DOCUMENT RESUME

ED 325 531 TM 015 820

AUTHOR	Daniel, Larry G.
TITLE	Common Factor Analysis or Components Analysis: An Update on an Old Debate.
PUB DATE	Nov 90
NOTE	30p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (19th, New Orleans, LA, November 14-16, 1990).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Comparative Analysis; Elementary Secondary Education; *Factor Analysis; Literature Reviews; *Research Methodology; *Statistical Analysis; Teachers
IDENTIFIERS	*Principal Components Analysis

ABSTRACT

Both common factor analysis and components analysis are useful techniques for reducing the number of variables in a data set or for identifying underlying covariance structures that exist among a set of variables. Although researchers have for some years debated the appropriateness of selecting one of these methods over the other, components analysis has traditionally been the dominant strategy among educational researchers. Following a brief overview of the logic of factor analysis, a review of several studies comparing the common factor and principal components methods is presented. Actual educational research data are used to demonstrate cases in which the two methods will produce different results. Data used in these analyses were collected from 70 preservice and inservice teachers using a 3-subscale "logic of confidence" measure developed by K. R. Okeafor and others (1987). Guidelines are presented to assist the researcher in determining which method to use in specific research situations. Six tables illustrate the analyses. A 49-item list of references is included. (Author/SLD)

facvcomp.lgd

ED325531

COMMON FACTOR ANALYSIS OR COMPONENTS ANALYSIS: AN UPDATE ON AN OLD DEBATE

Larry G. Daniel

University of Southern Mississippi

Paper presented at the annual meeting of the Mid-South Educational Research Association, November 14-16, 1990, New Orleans.

ABSTRACT

Both common factor analysis and components analysis are useful techniques for reducing the number of variables in a data set or for identifying underlying covariance structures that exist among a set of variables. Although researchers have for some years debated the appropriateness of selecting one of these methods over the other, components analysis has traditionally been the dominant strategy among educational researchers.

Following a brief, simplified overview of the logic of factor analysis, a review of a number of studies comparing the common factor and principal components methods is presented. Actual educational research data are used to demonstrate cases in which the two methods will produce different results. In addition, guidelines are offered to aid the researcher in determining which method to employ in specific research situations.

COMMON FACTOR ANALYSIS OR COMPONENTS ANALYSIS:
AN UPDATE ON AN OLD DEBATE

Factor analysis is an elegant statistical technique that may be used by the social scientist both in theory development and in the validation of measures of human behavior and abilities. Consequently, factor analysis has been described as "one of the most powerful tools yet devised for the study of complex areas of behavioral scientific concern" (Kerlinger, 1986, p. 689), and as "the furthest logical development and reigning queen of the correlational methods" (Cattell, 1978, p. 4). Factor analytic methods allow the researcher to reduce a set of observed variables to a smaller set of latent variables expressing common dimensions of the observed variable set. Nunnally (1967, p. 289) noted that this data reductive property of factor analysis "is useful in the explication of constructs." Likewise, Harman (1967, p. 4) noted, "The chief aim [of factor analytic methods] is to attain scientific parsimony or economy of description."

Cattell (1978, p. 5) has proposed a two-pronged application of factor analysis in the behavioral sciences:

> One important use of factor analysis is in finding the "significant dimensions". . .in a jungle of variables. . . .A second major service of factor analysis is in exposing and counteracting that vice of the human mind which. . .assumes that when there is one word [for a given construct] there must be one thing corresponding to it in the real world.

Despite the general agreement among social science researchers (e.g., Bookstein, 1990; Cattell, 1952, 1978; Gorsuch, 1983, 1990; Harman, 1967; McArdle, 1990; Mulaik, 1972, 1990; Nunnally, 1967; Steiger, 1990; Velicer &

Jackson, 1990) that factor analysis is an important research tool, there is considerable controversy regarding the appropriateness of various techniques for determining factor solutions. More specifically, the controversy centers around whether principal components analysis (PCA) or common factor analysis (CFA) is the most appropriate technique to use in research situations employing structural factor analysis.

The purpose of the present study is to review some of the arguments of the proponents of PCA and CFA, and to offer some guidelines for the appropriate choice of method in various research situations. A brief overview of the logic underlying factor analysis along with a simple heuristic example is provided for the novice reader. A review of various empirical studies which have compared PCA and CFA is presented. Finally, several factor analyses utilizing actual educational research data are presented to demonstrate cases in which the two methods will produce similar and different results.

The Logic of Factor Analysis--A Brief Overview

Social scientists (e.g., Cattell, 1952, 1988a; Rummel, 1970) have conceived of a three-dimensional model (often referred to as a "data box" or "data cube") for measuring and describing any given psychological or ideological phenomenon. The three dimensions (called modes) which constitute this model are generally considered to be persons, variables, and occasions of measurement (Cattell, 1952). Factor analytic techniques usually involve two of these three modes, one of which is factored across the other.

In conducting an "R-technique" factor analysis, the most commonly-used factor analytic technique, the researcher first selects a finite set of *p* variables from a universe of possible variables designed to measure a specific construct. The choice of variables may be based on theoretical considerations or on the researcher's own hypothetical notions regarding the nature of the construct. Data are collected on these variables from a sample of persons thought to be representative of the researcher's population of interest. A square ($p \times p$) matrix of association (correlation matrix) is constructed to determine the intercorrelations among the p variables. On the basis of these correlations, a new rectangular correlation matrix is constructed with the p variables serving as the rows and m common factors serving as the columns.

The researcher's goal in R-technique factor analysis is to reproduce as much of the variance in the original set of p variables as possible in m interpretable factors ($m < p$). By examining the content of the variables which correlate most highly with each factor (determined by consulting the factor structure coefficients or "loadings" of each variable with each of the factors), the researcher then attempts to give a name to each of the m interpretable factors underlying the original variable set. If the researcher's goal is theory development, the resultant factors serve as indicators of the various dimensions underlying the construct the researcher is endeavoring to measure. If the goal is construct validation of a psychometric instrument, the factors indicate the various constructs the instrument is purported to measure.

In "Q-technique" factor analysis, the most commonly-used alternative to R, the same two dimensions (variables and people) are used as in R-technique although they are reversed (Comrey, 1973), i.e., the people are factored across the variables, yielding "person factors," or clusters of persons who behave or think differently than persons in other clusters in terms of the constructs being measured. Although the present study is limited to discussion of the R-technique factor analysis, detailed discussions of the logic underlying Q-technique are provided elsewhere (Kerlinger, 1986; Stephenson, 1953). A number of research studies in which Q-technique factor analytic methods have been employed are also available for further study by the interested reader (e.g., Carr, 1989; Daniel, 1989b; Thompson, 1980; Townsend, 1987).

As a heuristic example of the logic underlying R-technique factor analysis, consider a hypothetical research situation involving 10 variables collected on each of 15 subjects, as presented in Table 1. The square inter-item correlation matrix of association for these 10 variables is presented in Table 2, and the rectangular "factor solution" matrix is presented in Table 3.

INSERT TABLES 1, 2, AND 3 ABOUT HERE

An analysis of the factor structure coefficients of the variables with the identified factors as shown in Table 3 indicates that two distinct factors are being measured, with Items 1, 5, 6, 7, and 9 correlating highly with Factor I, and the remaining items correlating highly with Factor II. In an actual research situation, the next two logical steps would be to examine the items identified with each factor, and then to name the factors.

Methods for Determining Factor Solutions

The most difficult (and perhaps the most subjective) part of conducting a factor analysis is determining the factor solution. Indeed, some have argued that factor analysis is so subjective that it should be used with extreme caution. For instance, Armstrong (1967) in an amusing article subtitled "Tom

Swift and His Electric Factor Analysis Machine" warned of the danger of routinely using exploratory factor analysis to drive theory development in lieu of forming _a priori_ theoretical assumptions about the relationships among variables. However, as explained by Daniel (1989a), these exploratory factor analytic methods can be followed up with _confirmatory_ methods which can directly take into account the _a priori_ relationships among the variables.

At least three processes are involved in determining a factor solution: (a) selecting from among several mathematical models for determining the values of factor structure coefficients, (b) determining if and how the factor results should be rotated, and (c) deciding on the number of factors to extract and interpret. In the present study, only the first of these three processes is considered. However, a considerable amount of writing has also been devoted to the issues of factor rotation (e.g., Cattell, 1978; Gorsuch, 1988; McDonald, 1985; Mulaik, 1972; Nunnally, 1978) and number of factors (e.g., Cattell, 1966; Kaiser, 1960; Linn, 1968; Zwick & Velicer, 1982, 1986).

Historically, there has been considerable controversy regarding the appropriateness of various techniques for determining factor solutions. Spearman's early ideas about mental abilities led to his development of the "_G_ factor" theory (Spearman, 1904), i.e., the notion that all mental abilities are ultimately explained by a single general or _G_ factor. Nunnally (1967, p. 334) has explained the _G_ factor theory as follows:

> The general factor was thought of as a type of mental yardstick of intelligence, and only one yardstick was thought necessary to explain the common ground among all forms of individual differences in abilities. Thus tests as diverse as tests of arithmetic, spelling, and the judgment of illusions were thought

> to share in G. In addition, it was theorized that each source of individual differences (test) possessed a unique factor. . . . The theory is sometimes called Spearman's two-factor theory, because it hypothesized that each test could be explained by a general factor and a unique factor. (emphasis in original)

Holzinger extended the G factor theory in his development of the "bifactor" solution (Holzinger, 1941; Holzinger & Swineford, 1937). This theory hypothesizes that the common variance among a series of tests (or variables) can be accounted for by both a general factor and by two or more "group" factors. Hence, all of the group factors correlate with the G factor, but not with one another. Nunnally (1967, p. 339) provides a simplified example of this factor solution.

By the 1930's, statisticians were beginning to develop the more sophisticated factor analytic models used most frequently today. Hotelling (1936) proposed methods for calculating the principal components solution. Principal components analysis (PCA) yields linear combinations of observed variables based upon a referent square matrix of correlation among the variables with unity values on the matrix diagonal. PCA assumes that the inter-variable correlation matrix produced by the data from a given sample perfectly reflects the population correlation matrix. In other words, PCA assumes that the variables have been collected without error (Gorsuch, 1988). The analysis yields n principal components from a set of n variables (i.e., the number of components equals the number of variables). Generally, the researcher selects from among these n factors the first several which account for a major portion of the variance. The selected factors are interpreted and named, whereas factors with negligible contributions to the variance are ignored. The result is a "truncated" component solution (Gorsuch, 1988).

In common factor analysis (CFA), the unity diagonal values on the inter-variable correlation matrix are replaced with communality (h^2) estimates of each variable. As noted by Gorsuch (1978, p. 24), variable communalities are the correlations of each variable with the "common elements" underlying the variable set. One frequently-used mathematical analog for determining lower-bound estimates of the communalities is to compute the squared multiple R of each variable with all the other variables. Since error variance is included in any data set, the communality estimates are less than unity in value.

CFA involves the computation of both common and unique factors. For n variables, there will be n unique factors that explain variance unique to each variable. CFA generally allows the researcher to extract k common factors for the n variables such that $1 \leq k \leq n$. As noted by Cureton and D'Agostino (1983), the number of common factors is generally less than half the number of variables ($k \leq n/2$). As each factor is extracted, a resultant "residual" matrix is calculated. This matrix expresses the relationships among the variables with the effect of the first factor removed from the analysis. In the "pure" mathematical model of CFA, factors are extracted until "some statistical test says the last residual matrix is essentially full or zeros" (Cattell, 1978, p. 30).

As previously noted, the controversy surrounding various factor analytic methods has yielded many diverse opinions of the viability of one method over another. Fairly early in the development of CFA and PCA, Cureton (1939) responded to the ongoing debates among the "two-factorists," "bi-factorists," "multiple-factorists," and "component analysts":

Factor theory may be defined as mathematical rationalization. A

> factor-analyst is an individual with a peculiar obsession regarding the nature of mental ability or personality. By the application of higher mathematics to wishful thinking, he always proves that his original fixed idea or compulsion was right or necessary. In the process he usually proves that all other factor-analysts are dangerously insane, and that the only salvation for them is to undergo his own brand of analysis in order that the true essence of their maladies may be discovered. (p. 287)

Interestingly, by 1967, one researcher stated, "The heated and inspired controversies about the 'best' method of factor analysis are over" (Harman, 1967, p. 9). However, even though the old G factor and bi-factor theories are now outdated, the controversy over common factor analysis and principal components analysis is very much alive. In fact, Bentler and Kano (1990) recently noted that the controversy is approximately 50 years old.

One reason for the controversy is the affinitive relationship between the two methods. Gorsuch (1990, p. 33) notes the similarity and distinction between the two:

> Common factor analysis is the general case of which component analysis is a special case. Common factor analysis includes variables with error and variables without error because, in the latter case, certain elements become zeros. Component analysis limits this broader model by an additional assumption: the variables are reproduced without error (i.e., without uniquenesses or residuals).

According to Acito and Anderson (1980) and Gorsuch (1983), principal components analysis is most useful when the number of variables is large and when variable communalities are relatively high (above 0.40) and homogeneous in value. As demonstrated by Gorsuch (1983), Harris (1970), and Velicer (1974, 1976, 1977), when these two assumptions are met, PCA and CFA yield nearly identical results. "However, with a smaller number of variables, the values in the diagonal assume greater importance and the loadings in the resulting factor pattern [using PCA] may tend to be inflated" (Acito & Anderson, 1980, p. 230).

Considering the problems associated with PCA, many have suggested that PCA should not be routinely used as a factor analytic procedure (e.g., Cureton, 1939; Cattell, 1988b; Gorsuch, 1990; Harman, 1967; McDonald, 1985; Mulaik, 1990). Cattell (1988b), for instance, notes that for PCA to be viable, the set of variables under consideration "would have to lie in a completely self-explanatory subuniverse, self-sufficient as a system entirely isolated from the rest of the universe. The components model must be considered a mere mathematical figment" (p. 134). Similarly, McArdle (1990, p. 81) notes that "PCA is a poor substitute for a more complete and reliable CFA." Despite these sentiments, PCA remains the preferred method in behavioral research (Pruzek & Rabinowitz, 1981). This preference may be due in part to the setting of PCA as the default factor extraction method in many popular statistical software packages such as SPSSX (Hubbard & Allen, 1987).

In opposition to the many writers who prefer CFA, there are a number of researchers who adamantly defend the viability of PCA. Chief among the defenders of PCA are Velicer and his associates. Velicer has sought through results of a host of empirical studies to show that PCA and CFA rarely produce notably different results (e.g., Velicer, 1974, 1976, 1977; Velicer &

Jackson, 1990). Wilkinson (1989) concurs with the findings of Velicer, and further notes that most attempts to discredit the merits of PCA focus upon mathematical artifacts associated with contrived example data sets. Wilkinson affirms that when analyzing real data, PCA and CFA generally do not differ enough to matter.

Others argue for the preference of PCA over CFA on the basis of parsimony. PCA is more mathematically direct and simple, thus it requires less computational energy (and less consumption of computer processing units). Gorsuch (1990), however, points out that elegance in scientific models does not necessarily always mean mathematical simplicity.

Gorsuch (1988) and Snook and Gorsuch (1989) note that preference of PCA over CFA, or vice-versa, is often a matter of paradigmatic biases of the researcher. PCA, notes Gorsuch (1988), follows a mathematical paradigm, i.e., it stresses derivation and exact computation of all procedures. CFA, on the other hand, follows a scientific, or statistical, paradigm. This latter paradigm stresses the value of estimation and replication of results across various samples. Preference for the CFA model is noted by McArdle (1990), who states, "The PCA always overestimates the true loading, and the accuracy of PCA becomes very poor as the true loadings get smaller. In these math models the CFA algorithms are always 100% accurate."

<u>Empirical Comparisons of the Two Methods</u>

A number of researchers have sought to empirically test the convergence of factor solutions offered by PCA and CFA. A number of these studies (e.g., Acito & Anderson, 1980; Borgatta, Kercher, & Stull, 1986; Linn, 1968; Snook & Gorsuch, 1989; Velicer, Peacock, & Jackson, 1982) have involved the factor analyzing of Monte Carlo generated of data. Acito and Anderson (1980) compared principal components analysis with two common factoring techniques (image and alpha factoring) using generated data sets of 50, 100, and 300 cases. The results suggested that image and alpha analyses were superior to principal components analysis in reproducing known factor patterns among variables. In fact, the Acito and Anderson analyses suggest that image analysis with 50 cases is almost as accurate as principal components analysis with 300 cases. In a similar study, Snook and Gorsuch (1989) found that common factor analysis reproduced known factor structures more accurately than principal components analysis when factor patterns identified with generated data were compared with population factor patterns.

Other comparative studies have involved actual data sets. For instance, Velicer (1977) compared two CFA techniques with PCA by reanalyzing data from nine studies. All the studies employed rather large sample sizes and large numbers of variables. Velicer's comparisons were based on observing similarities and differences in the factor patterns across the methods and on a summary statistic g based on summation of the squared differences in comparable factor structure coefficients across the methods. Velicer concluded that the patterns produced across the methods were remarkably similar, with the greastest differences noted in the last (weakest) factor extracted. A similar comparison study by Wilkinson (1989) confirms the Velicer findings. Wilkinson concludes that most differences between CFA and PCA are the result of "overfactoring" of data, i.e., extracting of too many meaningless factors (p. 456).

Hubbard and Allen (1987) also compared PCA with CFA using actual data, but their findings were in sharp contrast to those of Velicer and Wilkinson.

Hubbard and Allen provide empirical evidence across several studies that illustrates differences in the magnitudes of the factor structure coefficients using PCA and CFA. In several of the cases, these shifts in coefficients lead to differences in the interpretation of the factors. Similarly, Gorsuch (198ˆ‹ illustrated several factor analytic cases using a small number of variables in which the PCA and CFA methods yielded notably different results.

Without a doubt, the issue of which of these methods, if either, is superior to the other is far from being resolved. Opponents of PCA have often used either hypothetical data or else have used a rather small number of variables as weapons in their factor analytic arsenal. It could be argued that either of these types of data may lead to distortion of factor analytic results. On the other hand, PCA advocates often try to explain away the CFA advocates' arguments by suggesting that factor indeterminacy is responsible for most of the differences in the two methods.

Consequently, an empirical example is offered as a part of the present study in order to review and further explore these differences. An actual data set (Daniel & Okeafor, 1987) involving teachers' perceptions about their own professionalism on a "logic of confidence" scale was factor analyzed using both PCA and CFA. This data set serves as an interesting factor analytic example as attitudinal measures are frequently used in educational research.

Method

The data utilized in the following analyses were collected from 70 preservice and inservice teachers using a three-subscale "logic of confidence" measure authored by Okeafor, Licata, and Ecker (1987). This self-evaluation measure included 29 Likert-type items. The illustrated data set borders on being a "problem" data set for factor analytic purposes as the number of variables and subjects are below minimal criteria as established by Gorsuch (1983) and Comrey (1973).

Responses of the 70 subjects on the logic of confidence measure were factor analyzed using principal components analysis and principal axis factor analysis. Analyses were run using the SPSSx FACTOR procedure. Two alternate rules were used for determining the number of factors to extract in comparing the results of the analyses. The "eigenvalue greater than unity" rule (Guttman, 1954), was selected as it is a commonly used rule and as it is the default factor extraction method for many statistical software programs such as SPSSx. In a second round of analyses, Cattell's (1966) "scree" test was utilized for determining the number of factors. The scree test relies on a visual plot of the eigenvalues, with the number of factors to extract determined by a break in the curve of the line formed by the plotted values. Generally, scree is a more conservative test, suggesting that response variance can be adequately accounted for in relatively few factors. All results were rotated to simple structure using the varimax criterion.

Findings

Principal components and principal axis factor analyses using the default "eigenvalue greater than unity" extraction rule were run for the data at hand. Ten factors were extracted for each solution. Initial variable communality estimates for the principal factor solution were generally high, with only one of the 29 values falling into the "low" (< .40) range as defined by Gorsuch (1983), and with nearly all of the communalities (23 of 29) above .50.

Velicer (1977), Snook and Gorsuch (1989), and Hubbard and Allen (1987) propose several statistics by which various factor patterns may be compared.

However, the ability of a model to identify a viable factor structure is the most stringent test of a model, and, indeed, the aspect of the model of most interest to researchers employing factor analysis in empirical research. Hence, the more highly statistical means of comparing the two models were rejected for use in the present study in lieu of direct comparison of identified factor structures using the "eyeball" strategy. More specifically, the two models were evaluated and compared on the degree to which they produced factors which were uniquely saturated with given items. For the purposes of these comparisons, an item was considered uniquely associated with a factor if (a) it had a higher factor structure coefficient on that factor than on any other factor, and (b) its coefficient on that factor exceeded $|.40|$.

The "unique" items associated with these factors, along with the items' factor structure coefficients, are presented in Table 4. The two analyses yielded nearly identical results up through the fourth factor. Factors I and III, IV, and VI for both analyses included the exact same sets of items, and Factors II, and V varied by only one item each across the two solutions. Factors VIII and IX for the PCA analysis matched Factors VII and VIII, respectively, for the PAF analysis, and PCA Factor X matched PAF Factor IX with the exception of the former including one additional item. What is most notably different in the two analyses is the magnitude of the factor structure coefficients.

Also, each analysis yielded one rather unique factor. The most unique PCA factor was Factor VII, while Factor X was the most unique PAF factor. It could be effectively argued, however, that the last five factors on both analyses are hardly interpretable as most of them correlated adequately with only one or two items. These results suggest a problem with overextraction.

The "scree" test indicated a clear break between Factors III and IV; hence, a three-factor solution was run using each of the extraction methods. The resultant PCA and PAF factor matrices are presented in Tables 5 and 6, respectively. The factor structures are very similar. With the exception of Item 20, each of the items correlated most highly with the same factor across the two analyses. Again, the most notable difference between the two matrices was the magnitude of the structure coefficients. Since researchers rarely interpret items as defining a construct when structure coefficients are less than |.30|, a number of items might be omitted from interpretation if employing the PAF extraction method in the present research situation.

Discussion

The results of the foregoing analyses tend to confirm the assumptions (a) that the choice between PCA and CFA generally does not make a lot of difference, and (b) that differences shown by the two methods tend to be based on overextraction. The comparative ten-factor solutions using the two methods produced relatively similar results up until the seventh factor. By that point in the analyses, none of the remaining factors in either analysis was uniquely saturated with more than two items, making interpretation of these factors almost meaningless. Hence, the results would indicate that the "eigenvalue greater than unity" rule led to overextraction in both models.

The more viable three-factor solutions also yielded very similar results, suggesting that either method might yield an acceptable factor solution for the data at hand. However, these analyses illustrate that the magnitudes of the structure coefficients may be impacted upon by the choice of extraction method. In the present example, using an item saliency criterion of |.30|, the researcher would interpret seven more items with PCA than with PAF.

Thus, the choice of CFA or PCA is largely unimportant in terms of the models' abilities to produce factor structures, but may lead to differences in interpretation of results. However, the data employed in the present analysis may not be typical of data encountered in educational research. Despite the problems associated with the size of the sample and to some extent the number of variables in the set, the variables' communalities were relatively high, which may have accounted for the similarity of results across the two methods.

When employing a data set with low variable communalities, the researcher is generally advised to use CFA over PCA. However, it might be questionable whether the researcher should use _any_ type of factor analytic technique if a substantial number of the variables have low communalities. Communalities provided lower-bound estimates of the reliability of the variables (Cattell, 1978). Thus, if many of these variables have low communalities, the integrity of the instrument(s) being used is already at stake, and rather than representing viable constructs, the resulting factors are likely to be no more than mathematical artifacts.

Generally, CFA is the safer of the two techniques. This is especially true in cases in which there are only a few variables or in which variable communalities are poor. Nevertheless, the researcher would be amiss to blindly employ CFA without examining the variable communalities. Despite the fact that the two methods generally yield very similar results, "common factor analysis provides a fail-safe procedure in terms of more accurately reflecting the population factor pattern because it gives accurate results when component analysis produces biased results as well as when both methods produce reasonably accurate results" (Snook & Gorsuch, 1989, p. 153).

REFERENCES

Acito, F., & Anderson, R. D. (1980). A Monte Carlo comparison of factor analytic methods. Journal of Marketing Research, 17, 228-236.

Armstrong, J. S. (1967). Derivation of theory by means of factor analysis or Tom Swift and his electric factor analysis machine. American Statistician, 21(5), 17-21.

Bentler, P. M., & Kano, Y. (1990). On the equivalence of factors and components. Multivariate Behavioral Research, 25, 67-74.

Bookstein, F. L. (1990). Least squares and latent variables. Multivariate Behavioral Research, 25, 75-80.

Borgatta, E. F., Kercher, K., & Stull, D. E. (1986). A cautionary note on the use of principal components analysis. Sociological Methods & Research, 15, 160-168.

Carr, S. C. (1989, March). Effects of different data collection strategies on Q-technique factor analytic results. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Cattell, R. B. (1952). Factor analysis: An introduction and manual for the psychologist and social scientist. New York: Harper & Brothers.

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.

Cattell, R. B. (1988a). The data box: Its ordering of total resources in terms of possible relational systems. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate experimental psychology (2nd ed.) (pp. 69-130). New York: Plenum.

Cattell, R. B. (1988b). The meaning and strategic use of factor analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate experimental psychology (2nd ed.) (pp. 131-203). New York: Plenum.

Comrey, A. L. (1973). A first course in factor analysis. New York: Academic Press.

Cureton, E. E. (1939). The principal compulsions of factor-analysts. Harvard Educational Review, 9, 287-295.

Cureton, E. E., & D'Agostino, R. B. (1983). Factor analysis: An applied approach. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Daniel, L. G. (1989b, November). Stability of q-factors across two data collection methods. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED 314 439)

Daniel, L. G., & Okeafor, K. R. (1987, November). Teaching experience and confidence in teachers. Paper presented at the annual meeting of the Mid-South Educational Research Association, Mobile, AL. (ERIC Document Reproduction Service No. ED 292 763)

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

Gorsuch, R. L. (1988). Exploratory factor analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of multivariate experimental psychology (2nd ed.) (pp. 231-258). New York: Plenum.

Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. Multivariate Behavioral Research, 25, 33-39.

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

Harman, H. H. (1967). Modern factor analysis (2nd ed.). Chicago: University of Chicago Press.

Holzinger, K. (1941). Factor analysis. Chicago: University of Chicago Press.

Holzinger, K., & Swineford, F. (1937). The bi-factor method. Psychometrika, 2, 41-54.

Hotelling, H. (1936). Simplified calculation of principal components. Psychometrika, 1, 27-35.

Hubbard, R., & Allen, S. J. (1987). A cautionary note on the use of principal components analysis: Supportive empirical evidence. Sociological Methods & Research, 16, 301-308.

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

Kerlinger, F. A. (1986). Foundations of behavioral research (3rd ed.). New York: Holt, Reinhart, and Winston.

Linn, R. L. (1968). A Monte Carlo approach to the number of factors problem. Psychometrika, 33, 37-71.

McArdle, J. J. (1990). Principles versus principals of structural factor analysis. Multivariate Behavioral Research, 25, 81-87.

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

Mulaik, S. A. (1990). Blurring the distinctions between component analysis and common factor analysis. Multivariate Behavioral Research, 25, 53-59.

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

Okeafor, K. R., Licata, J. W., & Ecker, G. (1987). Toward an operational definition of the logic of confidence. Journal of Experimental Education, 56, 47-54.

Rummel, R. J. (1970). Applied factor analysis. Evanston, IL: Northwestern University Press.

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

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

Steiger, J. H. (1990). Some additional thoughts on components, factors, and factor indeterminacy. Multivariate Behavioral Research, 25, 41-45.

Stephenson, W. (1953). The study of behavior: Q-technique and its methodology. Chicago: University of Chicago Press.

Thompson, B. (1980). Validity of an evaluator typology. Educational Evaluation and Policy Analysis, 2, 59-65.

Townsend, J. (1987). Media selection in medical education: The development of a frame of reference. Unpublished doctoral dissertation, University of New Orleans, New Orleans.

Velicer, W. F. (1974). A comparison of the stability of factor analysis, principal component analysis, and rescaled image analysis. Educational and

Psychological Measurement, 34, 563-572.

Velicer, W. F. (1976). The relation between factor score estimates, image scores, and principal component scores. Educational and Psychological Measurement, 36, 149-159.

Velicer, W. F. (1977). An empirical comparison of the similarity of principal component, image and factor patterns. Journal of Multivariate Behavioral Research, 12, 3-22.

Velicer, W. F., Peacock, A. C., & Jackson, D. N. (1982). A comparison of component and factor patterns: A Monte Carlo approach. Multivariate Behavioral Research, 17, 371-388.

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

Wilkinson, L. (1989). A cautionary note on the use of factor analysis: A response to Borgatta, Kercher, and Stull, and Hubbard and Allen. Sociological Methods & Research, 17, 449-459.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. Multivariate Behavioral Research, 17, 253-269.

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

Table 1
Hypothetical Data Set for Factor Analysis Example

CASE	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10
1	2	1	1	1	2	1	1	1	2	1
2	2	2	2	2	1	2	2	1	2	2
3	2	1	2	1	2	1	2	1	2	1
4	1	1	2	1	1	2	1	1	1	2
5	2	1	2	1	2	1	2	1	2	2
6	1	2	2	2	2	1	2	2	2	2
7	2	1	2	1	2	1	2	2	2	2
8	2	1	1	1	1	2	1	1	2	1
9	2	1	2	2	2	2	1	1	1	2
10	2	1	2	1	2	2	1	2	2	2
11	2	1	2	1	2	2	1	2	2	2
12	1	1	1	1	1	2	1	1	1	2
13	2	1	2	1	2	1	2	1	2	2
14	2	1	2	1	1	2	1	1	1	2
15	2	1	2	1	2	1	2	1	2	2
16	2	1	1	2	2	1	2	1	2	2
17	2	1	2	2	1	2	1	2	2	2
18	1	1	1	2	1	2	1	1	1	2
19	2	2	2	2	2	2	1	2	2	2
20	1	1	2	1	1	2	1	2	1	2

Table 2
Inter-Item Correlation for Table 1 Data

	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9
ITEM2	-.08085								
ITEM3	.20000	.24254							
ITEM4	-.06052	.57248	-.06052						
ITEM5	.47140	.05717	.23570	-.04280					
ITEM6	-.23570	.05717	.00000	.17118	-.66667				
ITEM7	.23570	.22866	.23570	.04280	.45833	-.79167			
ITEM8	-.06052	.27890	.42366	.12088	.17118	.17118	-.17118		
ITEM9	.62994	.27501	.12599	.02288	.57907	-.53452	.53452	.25163	
ITEM10	-.24254	.17647	.40423	.30826	-.05717	.22866	.05717	.30826	-.27501

Table 3
Factor Solution Matrix for Factor Analysis Example[1]

	FACTOR1	FACTOR2
ITEM1	.63649	-.09264
ITEM2	.12548	.71718
ITEM3	.23546	.59894
ITEM4	-.09758	.59066
ITEM5	.80878	.09956
ITEM6	-.83154	.21605
ITEM7	.76894	.10756
ITEM8	.02084	.56631
ITEM9	.84367	.13154
ITEM10	-.24024	.66871

[1] These factors were extracted using principal components method, and were rotated to the varimax criterion.

Table 4
Items Associated with Factors Across Two Ten-Factor Solutions[1]

Factor I

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
9	(.86930)	9	(.84517)
18	(.83911)	18	(.82333)
14	(.82799)	14	(.80383)
16	(.77404)	16	(.76097)
26	(.53971)	26	(.47375)

Factor II

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
6	(.82072)	22	(.77825)
22	(.75780)	6	(.69581)
13	(-.57330)	28	(.53028)
24	(.55490)	24	(.48000)
28	(.51589)	25	(-.42278)

Factor III

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
7	(.77423)	7	(.74430)
23	(.77226)	23	(.73915)
27	(.62846)	10	(.61223)
10	(.62814)	27	(.42985)

Factor IV

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
8	(.70444)	5	(.78667)
5	(.66238)	8	(.48972)
20	(.54554)	20	(.48163)

Factor V

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
25	(.70828)	11	(.59110)
19	(.63434)	17	(.56193)
17	(.60123)	19	(.43543)

(cont.)

Factor VI

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
3	(.80552)	3	(.58259)
12	(.41247)	12	(.47724)

Factor VII

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
11	(.78386)	21	(.71080)
1	(.54778)		

Factor VIII

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
21	(.83544)	15	(.83399)

Factor IX

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
15	(.89427)	2	(.59057)
		4	(.46328)

Factor X

PCA extraction Item	(Struc. Coeff.)	PAF extraction Item	(Struc. Coeff.)
4	(.74118)	13	(.73277)
2	(.54477)		
29	(.48707)		

[1]Items are listed under each factor in descending order of the absolute values of their factor structure coefficients.

Table 5
Principal Components Three-Factor Rotated Solution Matrix

	FACTOR1	FACTOR2	FACTOR3
MYSELF1	.14288	-.52065*	.17304
MYSELF2	.41299*	.17433	.01133
MYSELF3	.31451*	.14906	-.00643
MYSELF4	.30699*	.08766	-.03429
MYSELF5	.02536	.66790*	.13844
MYSELF6	.10650	.09884	.69682*
MYSELF7	.12625	.61145*	-.09756
MYSELF8	-.12928	-.30911*	.02487
MYSELF9	.83499*	-.03698	-.10900
MYSELF10	.18936	-.75860*	-.09147
MYSELF11	-.16907	.33500*	-.11529
MYSELF12	.16244	-.02702	-.32851*
MYSELF13	-.07218	.22922	-.44816*
MYSELF14	.83174*	-.13026	-.00943
MYSELF15	.24205*	-.10032	-.00679
MYSELF16	.64798*	-.04798	-.09800
MYSELF17	-.24436	-.50795*	.26963
MYSELF18	.79585*	-.02196	-.31655
MYSELF19	-.09399	-.25246	.32288*
MYSELF20	-.22954	.43853	.43989*
MYSELF21	-.08753	.18514	.36532*
MYSELF22	-.03275	-.01547	.80000*
MYSELF23	.19866	.65412*	.22517
MYSELF24	-.28917	-.12307	.59293*
MYSELF25	.07943	-.43193	.47359*
MYSELF26	.59723*	-.22852	-.33556
MYSELF27	-.03055	.51746*	-.03223
MYSELF28	.28482	.17814	-.58234*
MYSELF29	-.02016	.37165*	6860

Note: Starred values indicate the factor with which the item is most highly correlated.

Table 6
Principal Axis Three-Factor Solution Matrix

	FACTOR1	FACTOR2	FACTOR3
MYSELF1	.09660	-.46131*	.13520
MYSELF2	.31552*	.13943	-.01636
MYSELF3	.24243*	.11808	-.02390
MYSELF4	.23112*	.06110	-.04882
MYSELF5	.00871	.60715*	.11227
MYSELF6	.05638	.08050	.62719*
MYSELF7	.10974	.55560*	-.09600
MYSELF8	-.11300	-.26093*	.02463
MYSELF9	.83380*	-.03624	-.08566
MYSELF10	.17546	-.74746*	-.08782
MYSELF11	-.12437	.27625*	-.07385
MYSELF12	.14578	-.02774	-.26842*
MYSELF13	-.02534	.19659	-.35905*
MYSELF14	.81483*	-.13213	.00375
MYSELF15	.18976*	-.08857	-.01629
MYSELF16	.57658*	-.04487	-.11001
MYSELF17	-.22177	-.45467*	.24037
MYSELF18	.80470*	-.01779	-.30109
MYSELF19	-.09070	-.21178	.25972*
MYSELF20	-.21556	.39585	.40198*
MYSELF21	-.10509	.15187	.28572*
MYSELF22	-.04702	-.02125	.78626*
MYSELF23	.17001	.61020*	.20555
MYSELF24	-.28994	-.11789	.53214*
MYSELF25	.04221	-.39546	.41046*
MYSELF26	.54865*	-.21093	-.31670
MYSELF27	-.02054	.45023*	-.02123
MYSELF28	.27635	.16718	-.52987*
MYSELF29	-.04701	.31418*	.29613

Note: Starred values indicate the factor with which the item is most highly correlated.