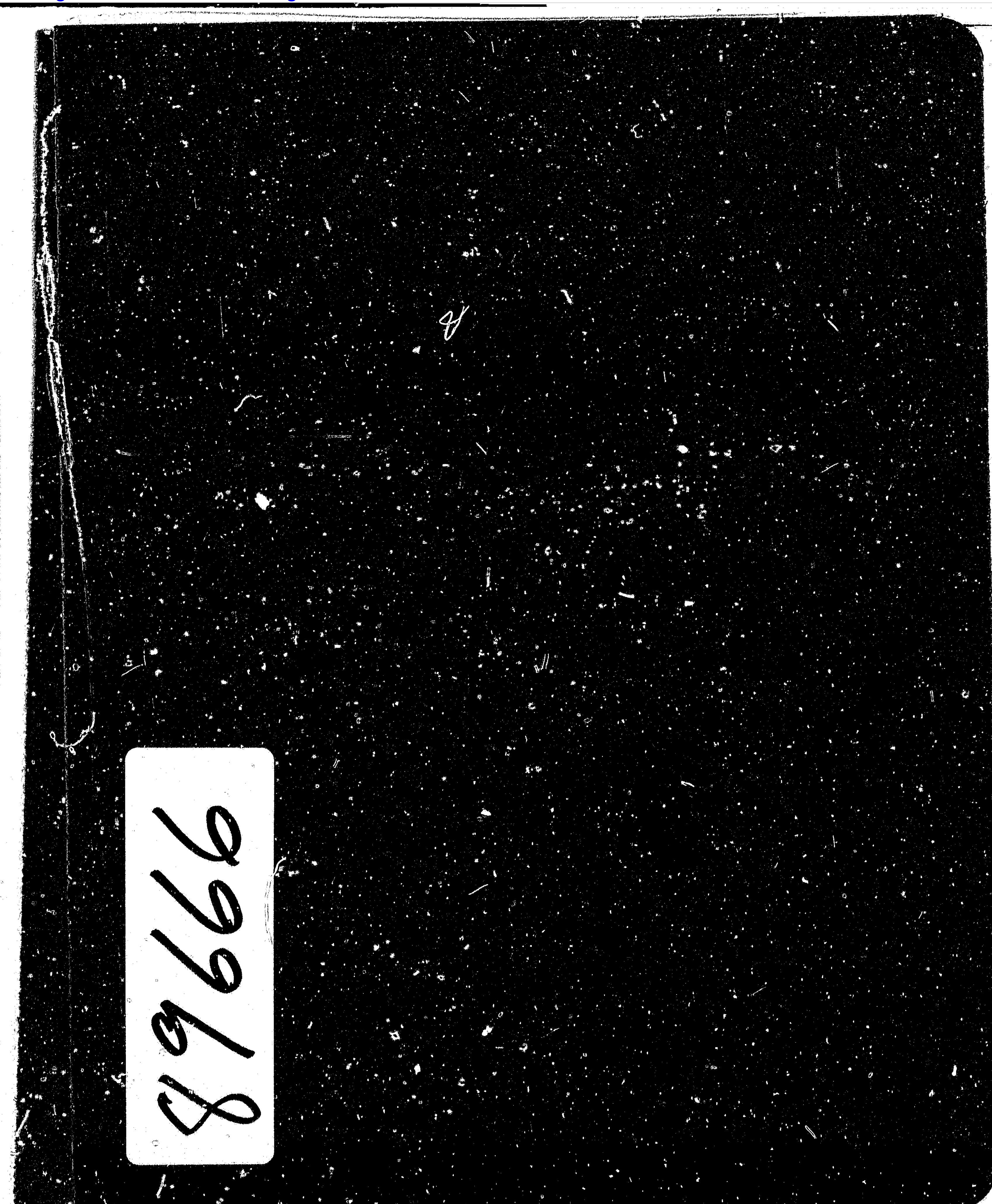
89966

29666

Inverted Factor Analysis: An Evaluation using Benchmark Data Sets

Craig Edelbrock

University of Pittsburgh School of Medicine

Michael Reed

Oregon State University

This research was supported by grant ~~#1-0260-8-PA-IJ~~ 81-IJ-CX-0059 from the Methodology Development Program, National Institute of Law Enforcement and Criminal Justice, to the first author. Correspondence may be addressed to Craig Edelbrock, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15261.

Abstract

Inverted factor analysis was evaluated on 20 previously studied multivariate mixtures. Two methods of determining number of factors and two rotational methods--orthogonal varimax and oblique direct quartimin--were compared. Objects were assigned to groups on the basis of highest absolute factor loadings, with the mimimum loading required for assignment systematically varied. Rotational methods did not differ significantly in either accuracy or coverage of the resulting classifications. Paradoxically, setting the number of factors equal to the number of underlying populations resulted in less accurate solutions than determining the number of factors empirically by Cattell's scree test. More importantly, the inverted factoring technique was found to be as accurate as the best hierarchial clustering algorithms previously tested on these mixtures. Thus, despite the implausibility of the factor analytic model for generating typologies--and numerous other problems and criticisms--inverted factor analysis appears to be a useful taxonomic tool.

Inverted Factor Analysis: An Evaluation using Benchmark Data Sets

Inverted factor analysis, also known as Q-factor analysis, inverse factor analysis, and profile factor analysis, is one of the oldest and most widely used procedures for constructing typologies in the behavioral sciences. The basic rationale of this procedure has not changed since Stephenson introduced the "inverted factor technique" in 1936. In fact, Stephenson's original articles (1936a,b) still provide a lucid introduction to the method. In the past 30 years, inverted factor analysis has been used in numerous studies to identify subtypes of individuals, particularly in the areas of psychiatry and deviant behavior (Butler & Adams, 1966; Collins, Burger & Taylor, 1976; Fleiss, Lawlor, Platman & Fieve, 1971; Guertin, 1952, Katz & Cole, 1963; Monro, 1955; Overall, Hollister, Johnson & Pennington, 1966; Raskin & Crook, 1963). The inverted factor technique has also been discussed in several methodological treatises (Baggaley, 1964; Broverman, 1961; Cattell, 1952; Morf, Miller and Syrotuik, 1976; Overall and Klett, 1972; Ross, 1963; Ryder, 1966; Stephenson, 1953).

Despite its historical precedence and diverse applications, inverted factor analysis has been strongly criticized as a method of generating typologies (Baggaley, 1964; Fleiss et al, 1971; Fleiss, 1972; Fleiss & Zubin, 1969; Jones, 1968; Lorr, 1966). A standard criticism has involved the use of the product-moment correlation index similarity between individuals. The correlation coefficient indexes similarity only in profile shape, not elevation or scatter. Moreover, a correlation of 1.00 does not necessarily indicate that two profiles have identical shape, but only that they are linear functions of one another (see Edelbrock & McLaughlin, 1980; and Fleiss & Zubin, 1969 for more detailed discussions). This criticism is not unique to inverted factor analysis, however, in that a variety of clustering methods can employ the

correlation coefficient as the measure of profile similarity (e.g., Carlson, 1972; Edelbrock, 1979; Lorr, Bishop & McNair, 1965; Lorr & Radhakrishnan, 1967).

Several other criticisms have been raised that are more pertinent to the inverted factoring technique. Fleiss and Zubin (1969), for example, have questioned the appropriateness of the linear model underlying factor analysis to the task of generating typologies of individuals. In particular, they have asked whether it makes any sense to say that an individual represents "X" amount of one type plus "Y" amount of another type, and so on. Lorr (1966) has further questioned the rationale of rotating Q-factors. Even if unrotated factor loadings represent similarity to underlying "types", what is the meaning of transforming such loadings so as to better approximate simple structure?

Fleiss and Zubin have also objected that the number of types one may identify is limited by the number of variables in the analysis. The maximum number of factors that can be extracted from a correlation matrix is equal to the rank of the matrix. For a matrix of Q-correlations, rank is at most p - 1, where p equals the number of variables. Fleiss and Zubin therefore reasoned that the maximum number of types one can identify is equal to the number of variables minus 1. This is obviously a problem when one has few variables with which to work, but seeks to identify several types of individuals.

Achenbach and Edelbrock (1981) have noted an additional problem involving procedures of factor extraction. Since the first factor typically extracts the most variance from the correlation matrix, it will encompass more individuals having high loadings than subsequent factors. This bias towards constructing one large group followed by successively smaller and smaller groups is rarely justified in taxonomic research. Clearly, the relative size of the groups should be determined by the data, not the taxometric procedure. Other methodological problems and issues include: (a) translating factor scores into discrete groups of individuals, (b) determining the appropriate number of factors, and (c) selecting a rotational procedure. The latter two problems also arise in regular R-factor analysis and have been discussed in detail elsewhere (cf. Mulaik, 1972; Harman, 1976).

Given these problems and criticisms, one would expect inverted factor analysis to have been laid to rest long ago--but this is not the case. More than 40 years after its inception, the technique is still in use. Furthermore, it has generated heuristically valuable and predictive typologies. A particularly good example is the nosology of depression constructed by Overall et al (1966). Using an inverted factoring procedure (Overall & Porterfield, 1963), three subtypes of depressed patients were identified, based on scores on the Brief Psychiatric Rating Scale. In a subsequent double-blind comparison, the three subtypes (labelled Anxious, Hostile, and Retarded) were found to differ markedly in terms of response to anti-depressant drugs. The value of inverted factor analysis in taxonomic research has been corroborated by several other recent studies (Collins et al, 1976; Evenson, Altman, Sletten & Knowles, 1973; Kunce, Ryan & Eckelman, 1976; Meyer & Kline, 1977; Raskin & Crook, 1976).

A Reconsideration

There are several compelling reasons for reconsidering inverted factor analysis as a taxonomic tool. For one, fruitful applications of the technique would appear to mitigate any methodological criticisms. Second, some points of criticism are patently wrong. For example, although the number of factors may be limited to $p - 1$, the number of types is not limited to the number of factors. In practice, inverted factor analysis may yield bipolar factors

comprised of both positive and negative loadings. Such bipolar factors have been taken to represent two underlying types manifesting opposite patterns of scores. Carlson et al (1976), for instance, obtained only four factors, but because each was bipolar, eight subtypes were identified. Third, some criticisms are based on dogma, not empirical facts. Some recent studies suggest that long-established psychometric dogma is in desperate need of revision. For example, despite the so-called "superiority" of distance measures for indexing profile similarity (e.g., Eades, 1965; Fleiss & Zubin, 1969: p. 239), recent Monte Carlo studies of hierarchical clustering methods have shown that correlation yields substantially better recovery of underlying mixture populations than Euclidean distance (Edelbrock, 1979; Edelbrock & McLaughlin, 1980). Finally, there have been very few attempts to test the inverted factoring technique empirically against other methods. One exception is the recent study by Blashfield and Morey (1980). Using Monte Carlo procedures, data sets designed to mimic MMPI psychotic, neurotic, and personality disorder patterns were generated then analyzed by inverted factor analysis, Lorr's non-hierarchical clumping procedure (Lorr et al, 1965), a hierarchical clustering algorithm called average linkage, and Ward's (1963) minimum variance technique. Blashfield and Morey concluded that the average linkage method yielded the best clustering solutions. For some data sets, however, the inverted factoring technique resulted in substantially fewer misclassifications than the other three methods.

Purpose of this Study

The purpose of this study was to evaluate inverted factor analysis on a standard set of multivariate mixtures. This research builds on Blashfield and Morey's recent study in the following ways: (a) a broad range of multivariate mixtures differing in number of variables, number of underlying populations, difficulty of solution, etc., were analyzed, (b) two methods of determining number of factors were tested, (c) two rotational procedures--one orthogonal the other oblique--were compared, and (d) the effects of varying the minimum loading required for classification were systematically evaluated. In addition, comparisons between the inverted factoring technique and several hierarchical clustering algorithms were made.

Methods

Data Sets

It has been argued previously (Edelbrock, 1979; Edelbrock & McLaughlin, 1980) that evaluations of taxometric methods should include test on "benchmark" data sets--that is, data sets have been well-characterized, are available to other investigators, and have been used in previous mixture model studies. Such benchmark data sets provide a common standard against which to compare clustering and classification methods and thus increase the generalizability of mixture model tests. With this in mind, 20 multivariate normal mixtures generated by Blashfield (1976) were selected for this study. These mixtures mimic real data in many ways, including (a) representative range of number of variables and populations, (b) quasi-normal distribution parameters, (c) addition of "measurement" error to scores, and (d) varying strength and complexity of the covaviance structure of the underlying populations. These data sets have also been used in previous tests of hierarchical clustering algorithms (Blashfield, 1976; Edelbrock, 1979; Edelbrock & McLaughlin, 1980), so direct comparisons across studies are possible.

Procedures

Data were double-centered according to the rationale and procedure given by Overall and Klett (1972; pp. 203-204). Variables were standardized (mean = 0, sd = 1) and scores were then standardized equivalently across objects.

Each of the 20 (object X variable) data sets was then inverted (i.e., to represent a variable X object matrix) and subjected to principal-components factor analysis using the BMDP4M - program. It is important to note that double-centering the data results in bipolarity of the unrotated factors. However, it does not necessarily result in bipolarity in the rotated factors, which were used here.

Two procedures were used to determine the number of factors. First, for each mixture, the number of factors was set to equal the number of underlying populations. Since the rotated factors were not bipolar, each factor comprised only one group of objects having high loadings in the same direction. Thus, determing the number of factors in this way is tantamount to setting the number of groups (j) equal to the number of underlying populations (k). These 20 analyses are subsequently designated by the notation $j = k$.

Second, the number of factors was determined by examining eigen values. For these data sets, the commonly used "eigen value greater than 1" rule resulted in considerable over-factoring. A few factors having large eigen values were obtained followed by several having eigen values slightly greater than 1.00. This problem was also encountered by Blashfield and Morey (1980). Following their procedure, Cattell's (1966) scree test was used to determine number of factors. In this study, both investigators examined the eigen value plot for each mixture and independently selected the number of factors. Although we agreed for all 20 mixtures, the number of factors indicated by the scree test did not always equal the number of underlying populations. For eight mixtures, the number of factors equalled one more than the number of underlying populations (i.e., $k + 1$). These 20 analyses are subsequently designated by the notation $j \neq k$ (i.e., the number of groups did not necessarily equal the number of populations).

One issue in factor analysis is whether to construct orthogonal (uncorrelated) or oblique (correlated) factors. This is an important consideration when deriving typologies because rotational procedures substantially affect final factor loadings, which are the basis for constructing groups. Most previous applications of inverted factor analysis (e.g., Blashfield & Morey, 1980; Collins et al, 1976; Fleiss et al, 1971; Katz & Cole, 1965) involved the varimax rotation--an orthogonal procedure. In this study, both varimax (orthogonal) ind direct quartimin (oblique) rotations were compared. This yields four analyses of 20 mixtures each: $j = k$ and $j \neq k$ with either varimax or direct quartimin rotation.

A crucial issue that arises in inverted factor analysis involves translating factor loadings into discrete groups of objects or individuals. A common procedure has been to assign individuals to groups on the basis of highest factor loadings (in terms of absolute value). Some investigators have specified a minimum loading required for classification. Fleiss et al (1971), for example, selected a minmum loading of .40. Individuals whose highest loadings were less than .40 were left unclassified. In their Monte Carlo study, Blashfield and Morey (1980) selected a minimum loading of .60, with the additional criterion that an object could not have a loading of .60 or higher on any other factor. These rather stringent criteria reduce coverage substantially, but result in more distinct and homogeneous groups.

In this study, objects were assigned to groups on the basis of their highest loadings. This is a simple procedure for constructing groups, but the coverage of the resulting classification can be manipulated by simply changing the minimum loading required for assignment. A low cutoff point results in the classification of a high proportion of objects into relatively heterogenous groups, whereas a high cutoff point results in the classification of a low proportion of objects into more distinct, non-overlapping groups. This

assignment procedure therefore makes it possible to evaluate classifications at several levels of coverage.

Calculating Accuracy

The accuracy of the inverted factor solutions was defined as the agreement between the obtained groups and the underlying populations in the mixtures. A wide variety of statistics have been used to measure accuracy in mixture model studies, and there is little consensus regarding the "best" accuracy measure. Kappa (Cohen, 1960) and Rand's statistic (Rand, 1971) have been used in many studies (e.g., Blashfield, 1976; Edelbrock, 1979; Edelbrock & McLaughlin, 1980; Kuiper & Fisher, 1975; Milligan & Isaac, 1980; Mojena, 1977; Rand, 1971). Both of these measures have drawbacks. Kappa has the advantage of correcting for chance level of agreement in a cross-classification, but it is appropriate only for square matrices (i.e., $j = k$). Rand's statistic does not require that $j = k$, but the scale is not uniform from matrix to matrix. That is, the lower bound of Rand's statistic is not zero but is determined by the marginal distributions of the cross-classification.

One way to overcome the idiosyncracies inherent in individual measures is to use multiple criteria for evaluating accuracy. Six measures, including kappa, Rand's statistic, asymmetric lambda, tau, Kramer's v, and the contingency coefficient were used in this study. We chose to report our main findings in terms of asymmetric lambda for several reasons. This statistic is appropriate for nominal level cross-classifications, has a range of zero to 1.00, and can be used with either square ($j = k$) or rectangular ($j \neq k$) matrices. The "asymmetrical" aspect of this statistic also seems well suited to the task of measuring accuracy. The term "asymmetrical" refers to the fact that lambda indexes the degree to which one classification predicts another, and not vice versa. In mixture model studies, the underlying populations comprise a fixed or dependent classification, predicted by empirically derived groups that are free to vary.

Although we report our main findings in terms of asymmetric lambda, we also report summary statistics in terms of kappa and Rand's statistic. This permits direct comparisons with previous studies. Finally, it is worth noting that our conclusions regarding the relative accuracy of various methods were identical for all six measures we explored. This is not surprising, since such measures are all founded on the same information extracted from the cross-classification matrix (cf. Hubert & Levin, 1976). Furthermore, in these analysis, the six measures of accuracy correlated >.95 with one another.

Statistical Analyses

For each of the 80 inverted factor solutions, objects were classified according to their highest loadings. Accuracy was then calculated at seven levels of coverage dictated by the following minimum loadings: .0, .4, .5, .6, .7, .8, and .9. These minimum loadings between were selected because: (a) all objects had highest loadings greater than .0, thus a cutoff point of .0 yields 100% coverage, (b) very few objects had highest loadings between .0, and .4 so accuracy and coverage varied little in this interval, and (c) there were too few loadings above .9 to calculate accuracy.

Accuracy and coverage values were analyzed in separate 2 x 2 x 7 analyses of variance representing: number of factors (j = k vs. j ≠ k) rotational methods (varimax vs. direct quartimin), and minimum loading (.0 to .9), respectively.

Results

Main results are portrayed graphically in Figures 1 and 2. These figures

----

Insert Figures 1 and 2 here

----

show the relations between the minimum loading required for classification and both accuracy (left axis) and coverage (right axis). Figure 1 depicts accuracy and coverage functions for the j = k solutions, whereas Figure 2 despicts results for the j ≠ k solutions. Overall, accuracy and coverage were significantly related to the minimum loading (p <.001), but in opposite ways. Raising the minimum loading uniformly increased accuracy, but decreased coverage to a greater and greater extent. No significant differences (F <1.00) were detected between varimax and direct quartimin rotations for either j = k or j ≠ k solutions. Varimax solutions resulted in consistently higher accuracy and coverage, however.

Paradoxically, j ≠ k solutions resulted in significantly <u>higher</u> accuracy and coverage than j = k solutions (p <.01). This was the case for both rotational methods. Figure 3 portrays accuracy differences between j = k and j ≠ k solutions in a manner that equates them for coverage. Accuracy is shown as a function of coverage, rather than as a function of the minimum loadings as in Figures 1 and 2. At all levels of coverage, j ≠ k solutions resulted in significantly higher accuracy and j = k solutions. Examination

----

Insert Figure 3 here

----

of the eight mixtures where j ≠ k confirmed that constraining the number of factors to equal the number of underlying groups substantially reduced accuracy. For these mixtures, higher accuracy was achieved when the number of groups was determined empirically by Cattell's scree test.

Comparisons with Other Methods

In a previous study (Edelbrock & McLaughlin, 1980), 18 hierarchical clustering algorithms were tested on the 20 benchmark mixtures. The algorithms included single, complete, average, and centroid linkage using either Euclidean distance, correlation, or the one-way or two-way intraclass correlation as the similarity measure; Ward's minimum variance technique; and a random algorithm used to establish a baseline control for evaluating methods. Two problems arise when making comparisons between inverted factor analysis and these hierarchical methods. First, the accuracy of each hierarchical method was calculated for j = k. That is, the number of clusters alwiys equaled the number of underlying populations. To make direct comparisons, it is necessary to select inverse factor solutions were j = k. This is unfortunate because j = k solutions were significantly less accurate than j ≠ k solutions. Comparisons are therefore based on a conservative estimate of the accuracy of the inverted factoring technique.

The second problem involves selecting the level of coverage at which to make comparisons. Whereas both inverted factor analysis and the hierarchical methods can yield classifications varying in coverage, this occurs in quite different ways. For inverted factor analysis, coverage depends upon the minimum loading required for assignment. For the hierarchical methods, coverage depends on the selection of the best j clusters at various levels in the hierarchical tree. This difference appears to represent a bias in favor of the hierarchical methods. For each mixture, the accuracy of the inverted factor solution is based on the <u>same</u> set of factors--only the minimum loading is varied. The accuracy of each hierarchical solution, on the other hand, is based on <u>different</u> sets of clusters, selected so as to maximize accuracy at each level in the hierarchical tree. This bias is evidenced by the fact that the accuracy of even the <u>random</u> hierarchical algorithm increases as coverage delcines (see Edelbrock & McLaughlin, 1980: p. 310).

To make comparisons between methods, accuracies of the j = k varimax solutions were calculated at 100% coverage. Focusing on 100% coverage eliminates the biases that can arise at lower levels of coverage. Furthermore, inverted factor analysis and the best hierarchical methods show uniform increases in accuracy as coverage declines. Thus, differences at 100% coverage are likely to be representative of differences at lower levels of coverage.

The mean kappa value for the varimax solutions equaled .65, which compares quite favorably with accuracies previously reported by Edelbrock & McLaughlin (1980: p. 310). Specifically, the inverted factoring technique was substantially more accurate than 10 of the 18 hierarchical algorithms: single and complete linkage using any of the four similarity measures, average and centroid linkage using Euclidean distance, and the random algorithm.

The j = k varimax solutions were also compared with the most accurate hierarchical algorithm--average linkage using the one-way intraclass correlation. Mean values for kappa, Rand's statistic, and asymmetric lambda, as well as paired <u>t</u>-test results, are shown in Table 1. According to all three measures, the average linkage algorithm was slightly more accurate than inverted factor analysis, but not significantly so.

Insert Table 1 here

As a final test, Rand's statistic was calculated for varimax j ≠ k solutions: which represent the highest accuracy attained by inverted factor analysis. Edelbrock and McLaughlin previously used Rand's statistic to evaluate the "best possible" clustering solutions attained by the 18 hierarchical methods they examined. Direct comparisons between methods are therefore possible. The average Rand value for the inverted factoring technique was .862. This is higher than 11 of the 18 hierarchical methods, and not significantly different than the most accurate hierarchical algorithm (see Edelbrock & McLaughlin, 1980: p. 311).

Discussion

Inverted factor analysis is one of the most widely used and widely criticized procedures for constructing typologies in the behavioral sciences. Unfortunately, some critics of the method have simply argued: "It shouldn't work, therefore it doesn't." Few commentators have backed up their criticisms with empirical evidence. In this evaluation, the inverted factoring technique yielded more accurate recovery of underlying populations than many previously studied hierarchical algorithms. Moreover, the inverted factor technique was found to be among the most accurate methods yet tested on these benchmark mixtures. These results agree with the previous study by Morf, Miller and Syrotuik (1976) who, on the basis of an objective comparison, concluded that inverted factor analysis was superior to the complete linkage algorithm in identifying subtypes of individuals. Thus, inverted factor analysis appears to be a useful taxonomic tool--despite the implausibility of the factor analytic model for generating typologies, the "inferiority" of the correlation coefficient as a measure of profile similarity, and numerous other problems (e.g., determining number of factors, assigning objects to groups, etc.).

In terms of recovering underlying mixture populations, differences between rotational methods were minimal. The more crucial methodological problem involved selecting the appropriate number of factors. Determining the number of factors empirically via Cattell's scree test resulted in more accurate solutions than the alternative procedure of setting the

number of factors equal to the number of populations. Blashfield and Morey (1980) also reported that the scree test was quite accurate in determining the correct number of populations in their MMPI Monte Carlo data. This is a potentially important finding because the scree test does not depend upon a priori knowledge regarding "true" underlying populations. Thus, this procedure may be useful in determining number of underlying groups in applications to real data. This is a major asset of the inverted factor technique. Hierarchical clustering algorithms, by contrast, do not produce a discrete number of clusters, but rather a hierarchical arrangement of objects and groups. Determining the appropriate number of clusters is an unsolved problem, although some work has been done on developing objective criteria for making this decision (e.g., Mojena, 1977).

The inverted factor technique also embodies a simple mechanism for manipulating the coverage of the resulting classifications. In this study, for example, objects were assigned to groups on the basis of their highest factor loadings. Raising the minimum loading required for assignment decreased coverage, but increased accuracy. The ability to vary coverage may be valuable in research applications. In an epidemiological study, for instance, high coverage may be desirable in order to account for the generality and distribution of phenomena in a population. In other situations, it may be advantageous to construct extremely homogeneous groups. This would dictate low coverage, but the resulting groups would encompass individuals representing relatively "pure types". Future research should explore different methods of translating factor loadings into groups. The dual cutoff criteria used by Blashfield and Morey (1980), for example, appear promising. Such stringent assignment rules result in reduced coverage, but yield more homogeneous and distinct groups. Moreover, such assignment rules may yield typologies that are more predictive of external criteria.

Finally, additional comparisons among clustering and classification methods are needed. There are few standard procedures for constructing empirically based taxonomies and little is known about the relative merits of different methods. Objective comparisons are necessary, not only to combat dogmatic arguments for or against specific approaches, but also to identify those procedures best suited to behavioral research. The results obtained here indicate that inverted factor analysis yields accurate recovery of underlying populations from multivariate normal mixutres. Evaluations on other types of mixtures and evaluations involving other criteria (e.g. replicability, sensitivity to data perturbation, etc.) would be valuable.

Achenbach, T.M., and Edelbrock, C. Taxonomic issues in child psychopathology. In T.H. Ollendick and M. Hersen (Eds.) Handbook of Child Psychopathology. New York: Plenum Press, in press, 1981.

Baggaley, A.R. Intermediate correlational methods. New York: Wiley, 1964.

Blashfield, R.K. Mixture model tests of cluster analysis: Accuracy of four agglomerative hierarchical methods. Psychological Bulletin, 1976, 83, 377-388.

Blashfield, R., & Morey, L.C. A comparison of four clustering methods using MMPI Monte Carlo data. Applied Psychological Measurement, 1980, 4, 57-64.

Broverman, D.M. Effects of score transformations in Q and R factor analysis. Psychological Review, 1961, 68, 68-80.

Butler, E.W., and Adams, S.W. Typologies of delinquent girls: some alternative approaches. Social Forces, 1966, 44, 401-407.

Carlson, K.A. Classes of adult offenders: A multivariate approach. Jouranl of Abnormal Psychology, 1972, 79, 84-93.

Cattell, R.B. The three basic factor-analytic research designs - their interrelations and derivatives. Psychological Bulletin, 1952, 49, 499-520.

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

Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.

Collins, H.A., Burger, G.K., and Taylor, G.A. An empirical typology of heroin abusers. Journal of Clinical Psychology, 1976, 32, 473-476.

Eades, D.C. The inappropriateness of the correlation coefficient as a measure of taxonomic resemblance. Systematic Zoology, 1965, 14, 98-100.

Edelbrock, C. Mixture model tests of hierarchical clustering algorithms: The problem of classifying everybody. Multivariate Behavioral Research, 1979, 14, 367-384.

Edelbrock, C., and McLaughlin, B. Hierarchical cluster analysis using intraclass correlations: A mixture model study. Multivariate Behavioral Research, 1980, 15, 299-318.

Evenson, R.C., Altman, H., Sletten, I.W., & Knowles, R. Factors in the description and grouping of alcoholics. American Journal of Psychiatry. 1973, 130, 49-53.

Fleiss, J.L. Classification of the depressive disorders by numerical typology. Journal of Psychiatric Research, 1972, 9, 141-153.

Fleiss, J.L., Lawlor, W., Platman, S.R., and Fieve, R.R. On the use of inverted factor analysis for generating typologies. Journal of Abnormal Psychology, 1971, 77, 127-132.

Fleiss, J.L. and Zubin, J. On the methods and theory of clustering. Multivariate Behavioral Research, 1969, 4, 235-250.

Guertin, W.H. An inverted factor--analytic study of schizophrenics. Journal of Consulting Psychology, 1952, 16, 371-375.

Harman, H.H. Modern factor analysis, 3rd Edition. Chicago: University of Chicago Press, 1976.

Hubert, L.J., and Levin, J.R. Evaluating object set partitious: Free sort analysis and some generalizations. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 459-470.

Jones, K.J. Problems of grouping individuals and the method of modality. Behavioral Science, 1968, 13, 496-511.

Katz, M.K., and Cole J.O. A phenomenological approach to the classification of schizophrenic disorders. Diseases of the Nervous System, 1963, 24, 147-154.

Kuiper, F.K. and Fisher, L. A Monte Carlo comparison of six clustering procedures. Biometrics, 1975, 31, 777-783.

Kunce, J.T., Ryan, J.J., and Eckelman, C.C. Violent behavior and differential WAIS characteristics. Journal of Consulting and Clinical Psychology. 1976, 44, 42-45.

Lorr, M. (Ed.) Explorations in typing psychotics. Oxford: Pergamon Press, 1966.

Lorr, M., Bishop, P.F., and McNair, D.M. Interpersonal types among psychotic patients. Journal of Abnormal Psychology, 1965, 70, 468-472.

Lorr, M., and Radhakrishnan, B.K. A comparison of two methods of cluster analysis. Educational and Psychological Measurement. 1967, 27, 531-543.

Meyer, L., and Kline, P. On the use of delegate scores in the G analysis of clinical data: A Q factor method for diagnostic classification. Multivariate Behavior Research, 1977, 12, 479-486.

Milligan, G.W. and Isaac, P. The validation of four ultrametric clustering algorithms. Pattern Recognition, 1980, 12, 41-50.

Mojena, R. Hierarchical grouping methods and stopping rules: An evaluation. Computer Journal, 1977, 20, 359-363.

Monro, A.B. Psychiatric types: A Q-technique study of 200 patients. Journal of Mental Science, 1955, 101, 330-343.

Morf, M.E., Miller, C.M., and Syrotuik, J.M. A comparison of cluster analysis and Q-factor analysis. Journal of Clinical Psychology, 1976, 32, 59-64.

Mulaik, S. The foundations of factor analysis. New York, McGraw-Hill, 1972.

Overall, J.E., Hollister, L.E., Johnson, M., and Pennington, V. Nosology of depression and differential response to drugs. Journal of the American Medical Association, 1966, 195, 162-164.

Overall, J., and Klett, C.J. Applied Multivariate Analysis, New York: McGraw-Hill, 1972.

Overall, J.E., and Porterfield, J.W. The powered vector method of factor analysis. Psychometrika, 1963, 61, 415-422.

Rand, W.M. Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association, 1971, 66, 846-850.

Raskin, A. and Crook, T.H. The endogenous-neurotic distincition as a predictor of response to antidepressant drugs. Psychological Medicine, 1976, 6, 59-60.

Ross, J. The relation between test and person factors. Psychological Review, 1963, 70, 432-443.

Ryder, R.G. Profile factor analysis and variable factor analysis. Psychological Reports, 1964, 15, 119-127.

Stephenson, W. Introduction to inverted factor analysis with some applications to studies of orexsis. Journal of Educational Psychology, 1936a, 27, 353-367.

Stephenson, W. The inverted factor technique. British Journal of Psychology, 1936b, 26, 344-361.

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

Ward, J.H. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 1963, 58, 236-244.

TABLE 1

Comparison between inverted factor analysis (varimax rotation) and the average linkage algorithm

Accuracy Measure	Method: Inverted Factor Analysis	Average Linkage	Paired t-value[a]
Kappa	.655	.793	1.49
Rand	.789	.864	1.44
Lambda	.656	.801	1.20

Note: Table entries are mean values for 20 mixtures. [a]df=19. None of the paired t-tests were significant (p >.10).

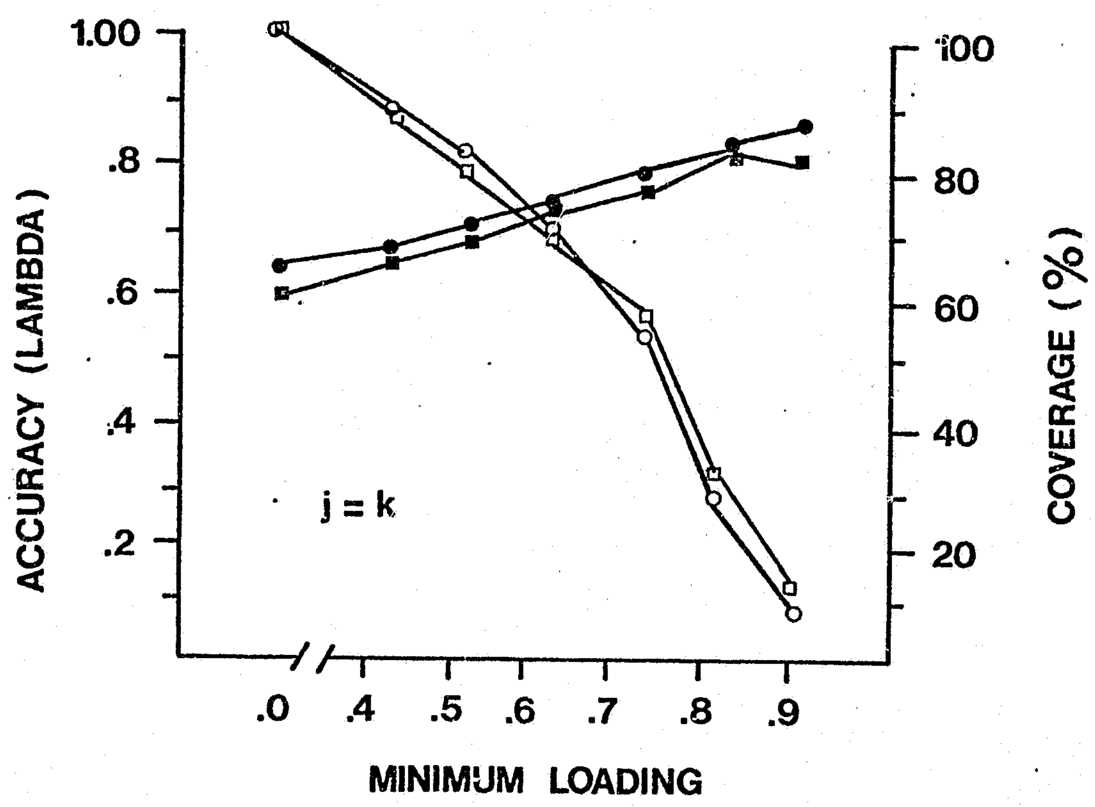

Fig 1. Accuracy and coverage functions for j = k factoring solutions (● varimax accuracy ○ varimax coverage ■ direct quartimin accuracy □ direct quartimin coverage).

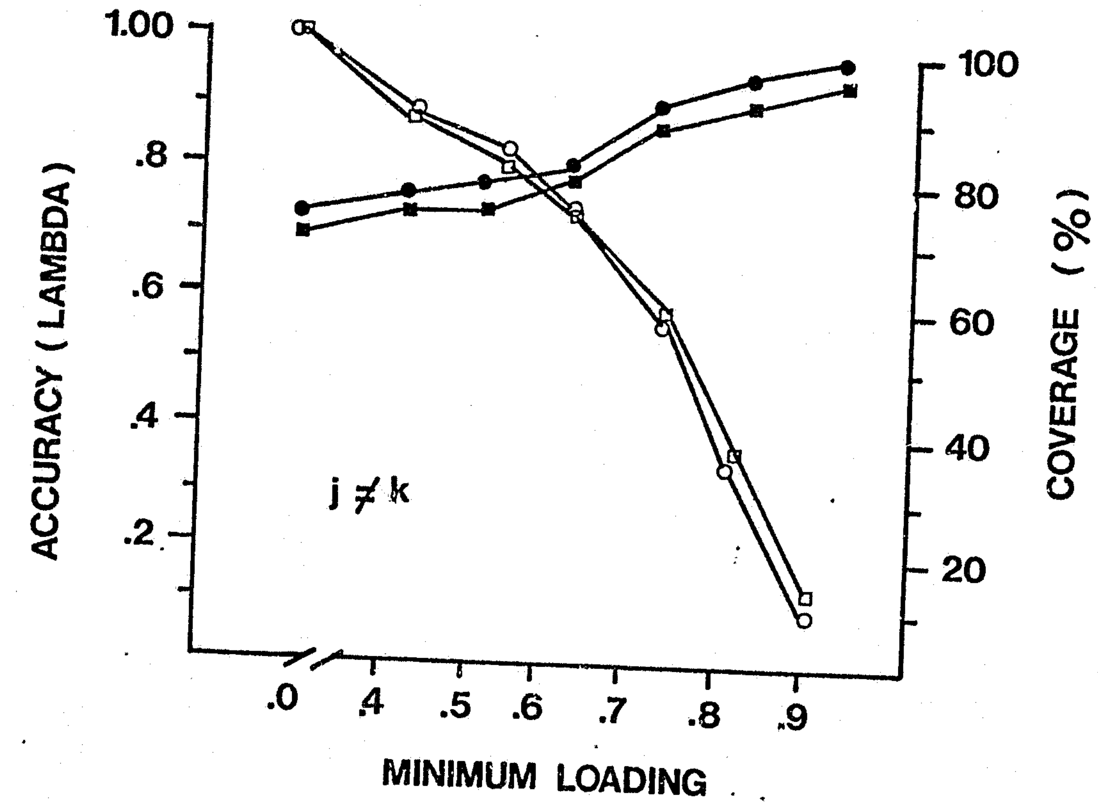

Fig 2. Accuracy and coverage functions for j ≠ k factoring solutions (● varimax accuracy ○ varimax coverage ■ direct quartimin accuracy □ direct quartimin coverage).

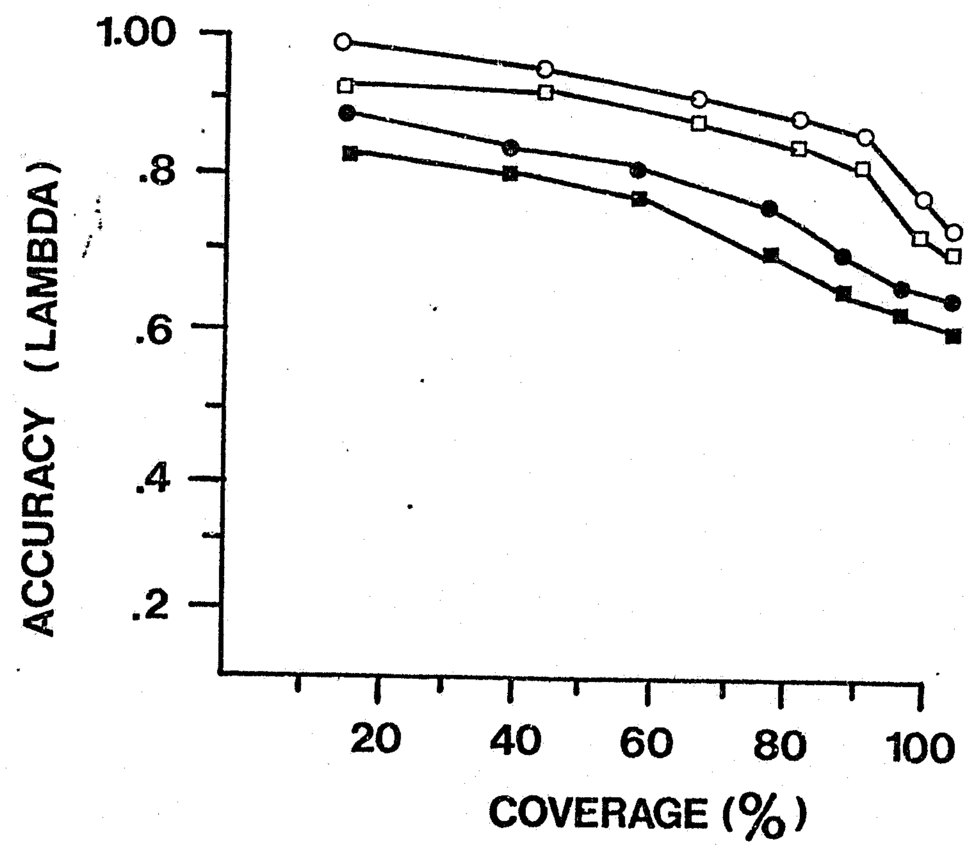

Fig 3. Comparison between j = k and j ≠ k factoring solutions (● j = k varimax ■ j = k direct quartimin ○ j ≠ k varimax □ j ≠ k direct quartimin).

END