A Method for Non-hierarchical Cluster Analysis

Based on Binary Relations and a Comparison

with Other Clustering Programs

Judith Toben Zatkin
Robert Cudeck
Douglas J. McCormick
Norman Cliff

Department of Psychology

University of Southern California

Technical Report 81-2
March, 1981

Funding for this research was provided by the National Institute of Justice Grant # 79-NI-AX-0065.

The authors wish to thank Linda M. Collins for several contributions to this research and Professors Peter M. Podgorny and Stephen A. Madigan for many valuable comments on earlier version of the report.

The purpose of this paper is to compare a new clustering program to existing programs. The new program was written to use a measure of association not available previously. All of the programs were used to analyze a single set of real data. The paper provides a review of literature on clustering, and the various aspects of an analysis are discussed. The new program is described, followed by a description of the data the programs will analyze. Finally, the results of all the clustering programs are given with a discussion of the similarities and differences among solutions.

Nature of Cluster Analysis

The term "cluster" has a very broad meaning, but one may say that clusters are groups of objects that are similar to each other. Objects with similar characteristics are grouped together and separated from those that have different qualities. Cluster analysis is used to find clusters from a set of objects, based on measurements of the objects, where the measurements can be either qualitative or quantitative. If the objects are measured on one dimension only, and there are not many objects, it is not too difficult to discover the groups of like objects, if the objects are well differentiated by the characteristic on which they were measured. But when items are measured on many dimensions and the relationships are complex, the bases on which items can be grouped may not be intuitively obvious.

Hartigan (1975) equated clustering with classification. He used the word as "a general term for formal, planned, purposeful, or scientific classification." Everitt (1974) pointed out that "some authors use the term classification to describe techniques for assigning individuals to groups having *a priori* labels." He described clusters as "continuous regions of (p-dimensional) space containing a relatively high density of points, separated from other such regions containing a relatively low density of points." This definition gives clusters a spatial sense, but deliberately does not restrict the shape of the clusters to being spherical. Everitt allows clusters to be elongated yet adjacent, with members of one cluster closer in a spatial sense to some members of another cluster than to all members of their own cluster.

Cluster analysis is used in many fields. For example, in medicine, diseases have been classified according to symptoms. Two recent papers in the psychiatric literature have dealt with clustering retarded adults and suicide attempters (Reid, et al, 1978; Paykel and Rassaby, 1978). A very extensive use of cluster analysis has occurred in biology where hierarchical techniques are often used to create taxonomies of plants and animals.

Although in biological work the objects are usually what is classified, it is also reasonable to cluster the variables. The relation between these two methods is analogous to that between Q and R-factor analysis. There are also procedures for clustering the variables and cases simultaneously into "blocks" (Hartigan, 1975).

Clustering Literature

There are a number of general reviews of cluster analysis. Sneath and Sokal (1973) wrote primarily about methods for biological taxonomies. They differentiated between classical taxonomies based on biological descent and phenetic relationships based on observed characteristics. They gave a clear explanation of the major types of cluster analysis, especially the hierarchical methods that are suited to the data of their interest.

Hartigan's *Clustering Algorithms* (1975) reviewed the basic issues and then presented a multitude of different types of clustering problems and FORTRAN programs that will cluster according to a variety of methods. He classified the methods not according to whether they seek to fit the data to a hierarchy, but rather by method of clustering.

Spath's (1980) recent book on clustering algorithms also included a series of FORTRAN programs. He considered hierarchical, non-hierarchical, and miscellaneous methods, and repeatedly analyzed a few data sets by many of the algorithms. Unfortunately, he did not give statistical comparisons of the varying solutions.

Everitt (1974) provided an excellent introduction to the issues of clustering. He described various types of algorithms and gave a detailed discussion of problems inherent in various methods of clustering and measurement. He pointed out limitations of some algorithms when attempting to recover unusually shaped clusters, and issue not dealt with in detail elsewhere. Everitt also included a chapter of practical guidelines for the user.

Components of Cluster Analysis

The user of cluster analysis must make a series of decisions that will affect the type of analysis to be performed and the solution obtained. Prior to the cluster analysis, the data must be collected. There are three things to consider when choosing the data: the variables, the subjects (or objects) they are measured on, and the scales of measurement. It is, of course, important to measure variables that are relevant to the purpose of the study. For instance, if the field of study is criminal behavior, eye color is probably irrelevant. Similarly, if the variables are being clustered, they should be observed on an appropriate set of objects. Of course, if the purpose of the study is to cluster patterns of criminal offenses, at least some of the subjects should be criminals.

Once the appropriate variables are selected and suitable objects or subjects chosen, the scales by which the data are measured must be considered. Data measured on any scale may be used, including nominal or dichotomous categories. If the data include variables measured on widely divergent scales, the investigator may wish to standardize or in some other way weight the variables.

The data is the first part of the problem; the clustering is the second. There are a number of parameters to a cluster analysis that the investigator must decide upon. They are the model, the index of measurement, and the method of clustering.

Model. Two basically different models have been discussed in the cluster analysis literature: hierarchical arrangements and non-hierarchical groups. In addition, the groups may or may not be overlapping. Although some data lend themselves more to one model or another, sometimes the choice of model is not clear. Unfortunately, the choice of model imposes a structure on the solution that might be inappropriate. Some guidelines can be found in the literature. "In any specific application, whether one uses hierarchical or non-hierarchical methods is largely dependent on the meaningfulness, in the particular situation, of the tree structure imposed by hierarchical clustering procedures" (Gnanadeskian, 1977). Everitt suggests that "Hierarchical techniques are probably best suited to biological types of data for which a hierarchical structure can safely be assumed to exist" (1974).

Indices of Association. There are many indices of measurement from which to choose to quantify the relationship between objects in the data matrix. One can measure closeness or similarity of items with an eye towards maximizing this measure within clusters and minimizing it between clusters, or one can measure distance, intending to minimize that measure within clusters. Similarity measures most often range between 0 and 1, whereas distance measures can take on any positive value.

A common index is simple euclidian distance, where d_{ij} is the distance bewteen two points, i and j,

$$d_{ij} = \{\sum_{k=1}^{r} (x_{ik} - x_{jk})^2\}^{\frac{1}{2}} \tag{1}$$

and,

x_{ik} = the value of the k^{th} variable for the i^{th} entity.

x_{jk} = the value of the k^{th} variable for the j^{th} entity.

When measurements have been taken on many variables, some weighting scheme may be necessary if very divergent units of measurement were used to measure the variables. Otherwise variables with larger units take on more importance than do variables measured on narrower scales.

The choice among measures of similarity is a theoretical one which depends largely on how the investigator conceptualizes the data. The Morisita index, which is available as an option in one of the computer packages (Smith, 1977) is computed as a similarity but then converted to a distance index. The similarity of entities i and j is

$$s_{ij} = \frac{2\{\sum^{k} (n_{ki} n_{kj})\}}{(l_i + l_j) m_i m_j} \qquad (2)$$

where,

k = number of attributes being compared

n_{ki} = the value of attribute k on entity i

m_i = the sum of all values in entity i

l_i = the probability of drawing the same attribute (assuming the values are counts) from entity i in two successive random draws without replacement. This probability is calculated as

$$l_i = \frac{\sum^{k} \{n_{ki} (n_{ki} - 1)\}}{m_i (m_i - 1)} \qquad (3)$$

Sneath and Sokal (1973) further differentiated between similarity measures of association and correlation. The difference here concerns scales of measurement. Association measures are calculated for binary or other qualitative data and are measures of agreement. Correlation measures include the Pearson product moment correlation coefficient and can be measured for any quantitative data. The correlation, r, between variables x and y is defined as

$$r_{xy} = \frac{\sum z_{xi} z_{yi}}{n}$$

z_{xi} = the standardized score for individual i on variable x.

z_{yi} = the standardized score for individual j on variable y.

Choosing an association index from among the variety which are available is not a trivial matter. "A serious difficulty in choosing a distance (function) lies in the fact that a clustering structure is more primitive than a distance function and that knowledge of clusters changes the choice of distance function" (Hartigan, 1975). This circularity is discussed by Spath (1980). "Choice of distance function is...determined by the success of the cluster algorithm." But the user must choose the function prior to using the algorithm.

Sneath and Sokal reluctantly offered the recommendation that, "of each type of coefficient considered, the simplest one should be chosen out of consideration for ease of interpretation" (1973).

Methods of Clustering

As though picking a model and an index were not sufficient headache, the investigator--who was trying only to simplify the data--must now face another decision: the method of clustering. Blashfield and Aldenderfer (1978) have surveyed the field of clustering and found two main categories of techniques which correspond to the two basic types of clustering models: hierarchical agglomerative methods and iterative partitioning methods. The hierarchical techniques are generally quicker and cheaper and more available in computer packages. They generally do not require repeated passes through the data, but suffer from being unable to reallocate items once they have been classified (Everitt, 1974).

Sneath and Sokal identified a number of aspects of clustering methods that can be combined to form various techniques. Agglomerative techniques start with two similar items and add additional items, whereas divisive techniques start with the set of all items and break down into subsets. Agglomerative techniques are more widely used.

The hierarchical techniques are used often in biology, and the following discussion of them comes largely from Sneath and Sokal (1973). Most widely used are the varieties of sequential, agglomerative, hierarchical, non-overlapping clustering methods. <u>Single-linkage clustering</u> (also known as nearest neighbor technique and minimum method) computes the similarity of an item to a cluster as the similarity between the item and its closest neighbor within the cluster. Therefore the connectedness of the cluster is based on these single links between two items. <u>Complete linkage</u> (also known as farthest neighbor and maximum method) uses the similarity of a new item to a cluster as the similarity between the item and the farthest member of the cluster. This method produces compact clusters whereas the single linkage method produces long, loose clusters, a phenomenon known as "chaining."

<u>Average linkage</u> takes an average between a new item and a cluster and can use a number of different ways of computing the average. Four averaging strategies and their various combinations are frequently discussed. <u>Arithmetic average</u> computes the total average between the item and all members of the cluster. <u>Centroid clustering</u> computes the average between the item and the center of the existing cluster. The other two possibilities are to give either equal or unequal weights to the original cluster members when computing the average.

Everitt (1974) discussed a number of iterative partitioning techniques. These techniques have three properties: a way to start clusters, a way of adding members to existing clusters, and a way to reallocate members to preferred locations. There are different strategies for choosing the number of starting clusters and which item will start them. Adding and subtracting items is done to maximize or minimize some measurement criterion over the set of items.

The <u>k-means</u> algorithm is a partitioning technique that is mentioned in a number of sources (Sneath and Sokal, 1973; Hartigan, 1975). It is now available in the BMDP computer package. It partitions some of the observations into K groups, then adds new members if they are close to the group, the group being defined by its mean. Groups are joined together if they are close or divided into two groups if the group becomes diffuse.

Blashfield and Aldenderfer (1978) pointed out in their survey of the literature that although a wide variety of partitioning techniques exist, about 75% of the studies use hierarchical techniques. They suggested three possible reasons for the popularity of the hierarchical agglomerative methods: they have been available the longest, users use what has been previously used in their literature, and more is known about these methods due to empirical analyses that have been done.

<u>Graphic Output</u>

A variety of ways of displaying the results of an analysis have been used. For hierarchical analysis, dendrograms or phenograms are frequently employed. A dendrogram is "a two-dimensional diagram illustrating the fusions or partitions which have been made at each successive level"

(Everitt, 1974). It is a tree-like sturcture with branches coming from or converging to a central point. Sneath and Sokal (1973) use the term phenogram to indicate that the relationships implied by the tree are observed (phenetic) relationships not implying biological discent.

For non-hierarchical programs, simple lists of the objects in a cluster suffice.

Software

The computations necessary to perform a cluster analysis on a data matrix of any size necessitate the use of computers and computer programs. The variety of techniques used in cluster analysis have resulted in a large number of programs being written. Some of these are available in widely used statistical packages; others have been published in books of clustering algorithms; still others are available from the authors of the programs.

Blashfield and Aldenderfer (1978) stated that 50 different clustering programs were mentioned by 53 respondents to a questionnaire sent to potential users. There appears to be no lack of algorithms available for clustering.

Types of Evaluation

Rarely does an investigator only perform a single cluster analysis. The usual case is to carry out a series of analyses (Everitt, 1974; Hartigan, 1975). It then falls to the investigator to assess each solution, to compare them to each other, and to decide their relative merits.

For instance, single linkage clustering is known to produce long chains rather than compact groups (Everitt, 1974; Sneath and Sokal, 1973). If the underlying structure consists of long chains, then this method will recover it. But if the structure is otherwise, then this method may produce a confusing solution.

If the data are artificially generated, then one knows whether the underlying strucutre has been discovered by the solution. But for data whose structure is unknown, the investigator must decide on the "correctness" by some other means. Does it confirm expectations? Does it suggest novel structure for the data? For example, Reid, Ballinger and Heather (1978) did a single hierarchical cluster analysis of 100 retarded adults and found eight clinically interesting and interpretable clusters. No statistical evaluation was done for either internal or external assessment, and still this was an interesting and revealing solution to the authors.

Another psychiatric study classified suicide attempters (Paykel and Rassaby, 1978). In this case four clustering methods were tried. Three of the methods produced only one cluster, the fourth method producing differentiable groups. This last method was used with an iterative reallocation procedure that minimized Euclidian distances from the cluster centroids to arrive at a three group solution. No statistical analyses were performed to see if the solution was a more accurate representation of the data structure than one group solutions. The three group solution was interpretable to the investigators and was accepted.

These uses of clustering are not inconsistent with Spath's (1980) advice: "Primarily, what makes an application of cluster analysis successful is the significant practical interpretation of the clusters it produces. For this reason it frequently makes sense to apply various methods, one after another and independently. Nevertheless one is sometimes happy enough simply to obtain a reasonable subdivision of the objects."

Hartigan showed a biological example in which the whales were grouped with deer and reindeer. The data used to derive this clustering were four constituents of mothers' milk: percent of water, protein, fat, and lactose. Hartigan suggested that the high fat content in the milk of these three animals may have something to do with living in cold conditions. This clustering has suggested something novel to this investigator that may not be seen by another.

A real difficulty in using cluster analysis is knowing when to accept a novel solution as "correct." The solution should be reasonable, but should also suggest something that was not obvious to the investigator before doing the analysis. The "unreasonableness" of categorizing whales with reindeer might lead one investigator to reject a solution that could be of significance to another. The dangers are that an "incorrect" solution may be accepted by an overly creative investigator while an intriguing, but peculiar, solution may be rejected by an overly dull one.

The problem of "finding" structure when none exists is not yet resolved in cluster analysis. An investigator is free to "interpret" any solution or to do repeated analyses until a solution is obtained that conforms to <u>a priori</u> expectations. At this stage clustering is strictly exploratory, and results should optimally be confirmed by other means.

An approach to this problem is to assess the solution statistically, either by external or internal criteria. Evaluation can be done on an individual solution as well as comparing several solutions to each other.

<u>External Criteria</u>

External criteria can be used both for validating a given subdiv-

ision against a known underlying structure or for comparing different solutions to each other.

Rand (1971) proposed a statistic of agreement between clusters that has been used widely in the clustering literature. It is essentially a counting method of agreement between two solutions that ranges from 0 for no similarities to 1 for identical solutions. Given two solutions, it is computed as follows:

$$\frac{a + d}{a + b + c + d} \tag{4}$$

where a,b,c,d represent the cells in a two by two table:

Soltuion 2

Solution 1	a	b
	c	d

a = the number of pairs of itmes that are clustered together in both solutions

b = the number of pairs of items that are clustered together in solution 1 but not in solution 2

c = the number of pairs of items that are clustered together in solution 2 but not in solution 1

d = the number of pairs of items that are not clustered together in either solution

Cohen's kappa (1960) has also been used in cluster analysis. It is computed as

$$k = \frac{P_o - P_c}{1 - P_c} \tag{5}$$

where,

p_o = the proportion of observed agreements between solutions,

p_c = the proportion of units for which agreement is expected by chance.

("Agreements" are items that are clustered in the same cluster.)

Milligan (In Press) has pointed out that the Rand statistic and Cohen's kappa have been found to correlate above .975.

Another external measure is the cophenetic correlation (Sneath and Sokal, 1973; Sokal and Rohlf, 1962). Cophenetic values are defined as the maximum similarity or minimum dissimilarity between any two observations in a dendrogram. A matrix of cophenetic values is derived by taking the value at the point at which each pair of items is joined. This matrix of cophenetic values is correlated with the original similarity matrix to arrive at a cophenetic correlation coefficient. Similarly it can be used as a measure of agreement between two dendrograms. The cophenetic correlation takes into account not only the number of items that did or did not go together, but also the level at which they joined.

There exists also a large number of internal measures (Sneath and and Sokal, 1973; Mazzich, 1978). Milligan (In Press) listed 30 measures, some for ordinal and nominal data, some for interval and ratio data. Examples of indices he examined are gamma, the point-biserial, correlation and Tau. Some internal measures are used in other contexts as optimality measures for clustering algorithms.

A New Clustering Program

Although there is a wide variety of choices already available to a potential user of cluster analysis, those choices generally demand some *a priori* knowledge on the part of the user. Deciding on a hierarchical medel imposes a very rigid structure on data that may result in a totally artificial solution. Partitioning, non-hierarchical techniques are less

restrictive, but can produce poor solutions because of poor starting points or initial partitions. Recently, an alternative program has been developed that is non-hierarchical, does not use a subset of items as starting points, and allows overlapping sets as a solution (McCormick, Cliff, Reynolds, Cudeck, and Zatkin, 1980). A solution to another problem--that of a measure of association--has also been offered.

Association measure. Cliff (1979) has proposed a quality index, q, for binary data that is a measure of consistency between items or between persons. Quality indices relate an observed value to a best case of perfect consistency and a worst case of total inconsistency. Perfect consistency is defined as Guttman scale data. Inconsistency, or the worst case, is defined as independence between items given fixed observed marginals.

$$q = \frac{t - t_w}{t_b - t_w} \quad (6)$$

t = an observed value of a statistic

t_w = the worst-case value of t

t_b = the best-case value of t

The statistic, t, proposed by Cliff, is a weighted sum of dominance relations measured for subjects on dichotomous variables. Dominance relations fall into three categories: redundant, contradictory, and unique relationships.

If a group of persons are scored on a single question from an achievement test, those that get the item correct "dominate" those that do not. The persons can be ordered by their score on this item. For a second test item the persons can again be ordered. If the two items order the persons in the same way, there is redundant information provided. If they order the people in the opposite way, the two items have furnished contradictory information. A third possibility is that an item provides not redundant nor contradictory information, but a new ordering scheme altogether. This information would be unique, and each item can offer a number of unique relationships relative to other items. A weighted combination of these three kinds of information can be calculated to measure consistency of a set of such items. That measure can then be related to perfect consistency or total independence by q, the quality index. High values of q indicate consistency and low values indicate independence.

The three types of relations can be calculated from a fourfold table like that mentioned previously, where entries refer to passing (endorsing) or failing (rejecting) two items. The entry a is the number who pass both, d is the number who fail both, and b and c are the numbers passing one and failing the other.

The number of redundant relations between the two items is ad; the number of contradictory is bc. The number unique to one item is ab + cd; the number unique to the other is ac + bd. The number of each kind of relation can be summed across all pairs of items to give a total number of redundant (r..), contradictory (c..) and unique (u..) relations for a set of items or other dichotomous measures.

Pearson r is sensitive to marginal frequencies in binary data, in that it will give low values when the marginals are very different even for items measuring the same property but at different difficulty levels. The index q is an attempt to overcome this shortcoming of r. Pearson r can be cast in terms of redundant, unique, and contradictory relations.

$$\text{Pearson } r = \frac{r_{jk} - c_{jk}}{(r_{jk} + c_{jk} + u_{jk})\ (r_{jk} + c_{jk} + u_{jk})} \qquad (7)$$

r_{jk} = relations redundant to items j and k

c_{jk} = relations contradictory in items j and k

u_{jk} = relations unique to item j and not found in k

u_{kj} = relations unique to item k and not found in j

Other measure of association can also be seen in terms of these relations. The Goodman-Kruskal gamma leaves out unique relations altogether.

$$\text{Goodman-Kruskal gamma} = \frac{r.. - c..}{r.. + c..} \qquad (8)$$

The Kuder-Richardson 20 formula can also be defined with these measures. KR20 has unique relations in the denominator only.

$$\text{KR20} = \frac{x(r..-c..)}{xr.. - (x-2)c.. + u..} \qquad (9)$$

where,

x = the number of items

r.. = the total number of redundant relations in the entire set

c.. = the total number of contradictory relations

u.. = the total number of unique relations

The advantage of q, then, is in its use of unique relations to help order items or persons better. Including unique relations in the denominator of an index gives these special pieces of information a negative weight; ignoring them, of course, gives them zero weight. q, as shown above, is a function of t. There are many possible ways of defining t. In this present research, the relation

$$t = r.. - c.. + (.25)u..$$

has been used. Preliminary experimentation and Monte Carlo studies led to this choice.

Clustering Algorithm. A clustering program has been written which uses q as the index for finding items that cluster together. The program is described in McCormick, et al (1980). The model allows for overlapping sets of items. It can recover a hierarchical arrangement, but does not impose that structure. The method used is to start each item as a cluster. Then a second item is joined to a cluster which has the highest pairwise q with the original item. From that point on, an item is added if it has the highest average q value with the cluster. Average is computed as the arithmetic average, similar to the average linkage clustering as described in Sneath and Sokal (1973). That is, q values are taken between each item in the cluster and a new item, then those q's are averaged. The item that is added to the cluster is the item that gives the highest q.

The program can compute other indices of association as well as q. Some other indices which have also been used are KR20, gamma, and Pearson r.

Stopping Rule. If the clusters are well differentiated and compact, then all items in the cluster will add on the other items from the cluster and be replications of each other. The program reports the histories for clusters only until they duplicate another cluster, then continues reporting only one of the replications. Items continue to be joined to clusters until all items from the item pool have been added.

When the first item from outside the cluster is added, the index value for the cluster should drop dramatically, if the cluster is well differentiated from outliers. This drop, if it occurs, can be used as an indication of the end of a cluster when one does not know the true cluster members. Other strategies can also be used for discovering the stopping point of clusters. One strategy that has been successful is to inspect the index values for clusters as items are added and pick a value that seems to indicate a large gap and to use that value for all the clusters. Although statistical methods for finding "large" gaps have been proposed (Wainer & Schacht, 1978) we have found the rule of using the absolute largest gap to be frequently effective.

Previous Evaluations. McCormick, et al. (1980) have reported a series of analyses which examined the performance of this mehtod using artificial data. In this study a number of different data sets were generated in which the sturcture of the data and the cluster memberships were known. Then the recovered clusters were compared to the known clusters, and the percentage of correctly classified objects was used as a measure of accuracy. Altogether there were 14 distinct conditions examined, and in each there were 24 objects to be classified.

Across all conditions the percentage of correctly classified objects was .84. This suggested that the method was successful at recovering data under artificial conditions. However, to study the utility of the method further, an evaluation with real data is required.

Method

Data. The data used consisted of adjective ratings by social workers of a group of mothers. The children of the mothers are from a birth cohort of all live births at Rigshospitalet in Cohpenhagen from September 1, 1959, to December 31, 1961 (Mednick, Mura, Schulsinger & Mednick, 1971). A great deal of data have been collected at various times on these subjects. Only data pertinent to the present study will be discussed here. Psychiatric histories for the parents of all the children in the cohort were collected. Seventy-two children with a parent who had a psychiatric admission of schizophrenia were found. Seventy-two additional children were matched to these who had a parent with a psychiatric admission of either character disorder or psychopath. One hundred-twenty one controls were found who had parents with no history of psychiatric admissions. The parents were interviewed when the children were between the ages of 11-13. The interviews were conducted by one of three social workers.

As part of this interview, the social workers scored the mothers on a check list of 154 descriptive words and phrases. An attempt was made to interview both parents of each child, and nearly all the mothers were contacted and interviewed. Much less success was achieved interviewing the fathers. Due to the incompleteness of information on the fathers, this study was confined to the ratings of the mothers only. It should be noted, however, that although the children are categorized as having a parent in one of the two disorder groups or in the control group, it is not necessarily the mother who was so classified.

The intention of this study was to cluster the adjectives that had been used as descriptors of the mothers. The first question was whether the clustering program could provide meaningful and interesting clusters when used on real data. The second question was how the solution from the new program would compare to solutions from other available clustering routines.

Item Pool. The group of 154 adjectives was considered too large to partition into interpretable clusters for purposes of a methodological study such as this. It was necessary to reduce the number by some means. Furthermore, some of the items had very low frequencies, and therefore would not cluster with other items. In order to reduce the quantity of items, the items were sorted and grouped intuitively by two independent judges. Those items that did not easily group with others were deleted. No effort was made to keep these groupings equal in size. The only criterion used was that the words formed cohesive groups on a rational basis. This first attempt to group the adjectives reduced the list to approximately 60 items. Next a check was made to determine how cohesive the items in each set were. This was measured by the q statistic. Items whose pairwise q with other items in their group was less than .2 were deleted. This step eliminated one entire group and reduced others by one or two items; three of the groups remained intact. The rest of the analyses were performed on the remaining 51 items. The final item pool is shown in Table 1.

Insert Table 1 about here

Procedure. The APL routine to perform non-hierarchical cluster analysis was essentially the same as that reported in McCormick, et al (1980), except for an additional section. In order to better summarize the final results, a "second order" analysis (Cliff, et al, Note 1) was performed on the solution from the original clusters. This occurred in the following manner. First, a subjective assessment of the point at which the true cluster members ended was made for each of the 54 clusters. Then a binary membership matrix was created which was of order 54 by 54, where a 1 denoted that the item on the row was part of the cluster begun by the item of the column. This binary membership matrix was subsequently analyzed by the same method which produced the first order solution. However, the second order clusters which were formed denote those clusters which had substantially the same items as members from the first order analysis. It will be seen that this additional analysis merely served as a way to clarify the form of the clusters derived in the first order procedure.

In order to compare the solution obtained with this program to other possible solutions, a series of other clustering routines were also used. A list of these programs are summarized in Table 2.

Insert Table 2 about here

In addition to the q index, the APL program also was used with Pearson r as a measure of association. With r, it was possible to use the largest gap cut-off rule, although a second-order clustering was again used to obtain an interpretable solution. NT-SYS (Rohlf, Kishpaugh, and Kirk, 1974) was used to represent a hierarchical program using correlations as the similarity index and complete linkage as the clustering procedure. BMDP1M (Dixon and Brown, 1979) was used for another hierarchical clustering. Here, too, correlations were used, but the method was group-average linkage. The hierarchical procedure in the EAP (Smith, 1977) package (PROC DENDRO) was also used. Group-average linkage was the procedure, and the Morisita index was used for association. A non-hierarchical solution was obtained from the EAP (Smith, 1977) reallocation procedure (PROC GRSIM), this procedure using the similarity matrix from the previous procedure and group-average similarities.

In addition to the cluster analyses, a factor analysis was performed using the SAS (Barr, Goodnight, and Sall, 1979) package and varimax rotation. The first analysis produced fourteen eigenvalues greater than unity. Square roots of the eigenvalues were plotted, showing a drop after the eighth. A second analysis limited the program to eight factors.

In addition to the solutions that were interpretable several others solutions were discarded for a variety of reasons. A procedure in the NT-SYS (Rohlf, Kishpaugh, and Kirk, 1979) package called "Subsets" produced a non-hierarchical solution of the most distinct subsets. The subsets obtained contained at most three members and this procedure failed to include 26 of the items. A second non-hierarchical program in the same package iterated many times but failed to find a set of definable clusters. The EAP (Smith, 1977) program for hierarchical clustering was attempted with other distance measures including the Euclidian distance and the Manhattan metric, but these produced confusing groups. The BMDP (Dixon and Brown, 1979) program using absolute value of the correlation as the index produced clusters of antonyms; this solution, although interesting, was so different from the other solutions, it was felt it was not comparable.

To summarize, six different clustering programs were examined. Five additional routines were also tested but the solutions of these were judged

to be unusable for a variety of reasons. Of the six programs which were compared, two used the clustering method described in McCormick, et al (1980), but one utilized the q index while the second used Pearson r. In addition to the six clustering routines, a standard principal component analysis was also undertaken.

Results

We will discuss the results in several stages. First, we will describe the nature of the results from the primary non-hierarchical clustering program, using q as the measure of association. Second, we will outline how the second-order clustering occurs and discuss the second-order analysis results. Then, we will compare the second-order cluster solution using q as the association index to that which is obtained using Pearson r as the index. Finally, we will compare these two solutions to those which were obtained from the other traditional clustering results.

Primary Non-hierarchical Clustering

Table 3 containes a sample of the results for the primary non-hierarchical clustering using the q index. As discussed previously, each object

Insert Table 3 about here

in the set to be clustered begins its own cluster, so that with k objects, the method produces k clusters. Table 3 shows the cluster histories for the adjectives Inhibited, Shy and Has Initiative, which are the 1st, 2nd, and 10th items.

Consider the history of the cluster begun by item 1, Inhibited. On the first cycle, item 6, Insecure, was added and the modified within-cluster q was .71. In the second cycle, item 40, nervous, was selected, which slightly increased the within-cluster q to .73. After that point each successive item added to the cluster gradually decreased the value of the within-cluster q, until after the 20th addition, its value was .02. Clearly, the first items added to this cluster are most similar to each other, while those added later are less similar. A subjective decision was made to

include only the first 5 items in this cluster. The index at that point was .63.

Next, the second cluster may be examined. It was begun by the adjective Shy, then proceeded by adding Nervous, Insecure and Timid, which produced within-cluster q's of .54, .72 and .67, respectively. On the eighth step, both clusters 1 and 2 had exactly the same members. Thereafter, the two clusters necessarily added identical items and had identical within-cluster q's. The behavior of q for the second cluster indicated that Shy was probably an outlier in the set of objects. The first item added produced a relatively low q value. The second item raised the within-cluster average. Thus Shy was probably located farther away from Nervous and Insecure than either of these latter items were from each other. Thus we closed this cluster with just a single member, the original item Shy.

Cluster 10 produced a set of items which was not similar to either the first or second clusters. At the Fourth cycle, it contained the items Has Initiative, Active, Energetic, and Outgoing, at which point the within-cluster average was .65.

In the same way, all 51 items began separate cluster, and a subjective decision was made as to the appropriate place to close the clusters. Generally, this decision was governed by the point at which the largest drop in the within-cluster average occurred. However, the rule is not infallable, and inevitably some clinical judgment must be exercised. The complete results of the primary clustering reduced the orininal matrix of ratings for the 265 parents on the 54 adjectives to a square membership matrix of order 54 which summarized the analysis. The membership matrix contains a 1 whenever the item of the row was a member of the cluster of the column. Tables 4 and 5 give the primary membership matrices for the q and r results.

Insert Tables 4 and 5 about here

Second-Order Analysis

Although the membership matrix (Table 4) contains all the information about cluster composition, in practice it is difficult to interpret the results in this form. One solution is to perform a second analysis, using the same method as outlined above, with the membership matrix as data. This strategy is intended to group together those clusters which are composed of the same items. Clusters which contain most of the same items can be considered secondary "super-clusters." For example, if clusters 1, 2, 7, 15 and 23 all contain nearly the same items, then these clusters would form a super-cluster. A summary matrix can then be constructed which contains items as rows and super-clusters as columns. An entry in this matrix denotes the percentage of times the row item occurred in the particular super-cluster. For example, if the adjective Inhibited was found in 4 of the 5 clusters above, then its entry would be .80.

The results from this analysis are displayed in Table 6. The matrix

---

Insert Table 6 about here

---

has been rearranged so that similar items are placed together as much as possible. Of the eight super-clusters produced, the first two are themselves highly similar. They differ primarily only with regard to items which occured in one or two clusters. These groupings contain positive attributes. The third grouping also contains positive characteristics, which appear to be mainly of a social nature. Super-clusters 4, 5, and 6 are also very similar, containing the negative traits Nervous, Insecure, Anxious, Timid and Inhibited. The seventh group has four items denoting a boisterous type. Finally, Stubborn and Headstrong composed the subset of items in the last group. About half the items were not classified with a large amount of frequency in any of the second-order clusters, although only the item Alert was unplaced.

Second-Order Clusters with Pearson r Association Index

Using Pearson r rather than q produced the super-clusters shown in Table 7. This approach gave a somewhat different perspective. The first

---

Insert Table 7 about here

---

two groupings contained almost all the positive and negative items, respectively. The third column isolated four boisterous characteristics, although not all were the same as in the previous solution using the q index. The items Headstrong and Stubborn re-emerged as a separate group in column 4. Clusters 5 through 8 were sub-groups of the positive characteristics. Finally, Withdrawn and Silent were also clustered together.

Table 8 contains the number of objects which were classified by each cluster for the solution using q as index or using Pearson r. As can be seen, the large cluster of positive and negative items in the solution using

---

Insert Table 8 about here

Pearson r almost entirely overlapped with the q clusters 1 through 6. However, many of the other clusters in the Pearson r solution had their counterparts in the q solution also. The value of the Rand (1971) measure of correspondence between solutions was .79. This suggested that there is a great deal of overlap between the two approaches, but no exact correspondence.

Traditional Clustering Methods

Results from the analysis of these data using the four traditional clustering methods plus the factor analysis will not be presented in detail. Rather we will attempt to highlight these solutions as they relate to the two above analyses, which were the main points of focus here.

For hierarchical soltuions, the investigator must decide the correct number of clusters. Everitt (1974) suggested that "an examination of the dendrogram for large changes between fusions would be useful." This suggestion was followed. It was not felt necessary to cut across the entire dendrogram to get clusters that had the identical degree of density. In some places, the clusters were taken at lower levels; in other cases, more diffuse groups were considered a cluster. Thus the decisions on where to cut the dendrogram were essentially subjective.

Each of these different methods produced solutions with 6 clusters. In the non-hierarchical EQP program, GRSIM, 10 clusters were found. It should be noted that this finding of 10 clusters is similar to the number of clusters in the two APL solutions discussed above. It could be that the higher number of clusters required by all these non-hierarchical approaches is a methodological feature of this kind of approach.

In the factor analysis solution, seven factors were retained for final study. Adjectives were assumed to belong together if they had loadings larger than .42. This value was judged to give the clearest solution in the seven factor results.

In Tables 9 and 10 the number of adjectives that were clustered in corresponding clusters of either the APL-q or APL-r solutions with the five

Insert Tables 9 and 10 about here

traditional methods are shown. The greatest number of matching objects was for the large clusters of positive and negative items. In terms of absolute size, the factor analysis seems to have the least overlap with APL-q and APL-r. This again may well be due to a basic methodological difference between the two models.

To compare the solutions statistically, the Rand (1971) statistic (Equation 4) was computed between all pairs of solutions. The results are shown in Table 11. Rand (1971) does not give a table of significance levels. Cohen

Insert Table

(1960), discussing kappa states, "it is generally of as little value to test kappa for significance as it is for any other reliability coefficient--to know merely that kappa is beyond chance is trivial since one usually expects much more than this in the way of reliability in psychological measurement." However, the conclusion can be made that many aspects of all the solutions have elements in common.

Discussion

Most of the clusters in most of the solutions had a certain degree of plausibility, but the second-order q analysis seems particularly reasonable. Forming an intuitive "tertiary" analysis of Table 6, suggests strongly some very interpretable clusters. There is the rather large group represented by super-clusters 1 and 2. This includes the ten socially desirable adjectives. Cluster 3 is a subset of these: kind, cooperative, and friendly, plus five others, charming, quick and outgoing, lively, and cheerful, a very nice interpersonal attractiveness cluster. Clusters 4, 5, and 6 are virtually identical, defining a neurotic cluster. Cluster 7 is quite a clear extroversion cluster, and cluster 8 is a stubborn-headstrong cluster. All of these are highly interpretable, and their partially overlapping nature adds to their interpretability, reflecting, apparently, the structure of the language.

The primary disappointment of this analysis is its failing to include many of the adjectives in any of the superclusters. The possible explanation is that, for the most part, the adjectives constitute a diffuse structure. The unincluded are simply those toward the periphery of this structure.

It is not totally unexpected that all the Rand statistics between the various solutions were fairly high. Milligan (In Press) has discussed this situation: "A particularly troublesome issue in clustering is the discrimination between two fairly similar solutions. For example, a researcher may need assistance in determining how many clusters are present in a data set. Such a situation occurs with the use of a hierarchical clustering algorithm where an applied researcher usually wants to select a specific partitioning level as the final solution. Procedures for determining which hierarchy level is the best representation of the data have not been well developed. This is partially due to the fact that the discrimination between two similar solutions can place an extreme performance demand on any recovery measure."

Examining the Table 11 shows a number of interesting findings. The APL program gave different results depending on the measure of association. The Rand Statistic between APL-q and APL-r was .788. Two hierarchical programs, which used average linkage but different indices (BMDP1M and EAP-DENDRO) gave a Rand statistics of .93. One of the highest values obtained was between the two EAP programs (.947); they used the same index and somewhat similar methods although one was hierarchical and the other not.

Thus there seems to be a main effect for the type of index used, which reinforces the notion that assuming the typical "default option" of Pearson r should be critically considered before a cluster analysis is performed.

The study has shown that the clustering program and the q statistic do produce an interesting and different solution. One cluster was not found elsewhere--the one containing Outgoing, Lively, Blustery, and Loud. These traits can be seen as occurring as points along a continuum of activity. The continuum can be seen as an ordering analogous to item difficulties on a test. In testing, a person who gets a positive score on a more difficult item will also have positive scores on the less difficult items. The distribution of scored for items at different difficulty levels all measuring the same trait will be positively skewed. This type of distribution is difficult for many measures of association to recover when one is trying to find the group of items that measure the same trait regardless of difficulty level. As can be seen from this study, the other measures of association failed to group these items together while q was successful.

The solution obtained is different from others and adds to the field of cluster analysis another option not previously available, expecially for data that consists of items distributed in a non-normal manner. Data exist that are more complelling than adjective check lists of Danish mothers. One type of data of great interest that has a non-normal distribution is crime. Most people commit few or no crimes with a few people commiting the majority, although frequencies of crimes may still be small. Q may prove to be superior to other indices of association for the investigator studying data distributed like crime.

Reference Note

1. Cliff, N., Cudeck, R., McCormick, D. J., Collins, L. Clustering binary data using consisterey indices: First and second order analysis. In preparation.

References

Barge, Sylvia J., And Gregory A. Marks Osiris III, University of Michigan, 1973.

Barr, Anthony J., Goodnight, James H., and Sall, John P. Statistical Analysis Systems, SAS Insititue, Inc., 1979.

Blashfield, Roger The equivalence of three statistical packages for performing hierarchical cluster analysis. Psychometrika, Vol. 42, No. 3, September 1977, 429-431.

Blashfield, Roger K., and Aldenderfer, Mark S. The literature on cluster analysis. Multivariate Behavioral Research, 1978, 13, 271-295.

Cliff, Norman. Test theory without true scores? Psychometrika, 1979, 44, 373-393.

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

Curry, David J. Some statistical considerations in clustering with binary data. Multivariate Behavior Research, Vol. 11, No. 2, April-1970, 175-188.

Dixon, W. J., and Brown, M. B., Editors Biomedical Computer Programs, P-Series. University of California Press, Berkeley, 1979.

Edelbrook, Craig Mixture model tests of hierarchical clustering algorithms: The problem of classifying everybody. Multivariate Behavior Research, 1979, 14, 367-384.

Everitt, Brian Cluster Analysis, Heinemann Educational Books Ltd., London, 1974.

Gnanadesikian, R. Methods for Statistical Data Analyses of Multivariate Observations. John Wiley & Sons, New York, 1977.

Gruvaeus, Gunnar, and Wainer, Howard Two additions to hierarchical cluster analysis. British Journal of Mathematical Statistical Psychology. 1972, 25, 200-206.

Hartigan, John A. Clustering Algorithms. John Wiley and Sons, New York, 1975.

Hays, William L. Statistics for the Social Sciences, 2nd ed. Holt, Rinehart, and Winston, Inc., New York, 1973.

Johnson, Stephen C. Hierarchical clustering schemes. Psychometrika, Vol. 32, No. 3, September 1967, 241-254.

McCormick, Douglas, Cliff, Norman, Reynolds, Thomas, Cudeck, Robert, and Zatkin, Judith Isolating Homogeneous Subscales of Dichotomous Items, presented at meetings of Psychometric Society, May 1980.

Mazzich, Juan E. Evaluating clustering methods for Psychiatric Diagnosis? Biological Psychiatry, Vol. 13, No. 2, 1978.

Mednick, S. A., Mura, E., Schulsinger, F. & Mednick, B. Perinatal conditions and infant development in children with schizophrenic parents. Social Biology, 1971, 18, 5103-5113.

Milligan, Glenn W. A Monte Carlo study of thirty internal criterion measures for cluster analysis. Psychometrika, 1981, 46, In Press.

Paykel, E. S. and Rassaby, E. Classification of suicide attempters by cluster analysis. British Journal of Psychiatry, 1978, 133, 45-52.

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

Reid, A. J., Ballinger, B. R. and Heather, B. B. Behavioral syndromes identified by cluster analysis in a sample of 100 severely and profoundly retarded adults. Psychological Medicine, 1978, 8, 399-412.

Rohlf, F. James, Kishpaugh, John, and Kirk, David Numerical Taxonomy System of Multivariate Statistical Programs, 1974.

Smith, Robert Ecological Analysis Package, 1977.

Sneath, Peter A., And Sokal Robert R. Numerical Taxonomy: The Principles and Practice of Numerical Classification. Wilt Freeman and Company, San Francisco, 1973.

Sokal, Robert R., and Rohlf, F. James The comparison of dendrograms by objective methods. Iaxon Vol XI, No. 2, February 1962, 33-39.

Spath, Helmuth Cluster Analysis Algorithms for Data Reduction and Classification of Objects. Translated by Ursula Bull, Ellis Horwood Limited, England, 1980.

VanEmden, M. H. An Analysis of Complexity. Mathematisch Centrum Amsterdam, 1971.

Wainer, H. & Schacht, S. Gapping, Psychometrika, 1978, 43, 203-212.

Table 1

Final item pool

Cooperative
Friendly
Kind
Responsible
Conscientious
Reliable
Independent
Understanding
Trusting
Helpful
Adaptable
Honest
Sincere
Loyal
Mature
Versatile
Warmhearted
Active
Energetic
Quick
Charming
Outgoing
Lively
Talkative
Industrious
Alert
Headstrong
Cheerful
Softhearted
Initiative
Insecure
Nervous
Timid
Anxious
Worrying
Inhibited
Shy
Withdrawn
Silent
Gloomy
Yielding
Passive
Tense
Sensitive
Preoccupied
Blustery
Loud
Quarrelsome
Aggressive
Stubborn
Restless

Table 2

Clustering Programs

Program	Model	Method	Index
APL-q	non-hierarchical	group-average linkage	q
APL-r	non-hierarchical	group-average linkage	r
NT-SYS	hierarchical	complete linkage	r
BMDP1M	hierarchical	group-average linkage	r
EAP-DENDRO	hierarchical	group-average linkage	Morisita
EAP-GRSIM	hierarchical	group-average linkage	Morisita

Table 3

Sample Results from APL-q
for 3 Clusters

Cluster 1			Cluster 2			Cluster 3		
Inhibited			Shy			Initiative		
Item		Index	Item		Index	Item		Index
6	Insecure	.71	40	Nervous	.54	13	Active	.80
40	Nervous	.73	6	Insecure	.72	14	Energetic	.73
4	Timid	.68	4	Timid	.64	30	Outgoing	.65
42	Anxious	.63	42	Anxious	.63	15	Quick	.58
7	Gloomy	.53	1	Inhibited	.56	33	Charming	.61
2	Shy	.49	7	Gloomy	.51	35	Kind	.55
3	Withdrawn	.47	3	Withdrawn	.47	24	Cooperative	.56
5	Silent	.44	5	Silent	.44	28	Friendly	.53
43	Worrying	.41	43	Worrying	.41	16	Independent	.52
9	Yielding	.38	9	Yielding	.38	26	Versatile	.53
44	Tense	.34	44	Tense	.34	36	Understanding	.50
8	Passive	.26	8	Passive	.26	23	Mature	.51
41	Preoccupied	.23	42	Anxious	.23	34	Warmhearted	.50
39	Sensitive	.20	39	Sensitive	.20	22	Conscientious	.49
45	Restless	.11	45	Restless	.11	27	Adaptible	.49
50	Aggressive	.11	50	Aggressive	.11	18	Responsible	.49
47	Quarrelsome	.08	47	Quarrelsome	.08	17	Reliable	.47
38	Softhearted	.04	38	Softhearted	.04	25	Helpful	.39
48	Stubborn	.02	48	Stubborn	.02	19	Loyal	.40

Table 4

Primary Membership Matrix Using q

Nervous	00000000000111111111100000000000000000000000000000
Insecure	00000000000011111111100000000000000000000000000000
Anxious	00000000000001111111100000000000000000000000000000
Timid	00000000000001111111100000000000000000000000000000
Inhibited	00000000000000000011111100000000000000000000000000
Loud	00000000111000000000000000000000000000000000000000
Blustery	00000000111000000000000000000000000000000000000000
Talkative	00000000111000000000000000000000000000000000000000
Cheerful	11110000000000000000000000000000000000000000000000
Lively	11111000000000000000000000000000000000000000000000
Outgoing	11111111111000000000000000000000000000000000000000
Charming	11111111000000000000000000000000000000000000000000
Quick	11111111000000000000000000000000000000000000000000
Initiative	00001111000000000000000000000001000000000000000000
Energetic	00000111000000000000000000000000000000000000000000
Active	00000111000000000000000000000000000000000000000000
Restless	00000000000000001100000000000000000000000000000000
Sensitive	00000000000000000000000000000000000001100000000000
Stubborn	00000000000000000000000001100000000000000000000000
Headstrong	00000000000000000000000001100000000000000000000000
Quarrelsome	00000000000101000000000000000000000000000000000000
Aggressive	00000000000100000000000000000000000000000000000000
Preoccupied	00000000000010000000000000000000000000000000000000
Tense	00000000000000100000000000000000000000000000000000
Passive	00000000000000000000000000010000000000000000000000
Yielding	00000000000000000000000000001000000000000000000000
Silent	00000000000000000000000110000000000000000000000000
Withdrawn	00000000000000000000001110000000000000000000000000
Gloomy	00000000000000000000001000000000000000000000000000
Shy	00000000000000000000000000000100000000000000000000
Worrying	00000000000000010000000000000000000000000000000000
Industrious	00000000000000000000000000000010000000000000000000
Sincere	00000000000000000000000000000000000001100000000000
Honest	00000000000000000000000000000000000010000000000000
Loyal	00000000000000000000000000000000100000011000000000
Trusting	00000000000000000000000000000000100000010000000000
Helpful	00000000000000000000000000000000100000000000000000
Alert	00000000000000000000000000000000000000000000000000
Softhearted	00000000000000000000000000000000001000000000000000
Warmhearted	00000000000000000000000000000000011000000000000000
Independent	00000000000000000000000000000000010000001111111111
Adaptable	00000000000000000000000000000000011001111111111111
Conscientious	00000000000000000000000000000000111111111111111111
Versatile	00000000000000000000000000000000011111111111111111
Mature	00000000000000000000000000000000011111111111111111
Understanding	00000000000000000000000000000000011111111111111111
Reliable	00000000000000000000000000000000011111111111111111

Table 4 (Continued)

Responsible	00000000000000000000000000000000111111111111111111
Friendly	11111000000000000000000000000000111111111111111111
Kind	11111000000000000000000000000001111111111111111111
Cooperative	11111000000000000000000000000001111111111111111111

Table 5

Primary Membership Matrix Using Pearson r

Conscientious	111111111111111100000000000011100000000000000000000000
Responsible	111111111111111100000000000011100000000000000000000000
Reliable	111111111111111100000000000110000000000000000000000000
Understanding	111111111111111110000000000000000000000000000000000000
Versatile	111111111111111110000000000000000000000000000000000000
Mature	111111111111111110000000000000000000000000000000000000
Independent	111111111111111110000000000000000000000000000000000000
Adaptable	111111111111111100000000000000000000000000000000000000
Kind	111111111000000001110000000000000000000000000000000000
Friendly	111111111000000001110000000000000000000000000000000000
Cooperative	111111111000000001110000000000000000000000000000000000
Honest	111111111100000000000000000100000000000000000000000000
Sincere	111111111000000000000000000100000000000000000000000000
Warmhearted	111111111000000010000000000000000000000000000000000000
Loyal	111111111010000000000000000000000000000000000000000000
Helpful	111111111000000000000000000000000000000000000000000000
Trusting	111111111000000000000000000000000000000000000000000000
Quick	111111111000000000000000000000000000000000000000000000
Energetic	111111100000000000000000000000011110000000000000000000
Active	111111100000000000000000000000011110000000000000000000
Outgoing	111111100000000000000000000000010000000000000000000000
Initiative	111111100000000000000000000000001000000000000000000000
Charming	111111100000000000000000000000000000000000000000000000
Lively	111111100000000000000000000000000000000000000000000000
Headstrong	100000000000000000000110000000000000000000000000000000
Stubborn	000000000000000000000110000000000000000000000000000000
Loud	000000000000000000000001111000000000000000000000000000
Blustery	000000000000000000000001111000000000000000000000000000
Aggressive	000000000000000000000001100000000000000000000000000000
Quarrelsome	000000000000000000000001100000000000000000000000000000
Restless	000000000000000000000000000000000000000100000000000000
Softhearted	000000000000000000000000000000000000000000100000000000
Cheerful	010000000000000000000000000000000000000000000000000000
Talkative	000000000000000000000000000000000100000000000000000000
Alert	100000000000000000000000000000000000000000000000000000
Industrious	001000000000000000000000000000000000000000000000000000
Silent	000000000000000000000000000000000000011000000111111111
Withdrawn	000000000000000000000000000000000000011000000111111111
Preoccupied	000000000000000000000000000000000000000010000111111111
Insecure	000000000000000000000000000000000000000000011111111111
Nervous	000000000000000000000000000000000000000000010111111111
Sensitive	000000000000000000000000000000000000000001000111111111
Worrying	000000000000000000000000000000000000000000000111111111
Yielding	000000000000000000000000000000000000000000000111111111
Passive	000000000000000000000000000000000000000000000111111111
Gloomy	000000000000000000000000000000000000000000000111111111

Table 5 (Continued)

Shy	000000000000000000000000000000000000000000000111111111
Tense	000000000000000000000000000000000000000000000111111111
Anxious	000000000000000000000000000000000000000000000111111111
Timid	000000000000000000000000000000000000000000000111111111
Inhibited	000000000000000000000000000000000000000000000111111111

Table 6

Second-order Clusters using q

		Super-clusters							
	Adjectives	1	2	3	4	5	6	7	8
1	Kind	1.00	1.00	1.00					
2	Cooperative	1.00	1.00	1.00					
3	Conscientious	1.00	1.00						
4	Understanding	1.00	.95						
5	Friendly	1.00	.95	.63					
6	Versatile	1.00	.95						
7	Mature	1.00	.95						
8	Reliable	1.00	.95						
9	Responsible	1.00	.95						
10	Adaptive	.85	.86						
11	Independent	.60	.57						
12	Charming			1.00					
13	Quick			1.00					
14	Outgoing			1.00				1.00	
15	Lively			.63					
16	Cheerful			.50					
17	Nervous				1.00	.94	1.00		
18	Insecure				1.00	.94	.94		
19	Anxious				.93	.88	.88		
20	Timid				.93	.88	.88		
21	Inhibited				.53	.50	.50		
22	Loud							1.00	
23	Blustery							1.00	
24	Talkative							1.00	
25	Stubborn								1.00
26	Headstrong								1.00
27	Shy				.07	.06	.06		
28	Withdrawn				.02	.25	.19		
29	Silent				.02	.25	.19		
30	Gloomy				.02	.19	.19		
31	Passive					.06			
32	Yielding				.07	.06	.06		
33	Initiative		.05	.05					
34	Industrious	.05	.05						
35	Alert								
36	Active			.38					
37	Energetic			.38					
38	Loyal	.01	.14						
39	Trusting	.05	.01						
40	Honest	.05	.05						
41	Helpful		.05						
42	Warmhearted	.15	.14						
43	Sincere	.01	.01						
44	Softhearted	.01	.01						
45	Sensitive	.01	.01						
46	Preoccupied				.07	.06	.06		

Table 6 (cont.)

	Adjective	1	2	3	4	5	6	7	8
47	Worrying				.07	.06	.06		
48	Tense				.07	.06	.06		
49	Restless				.13	.13	.13		
50	Quarrels				.07	.06	.13		
51	Aggressive						.06		

Table 7

Second-order Clusters using r.

	Adjectives	Super-clusters 1	2	3	4	5	6	7	8	9
1	Understanding	1.00				1.00				
2	Versatile	1.00				1.00				
3	Adaptible	1.00				1.00				
4	Mature	1.00				1.00				
5	Independent	1.00				1.00				
6	Reliable	1.00				1.00	.33			
7	Responsible	1.00				1.00	1.00			
8	Conscientious	1.00				1.00	1.00			
9	Loyal	1.00				.14				
10	Honest	1.00				.14				
11	Cooperative	1.00						1.00		
12	Friendly	1.00						1.00		
13	Kind	1.00						1.00		
14	Sincere	1.00								
15	Warmhearted	1.00								
16	Helpful	1.00								
17	Trusting	1.00								
18	Quick	1.00								
19	Active	.78							1.00	
20	Energetic	.78							1.00	
21	Initiative	.78							.25	
22	Outgoing	.78							.25	
23	Lively	.78								
24	Charming	.78								
25	Withdrawn		1.00							1.00
26	Silent		1.00							1.00
27	Inhibited		1.00							
28	Shy		1.00							
29	Timid		1.00							
30	Insecure		1.00							
31	Gloomy		1.00							
32	Passive		1.00							
33	Yielding		1.00							
34	Sensitive		1.00							
35	Nervous		1.00							
36	Preoccupied		1.00							
37	Anxious		1.00							
38	Worrying		1.00							
39	Tense		1.00							
40	Blustery			1.00						
41	Loud			1.00						
42	Quarrels			.50						
43	Aggressive			.50						
44	Stubborn				1.00					
45	Headstrong	.11			1.00					

Table 7 (cont.)

	Adjective	1	2	3	4	5	6	7	8	9
46	Industrious	.11								
47	Alert	.11								
48	Talkative								.25	
49	Cheerful	.11								
50	Softhearted									
51	Restless									

Table 8

Number of Adjectives which were Classified by Clusters Using Pearson r or q

Clusters using q	Clusters using Pearson r 1	2	3	4	5	6	7	8	9
1	17	1			8	3	3	2	
2	16	1			7	2	1	3	
3	11						3	4	
4		13	1						2
5		14	1						2
6		13	2						2
7	1		2					2	
8	1			2					

Table 9

Number of Adjectives Classified by APL-q with Five other Solutions

Clusters from		Clusters Using APL Program With q Index 1	2	3	4	5	6	7	8
NT-SYS	1	13	10	3					
	2				13	14	13		
	3				2	2	3	3	2
	4	4	6	5					
	5			3				1	
	6	2	2						
BMDP1M	1	16	14	3					
	2				13	14	13		
	3		1	8				1	
	4				2	2	3	3	2
	5	1	1						
	6	2	2						
EAP-DENDRO	1	16	14	3					
	2				5	5	5		
	3			6				2	
	4				2	2	3	2	1
	5	2	3	2					1
	6				6	6	6		
EAP-GRSIM	1	16	14	3					
	2				7	7	7		
	3	1	1	5				1	
	4				1	1	1	2	
	5								1
	6				1	1	2		1
	7				2	3	2		
	8	2	2						
	9		2	3				1	
	10				4	4	4		
FACTOR ANALYSIS	1	5	5	1					
	2				8	8	8		
	3				1	1	1	2	1
	4	6	3	3			1		
	5	2	3						
	6			5				1	
	7				3	3	3		

Table 10

Number of Adjectives Classified by
APL-r with Five other Solutions

Clusters from		Clusters Using APL Program With Pearson r Index								
		1	2	3	4	5	6	7	8	9
NT-SYS	1	13				7	3	3	1	
	2		14							2
	3	1		4	2				1	
	4	11				1			4	
	5	3							1	
	6		1							
BMDP1M	1	17				8	3	3	2	
	2		14							2
	3	8							4	
	4	1		4	2				1	
	5	2								
	6		1							
EAP-DENDRO	1	17				8	3	3	2	
	2		5							
	3	6							4	
	4			4	1					
	5	5			1				1	
	6		6							2
EAP-GRSIM	1	17				8	3	3	2	
	2		7							
	3	6							1	
	4			2						
	5	2			1					
	6			2	1					
	7		3							2
	8		1							
	9	3							4	
	10		4							
FACTOR ANALYSIS	1	6				4				
	2		8							2
	3	1		3	1					
	4	6		1		3	3	3		
	5	3				1				
	6	5							1	
	7		2							

Table 11

Rand's Statistic for all Pairs of Solutions

	APL-r	NT-SYS	BMDP1M	EAP-DENDRO	EAP-GRSIM	Factor Analysis
APL-q	.788	.809	.873	.808	.816	.731
APL-r		.777	.800	.748	.756	.661
NT-SYS			.918	.843	.850	.837
BMDP1M				.930	.915	.838
EAP-DENDRO					.947	.846
EAP-GRSIM						.959

END