67579	-.27936	.05491	-.14756	.55953
T23	.75701	-.06298	-.04720	-.02451	.57986
T13	.56813	.17673	-.02777	.51382	.61879
					(Sum =)
Eigenvalues	3.99744	1.53436	1.27791	1.09216	7.90187
% of Variance	30.7	11.8	9.8	8.4	60.8

Table 5

Factor Correlation Matrix

	I	II	III	IV
Factor I	1.00000			
Factor II	.00000	1.00000		
Factor III	.00000	.00000	1.00000	
Factor IV	.00000	.00000	.00000	1.00000

Table 6

Factor Pattern/Structure Matrix Rotated to the Varimax Criterion

Variable	I	II	III	IV	h-squared
T1	.29746	.71254	.16817	.02997	.62537
T3	.09710	.68827	-.10955	.01932	.49552
T4	.02655	.72268	.27970	.03509	.60243
T7	.85134	.07042	-.11581	.10799	.75481
T8	.81827	.13268	.02134	.09026	.69577
T10	.04967	.00305	.05627	.88352	.78626
T13	.16236	.50177	.00800	.58360	.61879
T15	-.07740	.09731	.78645	.01706	.63425
T16	.24225	.33026	.60202	.04778	.53246
T17	.09254	-.08291	.57362	.51945	.61430
T19	.48274	-.00192	.41131	.01715	.40251
T22	.65956	.31295	.16239	-.01404	.55953
T23	.51801	.47819	.25162	.13984	.57986
					(Sum =)
Trace	2.60968	2.23006	1.70444	1.40959	7.90187
% of Variance	19.7	17.2	13.1	10.8	60.8

Note. Coefficients greater than |.30| are underlined.

Table 7

Factor Pattern Matrix Rotated to the Oblimin Criterion

Variable	I	II	III	IV	h-squared
T1	.19989	.10483	-.68259	-.01529	.62537
T3	.00964	-.15793	-.70851	-.00072	.49552
T4	-.09432	.24363	-.72701	.00121	.60243
T7	.88743	-.21232	.05608	.06136	.75481
T8	.83656	-.07057	-.00515	.03736	.69577
T10	-.01066	-.02467	.05456	.89644	.78626
T13	.06010	-.08385	-.46644	.57141	.61879
T15	-.14431	.80501	-.05699	-.01598	.63425
T16	.16866	.57154	-.26147	-.00245	.53246
T17	.03981	.53645	.16823	.50124	.61430
T19	.48319	.37328	.10672	-.03050	.40251
T22	.64195	.08869	-.21299	-.07086	.55953
T23	.45342	.17149	-.39272	.08624	.57986

Note. Communality coefficients are now computed differently. To compute the h-squared for variable T16, each pattern coefficient is multiplied by its corresponding structure coefficient and then summed across the rows. For T16, h-squared = (.16866)(.35519) + (.57154)(.64395) + (-.26147)(-.40110) + (-.00245)(.14579) = .53246 The new communality coefficient is identical to the value attained in the Table 4 analysis.

Table 8

Factor Structure Matrix Rotated to Oblimin Criterion

Variable	I	II	III	IV	h-squared
T1	.42244	.24691	-.75717	.11690	.62537
T3	.19256	-.04560	-.68666	.06272	.49552
T4	.17059	.33944	-.73684	.11229	.60243
T7	.84036	-.04333	-.18501	.16663	.75481
T8	.83084	.09461	-.25015	.16380	.69577
T10	.11483	.10405	-.04769	.88422	.78626
T13	.27784	.08913	-.54109	.62509	.61879
T15	.02280	.78407	-.13734	.09243	.63425
T16	.35519	.64395	-.40110	.14579	.53246
T17	.17288	.59560	.01139	.57057	.61430
T19	.51688	.44347	-.09308	.09348	.40251
T22	.71120	.23262	-.41116	.07385	.55953
T23	.61808	.33215	-.56630	.23490	.57986
					(Sum =)
Trace	2.57654	1.65173	2.21452	1.45908	7.90187
% of Variance	19.8	12.7	17.0	11.2	60.8

Note. Trace are computed differently than eigenvalues. To compute the trace for Factor I, each pattern coefficient is multiplied by its corresponding structure coefficient and then sum down the rows. For Factor I, Trace = (.16866)(.35519)+(.88743)(.84036)+(.48319)(.51688)+(-.01066)(.11483)+ (.64195)(.71120)+(.06010)(.27784)+(-.09432)(.17059)+(-.14431)(.02280)+ (.03981)(.17288)+(.19989)(.42244)+(.45342)(.61808)+(.00964)(.19256)+ (.83656)(.83084)= 2.57654. The trace still sum to 7.90187, the sum of communality coefficients in the Table 4 analysis.

Table 9

Factor Correlation Matrix after Rotation to the Oblimin Criterion

	I	II	III	IV
Factor I	1.00000			
Factor II	.18956	1.00000		
Factor III	-.30059	-.15612	1.00000	
Factor IV	.16350	.15535	-.12193	1.00000

Table 10

Selected Fit Statistics for Example CFA Analysis

	Model #1	Model #2
v	13	13
n	301	301
Null chi sq	1009.30	1009.30
Null df	78	78
Noncentrality	931.3	931.3
Model chi sq	165.16	240.42
Model df	58	60
Noncentrality	107.16	180.42[a]
NC / df	1.847586	3.007[b]
GFI	0.960	0.899
AGFI	0.877	0.829
Parsimony Ratio	0.637362	0.659340[c]
GFI*Pars Ratio	0.611868	0.592747[d]
CFI	0.884935	0.806270[e]
Parsimony Ratio	0.743589	0.769230[f]
CFI*Pars Ratio	0.658028	0.620208[g]
RMSR	20.717	20.741
RMSEA	0.006158	0.010023[h]

[a]Noncentrality = χ^2 - df

[b]Noncentrality / df

[c]Parsimony Ratio = Model df / [(variables * (variables + 1)) / 2]

[d]GFI * Parsimony Ratio

[e]$$CFI = \frac{[(\text{Null}\ \chi^2 - \text{Null df}) - (\text{Model}\ \chi^2 - \text{Model df})]}{(\text{Null}\ \chi^2 - \text{Null df})}$$

[f]Parsimony Ratio = Model df / [variables * (variables - 1)) / 2]

[g]CFI * Parsimony Ratio

[h]RMSEA = [Model χ^2 - Model df) / (Model df * (n-1))]$^{.5}$

Table 11

Maximum Likelihood Estimates for Model #1

	Factor				
Variable	Verbal	Visual	Speed	Memory	Math
T1	0.000	0.757	0.000	0.000	0.000
T3	0.000	0.458	0.000	0.000	0.000
T4	0.000	0.608	0.000	0.000	0.000
T7	0.850	0.000	0.000	0.000	0.000
T8	0.790	0.000	0.000	0.000	0.000
T10	0.000	0.000	0.635	0.000	0.000
T13	0.000	0.000	0.480	0.000	0.000
T15	0.000	0.000	0.000	0.438	0.000
T16	0.000	0.000	0.000	0.677	0.000
T17	0.000	0.000	0.000	0.455	0.000
T19	0.000	0.000	0.000	0.434	0.000
T22	0.000	0.000	0.000	0.000	0.717
T23	0.000	0.000	0.000	0.000	0.749

Table 12

Phi (Factor Correlation) Matrix for Model #1

	Factor				
Factor	Verbal	Visual	Speed	Memory	Math
Verbal	1.000				
Visual	0.425	1.000			
Speed	0.328	0.450	1.000		
Memory	0.360	0.606	0.544	1.000	
Math	0.693	0.786	0.461	0.644	1.000

Table 13

Maximum Likelihood Estimates for Model #2

	Factor			
Variable	Verbal	Visual	Speed	Math
T1	0.000	0.780	0.000	0.000
T3	0.000	0.465	0.000	0.000
T4	0.000	0.581	0.000	0.000
T7	0.679	0.000	0.000	0.000
T8	0.704	0.000	0.000	0.000
T10	0.000	0.000	0.637	0.000
T13	0.000	0.000	0.480	0.000
T15	0.000	0.000	0.000	0.463
T16	0.000	0.000	0.000	0.713
T17	0.000	0.000	0.000	0.436
T19	0.411	0.000	0.000	0.000
T22	0.686	0.000	0.000	0.000
T23	0.701	0.000	0.000	0.000

Table 14

Phi (Factor Correlation) Matrix for Model #2

	Factor			
Factor	Verbal	Visual	Speed	Math
Verbal	1.000			
Visual	0.667	1.000		
Speed	0.439	0.441	1.000	
Math	0.512	0.592	0.519	1.000

Table 15

Selected EFA and CFA Studies from The Journal of Counseling Psychology

Author	Year	Type of Analysis
Cole & Jordan	1989	CFA
Good et al.	1995	CFA
Hayes & Tinsley	1989	EFA
Larson et al.	1992	EFA
Lee & Robbins	1995	CFA
Rice, Cole & Lapsley	1990	EFA
Sodowsky, Taffe, Gutkin & Wise	1994	EFA
Tinsley, Bowman & York	1989	EFA
Tinsley, Roth & Lease	1989	CFA
Tracey, Glidden & Kokotovic	1988	CFA

Table 16

Analysis of Select Articles Based on Recommendations by Hetzel (1996) and Thompson and Daniel (1996)

Information	% of Studies Reporting		
	Total	CFA	EFA
Coefficients Termed "Loadings"	100%	100%	100%
Criteria for Factor Retention	50%	___	100%
Eigenvalues	50%	___	100%
Factors Correlated	70%	60%	80%
Final Communality Coefficients	30%	___	60%
Initial Communality Estimates Used	10%	___	20%
Interpret Several Fit Statistics			
Item Means and Variances	10%	___	20%
Matrix of Association			
Correlation Matrix	30%	40%	20%
Variance-Covariance Matrix	10%	20%	0%
Method of Factor Extraction	40%	___	80%
Method of Factor Rotation	50%	___	100%
Method of Sample Selection	60%	80%	40%
Multiple Models Specified	40%	80%	___
Number of Variables	80%	80%	80%
Reliability Estimate Provided	40%	40%	40%
Rotated Factor Pattern Matrix	30%	___	60%
Rotated Factor Structure Matrix	30%	___	60%
Sample Size	100%	100%	100%
Sample Composition	100%	100%	100%

Note. Blank line indicates that information does not pertain to the particular type of analysis.

AERA

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: SOME COMMENTS ON ANALYTIC TRADITIONS IN EFA AS AGAINST CFA: AN ANALYSIS OF SELECTED RESEARCH REPORTS	
Author(s): KEVIN M. KIEFFER	
Corporate Source:	Publication Date: 4/13/98

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] Sample sticker to be affixed to document

Check here

Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

KEVIN M. KIEFFER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document [ ]

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

or here

Permitting reproduction in other than paper copy.

Level 2

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature:	Position: RES ASSOCIATE
Printed Name: KEVIN M. KIEFFER	Organization: TEXAS A&M UNIVERSITY
Address: TAMU DEPT EDUC PSYC COLLEGE STATION, TX 77843-4225	Telephone Number: (409) 845-1831
	Date: 3/18/98