DOCUMENT RESUME

ED 314 447 TM 014 258

AUTHOR	Daniel, Larry G.
TITLE	Comparisons of Exploratory and Confirmatory Factor Analysis.
PUB DATE	Nov 89
NOTE	37p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Little Rock, AR, November 8-10, 1989).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Factor Analysis; Goodness of Fit; Junior High Schools; *Junior High School Students; Psychological Testing; *Research Methodology; *Statistical Analysis
IDENTIFIERS	*Confirmatory Factor Analysis; *Exploratory Factor Analysis; LISREL Computer Program

ABSTRACT

Historically, most researchers conducting factor analysis have used exploratory methods. However, more recently, confirmatory factor analytic methods have been developed that can directly test theory either during factor rotation using "best fit" rotation methods or during factor extraction, as with the LISREL computer programs developed by K. G. Joreskog et al. (1986). This study provides a heuristic example that illustrates some comparisons between statistical results obtained using exploratory and confirmatory factor analytic techniques with a data set of scores of 145 junior high school students on 26 psychological tests. The first analysis used exploratory factor extraction of orthogonal principal components followed by confirmatory rotation to determine goodness of fit. The second analysis used confirmatory extraction of five factors based on theoretical expectations. Results indicate that exploratory and confirmatory factor analytic procedures can yield different results. While exploratory methods are an invaluable aid in the initial structuring of behavioral constructs, confirmatory methods serve the researcher in the affirmation or rejection of models derived using theoretical considerations or exploratory factor analytic techniques. Seven tables present the data set. An appendix lists the tests taken by the students, and a second appendix gives the LISREL commands for running the second analysis. A 34-item list of references is included. (SLD)

factan89.lgd

Comparisons of Exploratory and Confirmatory Factor Analysis

Larry G. Daniel

University of Southern Mississippi

Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR, November 8-10, 1989.

Abstract

Historically, most researchers conducting factor analysis have employed what have come to be called exploratory methods. However, more recently, confirmatory factor analytic methods have been developed that can directly test theory either during factor rotation using "best fit" rotation methods, or else during factor extraction, as with the LISREL computer programs developed by Joreskog and his colleagues. The present study utilizes a small data set to compare results across various exploratory and confiimatory factor analytic procedures. The benefits and limitations of each approach are summarized.

Comparisons of Exploratory and Confirmatory Factor Analysis

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). The usefulness of factor analytic methods has been established through traditional applications of these methods both in theory development (e.g., Guilford's Structure-of-Intellect Model--Guilford, 1965, 1967; Guilford & Hoepfner, 1971) and in test validation. Most researchers employing factor analytic techniques have used what have come to be known as "exploratory" techniques. However, more recently, "confirmatory" factor analytic methods have been developed that can directly test theory either during factor rotation using "best fit" Procrustean rotation methods, or else during factor extraction through the use of special statistical software packages such as LISREL (Joreskog & Sorbom, 1986).

The purpose of the present study was to provide a heuristic example which illustrates some comparisons between statistical results obtained using exploratory and confirmatory factor analytic techniques. A data set consisting of the scores of junior high school students on a variety of psychological tests (Holzinger & Swineford, 1939) is utilized for this purpose. This data set has traditionally been used with some frequency to illustrate factor analytic applications (cf. Gorsuch, 1983; Joreskog, 1986). Following a brief overview of the processes and

the relative strengths and weaknesses of exploratory and confirmatory factor analytic methods, the results of various factor analytic procedures using the selected data set are presented. Based on the results of these analyses, conclusions are drawn relative to the usefulness of exploratory and confirmatory methods.

Exploratory Factor Analytic Methods

The conventional "exploratory" factor analytic methods are designed to analyze covariance structures among a common set of variables, and to explain the relationship among these variables in terms of a smaller number of unobserved latent variables called factors. These factors, if interpretable, constitute "the initial structuring of a field [of study]" (Cattell, 1952, p. 359). Thus, Kerlinger (1986) refers to factor analysis as a means of reducing complex data sets containing numerous variables to a more manageable size. This data reductive property of exploratory factor analytic methods serves to make these methods useful in establishing the construct validity or "constitutive meaning" (Kerlinger, 1986) of psychometric instruments. Hence, Nunnally (1978, p. 111) notes that construct validity has been referred to as "factorial validity" and "trait validity."

Similarly, Gorsuch (1983, pp. 350-351) suggests that

> 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. . . .If a theory has

> clearly defined constructs, then scales can be directly built to embody those constructs. However, it is often the case that the theories in a particular area are sufficiently undeveloped so that the constructs are not clearly identified.

In short, "factor analysis is intimately involved with questions of validity. . . .Factor analysis is at the heart of the measurement of psychological constructs" (Nunnally, 1978, pp. 112-113).

Despite their usefulness in addressing issues related to construct validity, exploratory methods are not without their shortcomings. For instance, exploratory methods do not directly address issues of theory development. Exploratory factor structures are determined on the basis of the mechanics of the method rather than on the basis of a priori considerations of the researcher (Cattell, 1978). Selection of the number of factors to extract, for instance, is based upon one of several mechanical procedures, such as Guttman's "eigenvalue greater than one" rule (Guttman, 1954) or Cattell's "scree" test (Cattell, 1966). Zwick and Velicer (1986) present evidence that there may be discrepancies as to the number of factors that should be extracted across these various procedures, thus further confounding the problem of the number of factors to extract. Similarly, Cattell (1978) discusses problems associated with extraction of either too many or too few factors.

Factor interpretation may also be a problem in exploratory

methods:

> The major disadvantage of pure exploratory factor analysis lies in the difficulty involved in interpreting the factors. This difficulty most often comes about because the researcher lacks even tentative prior knowledge about the processes which produce covariation among the variables studied and has no basis on which to make his interpretations. In these circumstances the interpretations given the factors may be nothing more than tautological transformations of the names of the original variables. Difficulty is also encountered when the factors obtained represent confounded effects and the researcher is unable to decide which of these effects is unique to the factor--a problem which may come about from random selections of variables. (Mulaik, 1972, p. 363)

Additional interpretation problems may result due to overinterpretation of factors with small factor structure coefficients (Nunnally, 1978), or if factor structures are considered confirmed based upon post hoc analysis of factor structures for a given data set (Gorsuch, 1983).

A third criticism of exploratory methods is that they impose upon the research situation certain assumptions which may not always honor the relationships among the variables in a given data set. These assumptions include requiring that all common

factors be correlated with one another in the oblique case, or that all common factors be uncorrelated with one another in the orthogonal case; requiring that all observed variables be affected by all common factors; and requiring that all common factors be uncorrelated with all unique factors (Long, 1983a, p. 12).

Despite these shortcomings, when used appropriately, exploratory factor analysis is useful in helping researchers to assess the nature of relationships among variables within a given set, and, consequently, to establish the construct validity of tests. Hence, exploratory methods are 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 Analytic Methods

Confirmatory factor analytic methods extend the usefulness of the exploratory methods by determining the extent to which "endogenous" latent variables, i.e., those variables actually occurring within the factor analytic model, can be explained by "exogenous" variables, i.e., a set of conceptual or theoretical variables determined outside the model (Long, 1983b). Hence, confirmatory methods are superior to exploratory methods in that they may be utilized to test hypotheses regarding the nature of observed factors (Gorsuch, 1983). In order to perform these analyses, *a priori* factors are specified based upon theoretical expectations. Confirmatory methods then seek to optimally match the observed and theoretical factor structures for a given data

set to determine the goodness of fit of the predetermined factor model.

Considering that confirmatory methods capitalize more on theoretical considerations of a given data set than do exploratory methods, it is not surprising Gorsuch (1983) concluded that confirmatory methods should be the more widely used of the two major factor analytic approaches, with exploratory methods "reserved only for those areas that are truly exploratory, that is, areas where no prior analyses have been conducted" (p. 134). As opposed to the "canned" restraints inherent in exploratory methods, confirmatory methods allow the researcher to specify "substantively motivated" constraints which define the factor structural model (Long, 1983a). These constraints include, but are not limited to, predetermination of the degree of correlation, if any, between each pair of common factors, predetermination of the degree of correlation between individual variables and one or more factors, and specification as to which particular pairs of unique factors are correlated.

Mathematically, confirmatory factor analytic models are covariance structure models (Bock & Bargmann, 1966; Joreskog, 1973; Long, 1983b), or as they are more popularly known, "structural equation" or "LISREL" models. Bentler (1980), Baldwin (1989), and Long (1983b) offer coherent introductory discussions of these models. As Long (1983b, p. 12) points out, these models decompose the covariances among the variables in a given set in two conceptually distinct steps:

> First, the observed variables are linked to unobserved or latent variables through a factor analytic model, similar to that commonly found in psychometrics. Second, the causal relationships among these latent variables are specified through a structural equation model, similar to that found in econometrics.

According to Anderson (1987, p. 49), constructing a structural equation model "begins with a statement of a verbal theory that makes explicit the relations hypothesized among a set of variables," and then makes assumptions regarding a causal structure among the latent and observed variables in the set. In the confirmatory factor analytic case, the structural equation model can be used to compute maximum likelihood estimates of the degree of fit between observed and theoretical factor structures. As Bentler (1980, p. 420) described this process:

> The primary statistical problem [in structural equation modeling] is one of optimally estimating the parameters of the model and determining the goodness-of-fit of the model to sample data on measured variables. If the model does not acceptably fit the data, the proposed model is rejected as a possible candidate for the causal structure underlying the observed variables. If the model cannot be rejected statistically, it is a plausible [but not necessarily the most

appropriate] representation of the causal structure.

Confirmatory analysis can be utilized in two ways. First, exploratory methods can be used to determine the number of factors to extract, and then the factor results can be rotated to a "best fit" confirmatory position. This use of confirmatory analysis is especially useful when the researcher does not have a single theory in mind which can be theorized to explain the data, or when the theorized model fails to be confirmed (Long, 1983a). In this sense, the confirmatory model is actually used in an exploratory fashion. To purists this method might be considered questionable as it violates the assumption of independent sampling; however this problem can be overcome by splitting a given sample and using one cohort to determine the nature of the factor structure, and the other to determine the degree to which the model explains the data (Gorsuch, 1983).

A second way of using confirmatory methods is to employ them during factor extraction, with the number of factors being determined on the basis of previous exploratory studies, or upon theoretical considerations. Use of the model in this manner constitutes confirmatory factor analysis in the "pure" sense. The purpose of confirmatory factor analysis in this application is to determine the overall fit of the theoretical model to the data collected from the sample at hand.

In conducting a confirmatory factor analytic procedure of either type using LISREL VI (Joreskog & Sorbom, 1986), the user

decides which parameters of the model are to be "fixed" and which are to be "free." For instance, values on the factor structure matrix can be set at prespecified levels based upon expected correlation between individual variables and one or more factors. Fixed values are indicated by being given a zero value; free values are estimated during the analysis by fitting the model to the data in light of the constraints of theoretical expectation. The user may also specify whether the model is to be applied to the correlation or the covariance matrix. Although results are generally similar with either matrix, in certain research situations, the covariance matrix may better honor the reality of the data due to certain structural and distributional properties inherent to assessment and behavioral data (Cudeck, 1989). Finally, the degree of correlation, if any, between the factors can also be specified.

If the analysis succeeds in identifying all of the specified parameters, the model is identified. One check of the identifiability of the model is the test of positive definiteness of the information matrix. According to Joreskog and Sorbom (1986, I.23-I.24), "The information matrix is the probability limit of the matrix of second order derivatives of the fitting function used to estimate the model. . ." If the model is not identified, LISREL prints a warning message that the information matrix is not positive definite. Joreskog and Sorbom (1986, p. I.24) warn, however, that obtaining a positive definite information matrix does not necessarily insure that the model is

identified, although experience in using the LISREL software has indicated that this test is a highly reliable indicator of the model's goodness of fit.

Once a model is identified, the data set used for identifying the model would not be used by the purist to assess the fit of the model. This does not mean, however, that the estimators for the first data set should not be consulted. These estimators should be regarded as preliminary findings (Leamer, 1978), and "[t]he model selected must be viewed as tentative, in need of verification with a second, independent sample" (Long, 1983a, p. 68). LISREL VI offers three fitting functions for estimating parameters in confirmatory models. These functions are based on three distinct types of estimators, namely unweighted least squares, generalized least squares, and maximum likelihood estimators. Long (1983a) and Joreskog and Sorbom (1986) discuss the relative properties of each of these three estimators.

For the purposes of the present study, only maximum likelihood estimates will be utilized as these estimates are scale invariant, approximately unbiased, and approximately multivariate normally distributed. As with may other hypothesis testing statistical methods, confirmatory maximum likelihood analyses are useful in that they yield a complete analysis. Rotational procedures are not needed since maximum likelihood estimates are directly calculated in their final form (Gorsuch, 1983, p. 119). Hence, in addition to the advantages of

confirmatory methods over exploratory methods mentioned earlier, these methods are superior for yet another reason, namely that they rid the researcher of the problems inherent to the selection of an appropriate rotational procedure.

In addition to offering direct maximum likelihood estimators of each of the parameters of the model, LISREL VI also offers four statistics which indicate the degree of overall fit of the observed factor structure to the theoretical structure. The chi-square goodness of fit statistic tests the null hypothesis that there is no statistically significant difference in the observed and theoretical covariance structure matrices. The value of chi-square ranges from zero to infinity, with a zero value indicating perfect fit. Obviously, in the majority of cases, the researcher's goal is to not reject the null hypothesis, thereby confirming the expected structure. Since a statistically significant result yields a rejection of the fit of a given model, the chi-square statistic has been referred to as a "lack of fit index" (Mulaik, James, Van Alstine, Bennet. Lind, & Stilwell, 1989).

Bentler and Bonett (1980) point out that the chi-square test is subject to yielding statistically significant results (and consequently, failing to offer evidence of goodness of fit) when sample sizes are large. Thus, it may be unclear in many large sample research situations whether the statistical significance of the chi-square is due to poor fit of the model or to the size of the sample. Joreskog and Sorbom (1986, p. I.39) emphasize

that chi-sqare values may also escalate to statistically significant values when variables are not subject to multivariate normal distribution. Hence, it is true here as it is in many other research situations that reliance upon statistical significance testing may not always be the best way to evaluate statistical results (Carver, 1978; Thompson, 1987).

A second statistic for assessing overall fit of a model is the goodness of fit (GFI) index. According to Joreskog and Sorbom (1986, p. I.41), "GFI is a measure of the relative amount of variances and covariances jointly accounted for by the model." In addition to GFI, LISREL also computes an adjusted goodness of fit (AGFI) statistic based on a correction for the number of degrees of freedom in a less restricted model obtained by freeing more parameters. Unlike chi-square, GFI and AGFI are less dependent on sample size or departures from multivariate normality. Generally, these goodness of fit indicators range in value from zero to one, although negative values are also possible.

A third overall model fit statistic is the root mean square residual (RMR). RMR is "a measure of the average of the residual variances and covariances" (Joreskog & Sorbom, 1986, p. I.41), and since it is an indicator of measurement error, it is most useful when comparing two different models for the same data set. Consequently, researchers rely less frequently upon RMR than upon other goodness of fit indices.

A final statistic for assessing overall model fit is the

total coefficient of determination, a derivative of the multiple correlation coefficient (R^2) which originated with Specht (1975). Like GFI and AGFI, the coefficient of determination ranges between zero and unity. According to Anderson (1987, p. 53), "This statistic indicates the amount of variation in the endogenous variables jointly accounted for by the model." The coefficient of determination can also serve as a generalized measure of the reliability of the model (Joreskog & Sorbom, 1986, p. III.11).

Even when all of the overall model fit statistics indicate a high degree of fit between the observed and expected factor structures, it is important to remember that

> . . .the measures χ^2, GFI and RMR are measures of the overall fit of the model to the data and do not express the quality of the model judged by any other internal or external criteria. For example, it can happen that the overall fit of the model is very good but with one or more relationships in the model very poorly determined, as judged by the squared multiple correlations. . . (Joreskog & Sorbom, 1986, p. I.41)

It is also possible that more than one model can be determined to adequately fit the data (Biddle & Marlin, 1987; Thompson & Borrello, 1989). Thompson, Webber, and Berenson (1987) illustrate this point in the confirmatory analysis of the factor structure of a children's health locus of control measure.

Thus, obtaining good fit of a model to a particular data set does not mean that the model is confirmed, since other models (tested or not) may also fit the data.

When a model has been rejected (i.e., when the confirmatory analysis fails to adequately fit the observed factor structure with the theoretical structure), the researcher may wish to implement a "specification search." This procedure uses the estimates computed for the rejected model to suggest better fitting models. The LISREL VI package can be utilized to conduct a specification search once a model has been rejected. The procedure relaxes fixed parameters of the model one at a time on the basis of "modification indices." These indices reflect the changes of the chi-square goodness of fit test with each of the fixed parameters, in turn, relaxed, and then frees the one parameter which offers the greatest amount of improvement in the fit of the model.

A Heuristic Example Comparing Exploratory and Confirmatory Methods

Holzinger and Swineford (1939) conducted various factor analytic procedures on a data set consisting of scores of 317 junior high school students on 26 psychological tests measuring various ability variables. The "Grant-White" cohort of this data set (n = 145) will be utilized in the present study as a heuristic example illustrating differences in factor analytic results using exploratory and confirmatory methods. The 26 tests administered to the sample were theoretically designed to measure

five broad psychological constructs--visualization (Tests 1 through 4), verbal intelligence (Tests 5 through 9), test-taking speed (Tests 10 through 13), memory (Tests 14 through 19), and mathematical ability (Tests 20 through 26). A brief description of the content of each of the tests is presented in Appendix A. Discussions of Holzinger and Swineford's (1939) factor analytic results based on these data are presented by Gorsuch (1983) and Mulaik (1972).

As previously noted, the data set utilized in the present study included scores for 145 subjects on the 26 measures. Two separate factor analyses were performed using these data. The first analysis utilized exploratory factor extraction of orthogonal principal components followed by confirmatory rotation to determine goodness of fit. The second analysis utilized confirmatory extraction of five factors based upon theoretical expectations. In both analyses, factors were allowed to correlate during the confirmatory procedures.

Results

Analysis I: Exploratory Extraction, Then Confirmatory Rotation

An initial principal components exploratory factor analysis was run using the SPSSx FACTOR procedure. The analysis yielded six factors with prerotation eigenvalues greater than one. Analysis of the "scree" plot (Cattell, 1966) of the eigenvalues indicated the appropriateness of a four factor solution. These four factors accounted for 54.6% of the variance. Factor results were rotated to the varimax criterion. The resultant factor

matrix presented in Table 1.

INSERT TABLE 1 ABOUT HERE

Using a saliency criterion of |.40|, Factor I was most highly saturated with Tests 9, 7, 6, 5, 8, 20, and 22. The five tests correlated most highly with this factor (Tests 9, 7, 6, 5, and 8) were originally designed to measure abilities in the verbal domain. The remaining two tests (20 and 22) that correlated highly with this factor were designed as measures of mathematical ability. Hence, the observed structure for this factor does not exactly match the expected structure, although it does represent relatively well the expected verbal ability factor.

Factor II was most highly saturated with Tests 1, 4, 25, 26, 2, 3, and 23. Four of these tests (Tests 1 through 4) were designed as measures of visual ability. The remaining three tests were originally designed as measures of mathematical ability; however, since Tests 25 and 26 are alternate measures, respectively, for Tests 3 and 4, they can be regarded as tests of visual ability. Hence, this factor represents well the expected structure of the visual abilities factor. It is interesting that Test 23, the seemingly "misplaced" test, also correlated highly with Factor I (factor structure coefficient equal to .43), indicating that this test did not uniquely saturate the factor space occupied by either of the first two factors. Were this a typical exploratory construct validity study, a researcher might

be prone to delete this test from future validity studies.

Factor III was most highly saturated with Tests 10, 12, 13, 11, 24, and 21. Four of these tests (Tests 10 through 13) were designed as measures of test taking speed. The remaining two measures (Tests 24 and 21) were designed as measures of mathematical ability. Item 24 also correlated highly with Factor I (factor structure coefficient equal to .41), offering evidence for possible deletion of the measure from future analyses.

Finally, Factor IV was most highly saturated with Tests 17, 14, 15, 16, 18, and 19. This was perhaps the "cleanest" of the four factors, with all of the six highly correlated tests originally designed as measures of memory and recognition. One of the measures (Test 16), however, also correlated highly with Factor II (factor structure coefficient equal to .45).

In sum, this exploratory factor analytic procedure yielded interpretable results, with the four extracted factors accounting for all 26 of the test variables. Four of the five expected factors were identified, with each of the five tests of the remaining expected factor (mathematical ability) correlating highly across one or more of the first three factors. In interpreting these results, the researcher might wish to delete the last five measures from future analyses, or else re-evaluate the construct thought to be measured by the tests in this cohort. The fact that several of the tests in this cohort correlated well with more than one factor could serve as a justification for deleting those measures from subsequent analyses.

As a follow-up analysis to this exploratory procedure, a Procrustean rotation procedure was utilized to assess the viability of this four factor solution. Procrustean rotation procedures are used to assess the goodness-of-fit of observed factor structures to expected structures. One method for performing these procedures has been discussed by Gorsuch (1983) and psychometrically elaborated by Thompson (1986). This method 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 correlation coefficients, and hence provide estimates of the degree of goodness (or badness) of fit between the two factor solutions. These cosines may be computed using the computer program RELATE (Veldman, 1967) which also "recreates" the observed factor matrix to best fit the theoretical factor matrix.

In the present case, the Table 1 matrix served as the actual structure matrix. The "target" structure matrix was structured based upon the original theoretical factors suggested by Holzinger and Swineford (1939). Each of the 26 measures was expected to be univocal (i.e., to "speak" through only one factor); consequently, each variable was assigned a factor-structure coefficient of unity on the factor with which it was expected to correlate and a value of zero on all other factors. Since the mathematical abilities factor was not identified in the previous exploratory analysis, Tests 20 through 26 (originally

designed as tests of mathematical ability) were expected to correlate with Factor II (Visual Abilities). This placement was selected since three of these seven variables had correlated highly with Factor II in the previous exploratory analysis.

Table 2 presents the cosines among the factor axes across the expected and actual factor matrices. The reproduced observed factor matrix resulting from this analysis is presented in Table 3. These results suggest a high degree of fit of the actual structure to the expected structure, with all of the cosines among paired factor axes approaching unity (values ranging from .9752 to .9975), and with all but one of the measures (Test 24) highly correlating with the expected factors. In addition, each observed factor tended to correlate rather weakly with expected factors other than the one it was supposed to represent. Thus, these results tend to confirm the previous findings.

INSERT TABLES 2 AND 3 ABOUT HERE

Analysis II: Confirmatory Extraction

The second method used for analyzing the data consisted of confirmatory extraction of five factors based upon the theoretical model posited for the 26 measures. This procedure was a "pure" application of confirmatory factor analysis as no consideration was given to exploratory procedures prior to conducting the analysis. A target factor matrix was specified with free values for 26 parameter estimates based upon expected factor structures. All of the measures were predicted to be univocal, i.e., correlated with or "speaking through" only one

factor, although factors were allowed to correlate. The LISREL VI commands for running this analysis are listed in Appendix B.

The maximum likelihood estimates resulting from this analysis are presented in Table 4, and the interfactor phi or correlation matrix is presented in Table 5. The chi-square statistic for the model (<u>df</u> = 289) was 466.57 (probability level < 0.001). The total coefficient of determination and goodness of fit index, respectively, were 0.95 and 0.80. Interfactor correlations were relatively high, with off-diagonal coefficients ranging from 0.46 to 0.87, and with the highest interfactor correlations between the fifth and the remaining four factors. These results indicate that the expected five-factor model is identified, although the fit of the model is less than adequate.

<u>INSERT TABLES 4 AND 5 ABOUT HERE</u>

The LISREL model specified for Analysis II included a specification search based upon automatic model modification. Allowing for these modifications essentially made the confirmatory model an exploratory one since these modifications were <u>not</u> based upon theoretical considerations. Automatic modification indices were utilized to free the fixed model parameters (13, 2), (24, 3), (21, 3), (17, 2), (10, 2), and (24, 2) in an attempt to improve the fit of the model. Modifications through parameter (21, 3) were considered noteworthy. The three remaining modifications were not considered as the modification of parameter (17, 2) yielded an unrealistic value for parameter

(17, 4) (maximum likelihood estimate greater than one). Since this value could not be interpreted, the remaining modifications were not considered (Joreskog & Sorbom, 1986, p. I.42).

The maximum likelihood estimates and interfactor phi matrix which resulted from the first three modifications of the model are presented in Tables 6 and 7, respectively. The chi-square statistic for the model (df = 286) was 413.22 (probability level < 0.001). The total coefficient of determination and goodness of fit index, respectively, were 0.98 and 0.83. Thus, the overall fit of the model improved with these modifications, although the improvement was minimal. In addition, the interfactor phi-matrix correlations were adjusted downward, with off-diagonal coefficients ranging from 0.37 to 0.85.

INSERT TABLES 6 AND 7 ABOUT HERE

The goodness of fit is still not adequate enough to assume the appropriateness of the model. Interestingly, two of the three measures for which Analysis II model parameters were modified (Tests 21 and 24) were in the set of items expected to measure mathematical ability. This finding, along with the high interfactor correlations between the MATH factor and each of the other factors, suggests that the mathematical abilities construct may not be as discretely identified as the other four factors.

Discussion

The results of the analyses presented above indicate that exploratory and confirmatory factor analytic procedures can yield

different results. In the present example, the atheoretical, mechanical factor extraction procedures utilized in the exploratory factor analytic procedure (Analysis I) yielded a different number of factors than the theoretically driven procedures utilized in the confirmatory procedure (Analysis II). It is interesting to note, however, that the seeming accuracy of the Procrustean-rotated initial exploratory factor matrix in matching the expected target matrix may not necessarily indicate the viability of the four factor structure over the five factor structure. Gorsuch (1974) points out that Procrustean methods not only tend to capitalize on chance, but also rely heavily on "stretching and trimming the data to fit the hypotheses" (p. 167). Horn (1967) provides psychometric evidence to illustrate this property of Procrustean methods, showing that these methods can sometimes lead to erroneous confirmatory results even when factor structures are determined using totally random numbers!

The identification of a five-factor model using the LISREL confirmatory procedures could be regarded as an affirmation of Holzinger and Swineford's (1939) classifications of the various ability measures. However, the less than desirable fit of this model to the data indicates the need for additional exploratory studies with other data sets to either confirm or challenge the five-factor theoretical structure posited for these measures.

The results of the exploratory procedures utilized in Analysis I illustrate a valuable point, namely that initial exploratory factor structure should not be regarded as

confirmation of expected factor structure simply because some or all of the observed factors match expected factors. Additional confirmatory factor analytic studies that establish goodness of fit across several different data sets are needed to genuinely confirm initial exploratory factor analytic results. As Gorsuch (1983, p. 335) put it, "To the extent that invariance can be found across systematic changes in either variables or the individuals, then the factors have a wider range of applicability as generalized constructs." In the Analysis I example, the follow-up confirmatory procedure challenged the exploratory factor solution even though the original exploratory solution yielded interpretable factors and appeared to account well for the variance in the factor space.

Another point illustrated by the analyses conducted here is that confirmatory methods may be used in an exploratory fashion. Confirmatory methods can be used as (a) follow-up procedures to exploratory factor identification (as illustrated in Analysis I), or (b) effective exploratory methods for examining modification of confirmatory models which are either not identified or do not adequately fit a given data set (as illustrated in Analysis II). This modification funtion of confirmatory models can provide the researcher with clues as to how best to restructure theory. In cases in which models are specified so as to allow factors to correlate, modification can also be used to determine how changes in the model affect these correlations, with the possibility of identifying nested effects of some factors within others.

Finally, even though neither the four- nor five-factor solution models proposed herein achieved very good fit with the selected data set, both models were viable. This finding illustrates another strength of confirmatory factor analysis, i.e., the identification of multiple plausible models for explaining the relationships among a given set of variables. This aspect of confirmatory models makes the models especially useful in research situations in which the theoretical structures underlying observed behaviors are weak or nonexistent, or in which there are competing theoretical models for explaining relationships among behavioral variables.

Summary

The present study has illustrated several ways in which confirmatory factor analytic methods may be useful in behavioral research. Advantages of confirmatory techniques over traditional exploratory factor analytic techniques include the ability to directly test theory and the ability to generate various statistics for determining goodness of fit of theoretical models to actual data. Used either 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. Whereas exploratory methods are an invaluable aid in the initial structuring of behavioral constructs, confirmatory methods serve the researcher in the affirmation or rejection of models derived using theoretical considerations or exploratory factor analytic techniques.

References

Anderson J. G. (1987). Structural equation models in the social and behavioral sciences: Model building. Child Development, 58, 49-64.

Baldwin, B. (1989). A primer in the use and interpretation of structural equation models. Measurement and Evaluation in Counseling and Development, 22, 100-112.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. Annual Review of Psychology, 31, 11-21.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Biddle, B. J., & Marlin, M. M. (1987). Causality, confirmation, credulity, and structural equation modeling. Child Development, 58, 4-17.

Bock, R. D., & Bargmann, R. E. (1966). Analysis of covariance structures. Psychometrika, 31, 507-534.

Carver, R. P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

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(2), 245-276.

Cattell, R. B. (1978). The scientific use of factor analysis in

behavioral and life sciences. New York: Plenum.

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

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

Guilford, J. P. (1965). Intelligence: 1965 model. American Psychologist, 21, 20-26.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw-Hill.

Guilford, J. P., & Hoepfner, R. (1971). The analysis of intelligence. New York: McGraw-Hill.

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

Holzinger, K. J., & Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution. Chicago: University of Chicago.

Joreskog, J. A. (1969). A general approach to confirmatory factor analysis. Psychometrika, 34, 183-202.

Joreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger, and O. D. Duncan (Eds.), Structural equation models in the social sciences (pp. 85-112). New York: Seminar.

Joreskog, 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.

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

Leamer, E. E. (1978). Specification searches. New York: John Wiley.

Long, J. S. (1983a). Confirmatory factor analysis: A preface to LISREL. Beverly Hills: SAGE.

Long, J. S. (1983b). Covariance structure models: An introduction to LISREL. Beverly Hills: SAGE.

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

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N. ind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. Psychological Bulletin, 105, 430-445.

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

Specht, D. A. (1975). On the evaluation of causal models. Social Science Research, 4, 113-133.

Thompson, B. (1986, April). A partial test distribution for cosines among factors across samples. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Thompson, B. (1987, April). The use (and misuse) of statistical significance testing. Paper presented at the annual meeting of

the American Educational Research Association, Washington. (ERIC Document Reproduction Service No. ED 287 868)

Thompson, B., & Borrello, G. M. (1989, January). A confirmatory factor analysis of data from the Myers-Briggs Type Indicator. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston.

Thompson, B., Webber, L., & Berenson, G. (1987). Factor structure of a children's health locus of control measure: A confirmatory maximum-likelihood analysis. Educational and Psychological Measurement, 47, 1071-1080.

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
Varimax Rotated and Sorted Exploratory Factor Matrix

Variable	I	II	III	IV
T9	.84368	.16866	.05732	.18178
T7	.84067	.16193	.15760	.07190
T6	.80078	.18593	.06986	.19972
T5	.77695	.20061	.20741	.08979
T8	.63806	.31948	.24004	.10541
T20	.44767	.38582	.07164	.32942
T22	.43534	.36265	.11846	.33263
T1	.16061	.69723	.19667	.15574
T4	.22608	.68976	.10959	.07172
T25	.14618	.63062	-.06440	.13987
T26	.24343	.62363	.07000	.04550
T2	.07949	.58449	.07805	.03359
T3	.13273	.54714	.13808	.12329
T23	.42674	.48382	.22395	.23677
T10	.17895	-.11212	.83838	.12053
T12	.02371	.20235	.78960	.04919
T13	.17549	.45558	.61988	.02471
T11	.18564	.07723	.61887	.34552
T24	.41126	.10006	.54632	.28598
T21	.17557	.43149	.48810	.19031
T17	.16494	-.00857	.24505	.69260
T14	.20588	.02369	.04536	.68092
T15	.08580	.10845	.04312	.65548
T16	.05532	.44604	.05855	.57261
T18	-.00345	.34462	.38085	.47471
T19	.15941	.18196	.19371	.45310

Table 2
Cosines Among Axes of Observed and Theoretical Factor Structures

	1	2	3	4
1	.9209	.3488	.1322	.1129
2	-.3417	.9346	-.0006	-.0988
3	-.1140	-.0494	.9894	-.0750
4	-.1485	.0500	.0601	.9858

Table 3
The Table 1 Matrix Rotated to "Best Fit"
with the Expected Target Matrix*

	VERBAL	VISUAL	SPEED	MEMORY
T1	.4347	.5813	.1302	.1764
T2	.2912	.5157	.0368	.0552
T3	.3452	.4537	.0852	.1375
T4	.4714	.5602	.0432	.0782
T5	.8203	-.0870	.1000	-.0044
T6	.8341	-.1197	-.0463	.0915
T7	.8596	-.1431	.0467	-.0364
T8	.7427	.0700	.1411	.0396
T9	.8693	-.1487	-.0614	.0658
T10	.2501	-.1783	.8056	.1370
T11	.3187	-.0257	.5614	.3541
T12	.2023	.1757	.7649	.1025
T13	.4052	.3630	.5690	.0583
T14	.2808	-.1155	-.0308	.6446
T15	.1966	.0073	-.0216	.6415
T16	.2789	.3414	-.0133	.5821
T17	.2595	-.1329	.1722	.6726
T18	.2210	.2762	.3246	.5086
T19	.2870	.0707	.1305	.4437
T20	.5935	.1750	-.0239	.2819
T21	.3982	.3242	.4274	.2124
T22	.5806	.1572	.0247	.2885
T23	.6181	.2828	.1313	.2076
T24	.5181	-.0756	.4673	.2587
T25	.3619	.5256	-.1220	.1439
T26	.4561	.4951	.0073	.0441

*The largest factor structure coefficient for each variable across the four factors is printed in bold type.

Table 4
Maximum Likelihood Factor Matrix for Analysis II

	VERBAL	VISUAL	SPEED	MEMORY	MATH
T1	0.000	0.704	0.000	0.000	0.000
T2	0.000	0.498	0.000	0.000	0.000
T3	0.000	0.502	0.000	0.000	0.000
T4	0.000	0.681	0.000	0.000	0.000
T5	0.803	0.000	0.000	0.000	0.000
T6	0.816	0.000	0.000	0.000	0.000
T7	0.838	0.000	0.000	0.000	0.000
T8	0.702	0.000	0.000	0.000	0.000
T9	0.844	0.000	0.000	0.000	0.000
T10	0.000	0.000	0.659	0.000	0.000
T11	0.000	0.000	0.683	0.000	0.000
T12	0.000	0.000	0.714	0.000	0.000
T13	0.000	0.000	0.743	0.000	0.000
T14	0.000	0.000	0.000	0.509	0.000
T15	0.000	0.000	0.000	0.501	0.000
T16	0.000	0.000	0.000	0.609	0.000
T17	0.000	0.000	0.000	0.624	0.000
T18	0.000	0.000	0.000	0.637	0.000
T19	0.000	0.000	0.000	0.502	0.000
T20	0.000	0.000	0.000	0.000	0.654
T21	0.000	0.000	0.000	0.000	0.616
T22	0.000	0.000	0.000	0.000	0.655
T23	0.000	0.000	0.000	0.000	0.728
T24	0.000	0.000	0.000	0.000	0.602
T25	0.000	0.000	0.000	0.000	0.462
T26	0.000	0.000	0.000	0.000	0.544

Table 5
Interfactor Phi Matrix for Analysis II

	VERBAL	VISUAL	SPEED	MEMORY	MATH
VERBAL	1.000				
VISUAL	0.590	1.000			
SPEED	0.462	0.574	1.000		
MEMORY	0.502	0.637	0.599	1.000	
MATH	0.757	0.872	0.645	0.762	1.000

Table 6
Modified Maximum Likelihood Factor Matrix for Analysis II

	VERBAL	VISUAL	SPEED	MEMORY	MATH
T1	0.000	0.725	0.000	0.000	0.000
T2	0.000	0.495	0.000	0.000	0.000
T3	0.000	0.517	0.000	0.000	0.000
T4	0.000	0.688	0.000	0.000	0.000
T5	0.803	0.000	0.000	0.000	0.000
T6	0.816	0.000	0.000	0.000	0.000
T7	0.838	0.000	0.000	0.000	0.000
T8	0.703	0.000	0.000	0.000	0.000
T9	0.844	0.000	0.000	0.000	0.000
T10	0.000	0.000	0.786	0.000	0.000
T11	0.000	0.000	0.663	0.000	0.000
T12	0.000	0.000	0.703	0.000	0.000
T13	0.000	0.325	0.466	0.000	0.000
T14	0.000	0.000	0.000	0.508	0.000
T15	0.000	0.000	0.000	0.499	0.000
T16	0.000	0.000	0.000	0.604	0.000
T17	0.000	-0.491	0.000	0.629	0.000
T18	0.000	0.000	0.000	0.639	0.000
T19	0.000	0.000	0.000	0.503	0.000
T20	0.000	0.000	0.000	0.000	0.670
T21	0.000	0.000	0.349	0.000	0.423
T22	0.000	0.000	0.000	0.000	0.668
T23	0.000	0.000	0.000	0.000	0.733
T24	0.000	-0.680	0.486	0.000	0.333
T25	0.000	0.000	0.000	0.000	0.498
T26	0.000	0.000	0.000	0.000	0.566

Table 7
Modified Interfactor Phi Matrix for Analysis II

	VERBAL	VISUAL	SPEED	MEMORY	MATH
VERBAL	1.000				
VISUAL	0.574	1.000			
SPEED	0.401	0.369	1.000		
MEMORY	0.502	0.596	0.554	1.000	
MATH	0.746	0.848	0.400	0.719	1.000

Appendix A
Tests Included in Holzinger and Swineford (1939) Data Set

Tests of Visualization

TEST 1--Visual Perception Test (from Spearman VPT, Part III)
TEST 2--Cubes Test (Simplification of Brigham's Spatial Relations Test)
TEST 3--Paper Form Board Test (Shapes that can be combined to form a targeted criteria)
TEST 4--"Lozenges" Test (from Thorndike--Shapes that can be inverted and then combined to identify a targeted criteria)
TEST 25--Paper Form Board Test (Revision of TEST 3)
TEST 26--"Flags" Test (Possible Substitute for TEST 4--Lozenges)

Tests of Verbal Intelligence

TEST 5--General Information Verbal Test
TEST 6--Paragraph Comprehension Test
TEST 7--Sentence Completion Test
TEST 8--Word Classification Test (Selecting word which does not belong in a set)
TEST 9--Word Meaning Test

Tests to Determine Test-taking Speed

TEST 10--Speeded Addition Test
TEST 11--Speeded Code Test (Transforming shapes into alpha with code)
TEST 12--Speeded Counting Test (Counting of dots in various shapes)
TEST 13--Speeded Discrimination Test (Discriminating straight and curved capitals)

Tests of Memory/Recognition

TEST 14--Memory of Target Words Test
TEST 15--Memory of Target Numbers Test
TEST 16--Memory of Target Shapes Test
TEST 17--Memory of Object-Number Association Targets Test
TEST 18--Memory of Number-Object Association Targets Test
TEST 19--Memory of Figure-Word Association Targets Test

Tests of Mathematical Ability

TEST 20--Deductive Math Ability Test
TEST 21--Math Number Puzzles Test
TEST 22--Math Word Problem Reasoning Test
TEST 23--Number Series Test (Completion of number series)
TEST 24--Woody-McCall Mixed Math Fundamentals Test

Appendix B
LISREL VI Commands for Running Analysis II

```
USERPROC NAME=LISREL
TI THEORETICAL CONFIRM ANALYSIS #2
DA NI=26 NO=0 MA=KM
NA
'T1' 'T2' 'T3' 'T4' 'T5' 'T6' 'T7' 'T8' 'T9' 'T10' 'T11' 'T12'
'T13' 'T14'
'T15' 'T16' 'T17' 'T18' 'T19' 'T20' 'T21' 'T22' 'T23' 'T24' 'T25'
'T26'
RA
MO NX=26 NK=5 PH=ST
LK
'VERBAL' 'VISUAL'      'SPEED'     'MEMORY'      'MATH'
PA LX
0             1             0             0             0
0             1             0             0             0
0             1             0             0             0
0             1             0             0             0
1             0             0             0             0
1             0             0             0             0
1             0             0             0             0
1             0             0             0             0
1             0             0             0             0
0             0             1             0             0
0             0             1             0             0
0             0             1             0             0
0             0             1             0             0
0             0             0             1             0
0             0             0             1             0
0             0             0             1             0
0             0             0             1             0
0             0             0             1             0
0             0             0             1             0
0             0             0             0             1
0             0             0             0             1
0             0             0             0             1
0             0             0             0             1
0             0             0             0             1
0             0             0             0             1
0             0             0             0             1
OU SE TV MR MI FS SL=.05 TM=1200
END USER
```