DOCUMENT RESUME

ED 402 332 TM 025 859

| | |
|---|---|
| AUTHOR | Huberty, Carl J.; And Others |
| TITLE | Behavioral Clustering of School Children. |
| PUB DATE | Apr 96 |
| NOTE | 47p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996). |
| PUB TYPE | Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150) |
| EDRS PRICE | MF01/PC02 Plus Postage. |
| DESCRIPTORS | Behavior Patterns; Behavior Rating Scales; *Classification; *Cluster Analysis; Elementary Education; *Elementary School Students; National Norms; Prediction; Research Methodology; *Responses; *Validity |
| IDENTIFIERS | *Behavior Assessment System for Children |

ABSTRACT

This article illustrates how a cluster analysis can be conducted, validated, and interpreted. Data normed for a behavioral assessment instrument with 14 scales on a nationally representative sample of U.S. school children were utilized. The discussion explores the similarity index, cluster method, cluster typology, cluster validity, cluster structure, and prediction of cluster membership. The Behavior Assessment System for Children (BASC) form that assessed 6- to 11-year-old students with the Teacher Rating Form (TRS-C) was used with a sample of 1,228 elementary school children. The clustering method involved a two-step procedure: a Ward hierarchical analysis followed by an iterative cluster partitioning via a K-means analysis. As illustrated, the following steps are suggested for cluster analysis: (1) select the study units (children); (2) choose the system of response variables; (3) decide how to measure the response variables; (4) select the similarity index; (5) select the cluster method; (6) determine the initial cluster typology; (7) provide some evidence of cluster validity; (8) interpret final cluster typology; (9) describe final cluster structure; and (10) develop a classification rule for new units. (Contains 2 figures, 9 tables, and 48 references.) (SLD)

# Behavioral Clustering

# of School Children

Carl J Huberty

Christine DiStefano          Randy W. Kamphaus

University of Georgia

**Paper presented at the annual meeting of the American Educational Research Association, New York, April 1996**

## Abstract

The intent of this article is to illustrate how a cluster analysis might be conducted, validated, and interpreted. Data normed for a behavioral assessment instrument with 14 scales on a nationally representative sample of U.S. school children were utilized. The analysis discussed covers the similarity index, cluster method, cluster typology, cluster validity, cluster structure, and prediction of cluster membership.

## Introduction

A list of 42 studies that attempted to identify learning disability subtypes via empirical grouping methods is given by Hooper and Willis (1989, pp. 64-68). The list may be categorized according to the type of response variables used: achievement (4 studies), neurocognitive (20 studies), neurolinguistic (2 studies), and 16 studies using some combination of variable types. Twelve of the studies used Q-type factor analysis and 30 studies used cluster analysis as the grouping method -- one study used both methods, and one study used multiple regression. In only three of the 25 post-1984 studies was Q-type factor analysis used -- an indication that cluster analysis is the preferred method of late.

The search for subtypes of children with special needs has been studied extensively during the past two decades. Some of the more recent reports have focused on substantive issues of subtyping (e.g., DeLuca, Rourke, & Del Dotto, 1991; Fuerst & Rourke, 1991; Glutting, McGrath, Kamphaus, & McDermott, 1992; Jenkins, Pious, & Peterson, 1988; Korhonen, 1991; McDermott, Glutting, Jones, & Noonan, 1989; Watson & Goldgar, 1988; Williams, Gridley, & Fitzhugh-Bell, 1992); some have focused on methodological issues (e.g., DeLuca, Adams, & Rourke, 1991; Fletcher, Morris, & Francis, 1991; Morris & Fletcher, 1988; Rourke, 1994; Speece, 1990); and some have dealt with both issues (e.g., Fletcher & Satz, 1985; Glutting & McDermott, 1990; Morris, 1988; Swanson & Keogh, 1990). The Hooper and Willis (1989) book might be added to the last list. There is some variation in these research efforts with regard to type of sample, size of sample, type of response variables, focus on type of disability, data analysis method, and validation of results.

Most of the studies cited above involved the clustering of children with some type of disability using, typically, non-behavioral response variables. McKinney (1989) summarizes a program of research on behavioral characteristics of children with learning disabilities. Speece and Cooper (1991, p. 45) support the inclusion of "normal" children in determining behavioral clusters. The study of normal children using behavioral variables is reviewed in some detail by Kamphaus, Huberty, and DiStefano (1996).

The intent behind the current paper is not to review the vast array of previous writings dealing with the grouping of children with special needs, but rather to report a study of subtype identification using behavioral measures on a nationally representative sample of U.S. school children, with an emphasis on data analysis strategy. Some of the data analysis techniques used have been suggested in the previous literature, while others have not.

## Instrumentation and Data

Behavior problems as well as assets of a representative national sample of U.S. children and youth were assessed via the Behavior Assessment System for Children (BASC; Reynolds & Kamphaus, 1992). The BASC has three rating forms: parent, teacher, and self. [In addition, there is a classroom observation system as well as a history form.] The first two forms were used with three groups of subjects: preschool (ages 4-5), 6-11 year old children, and adolescents (ages 12-18). The self form was used only with the latter two groups of subjects. Thus, in the norming process, eight data sets resulted. For purposes of the current study, the one data set containing assessments of the 6-11 year-old children with the teacher rating form (TRS-C) were utilized.

The BASC TRS-C norming data were collected at 116 sites representing various regions of the United States. The sites were selected in order to represent a diverse sampling of the population by geographic region, SES, ethnicity, and child exceptionality. The TRS-C sample used for these analyses included 1228 elementary school children (ages 6-11) who were attending both public and private schools. The TRS-C sample was formally stratified in order to approximate 1986-1988 U.S. Census Bureau statistics. Stratification variables included grade, gender, and ethnicity. African-American and Hispanic children were oversampled to a limited extent in order to ensure adequate representation. TRS-C data collection was conducted in the following manner (Reynolds & Kamphaus, 1992):

> At each participating institution, two classrooms were selected per grade. Within each classroom, two male and two female children were randomly selected for teacher ratings... (p. 85)

In addition, an attempt was made to include children with known exceptionalities in proportion to population characteristics. Characteristics of the normative sample closely approximate population attributes with respect to the distribution of parent education level and percent of children receiving special education services (5.8% females and 9.9% males) (Kamphaus & Frick, 1996).

The TRS-C has 148 items that are rated by the teacher on a four-point range of frequency, from "Never" to "Almost Always." The BASC-TRS was developed using a blend of rational/theoretical and empirical approaches to test development (Martin, 1988). Scales were selected a priori to assess a broad array of maladaptive and adaptive constructs; constructs with prior empirical support were favored. Measurement of the scales was considerably refined and modified based on content reviews and a variety of empirical studies that were conducted on two tryout samples and the normative sample (Reynolds & Kamphaus, 1992). The final 14 TRS-C scales and their descriptions are shown in Table 1.

Four sets of norm tables were developed based on a linear transformation of raw scores to T scores (mean = 50, standard deviation = 10): General, Female, Male, and Clinical. The General national norms were used for the current study for three reasons: (1) gender-separate norms mask gender differences (Kamphaus & Frick, 1996), gender differences on the scales were exceedingly small with the most exceptional cases approaching a difference of one-half of a standard deviation (Reynolds & Kamphaus, 1992); (2) preliminary cluster analyses conducted as part of this study produced highly similar typologies when gender norms were used; and (3) gender criteria for diagnosis are not used by major classification systems such as the DSM IV.

The BASC-TRS manual (Reynolds & Kamphaus, 1992) provides three types of reliability evidence: internal consistency, test-retest reliability, and interrater reliability. The internal consistency coefficient values and numbers of scale items are given in Table 2; seven of the total of 148 items are not associated with any particular scale. The manual presents evidence of factor analytic support for the construct validity of the scales using both principal axis and covariance structure analyses. The TRS scales also typically exhibit high correlations with analogous scales from other teacher rating instruments (Kamphaus & Frick, 1996). Several independent reviews of the BASC have noted that the TRS possesses adequate to good evidence of reliability and validity using a variety of indicators although, as a relatively new instrument, considerably more research is desirable (Adams & Drabman, 1994; Flanagan, 1995; Hoza, 1994, Kline, 1994; Sandoval & Echandia, 1994; Witt, 1994).

The data matrix considered for the current analysis was a 1228-by-14 matrix. There were

1228 children aged 6-11, on each of whom were obtained scores on the 14 Teacher Rating scales; each scale score was a T score. Correlations among the 14 scales, based on N=1228, are reported in Table 3; a summary of the distribution of correlation absolute values is:

Max = .82, $C_{75}$ = .60, $C_{50}$ = .47, $C_{25}$ = .33, Min = .04,

where $C_{75}$ denotes the 75th centile of the distribution. The Externalizing Problems scale subset (scales 1-3) and the Adaptive Skills scale subset (scales 11-14) are the most highly interrelated scale subsets.

All analyses for this study were done using the SAS statistical package, Version 6.08.

## Preliminary Data Analysis Considerations

In any multiple response variable research situation, an initial decision pertains to the choice of variables. In the current situation, the basic variable set was the collection of behavioral scales defined by the BASC. We decided not to consider using any masking variables or noise variables (as per Milligan & Cooper, 1987, p. 344). Another decision to consider, in general, is whether or not to standardize the variable measures (see, e.g., Milligan & Cooper, 1988). In our case, this decision was predetermined because T scores were the only measures used in norming the BASC. [There is some evidence (e.g., Edelbrock, 1979) that whether or not standardizing variable scores is desirable is a nonissue.]

A data-preparation consideration to be made prior to conducting a cluster analysis is the completeness of the data matrix. That is, a search for missing data needs to be conducted. For the current data set, there were no 6-11 aged children with missing behavioral measures. Thus, no data imputation methods were needed -- such methods are discussed by Little and Rubin

(1987) and Reilly (1993).

Another consideration to be made prior to a cluster analysis is the existence of outlying children. There are numerous methods one can use in detecting outliers (see, e.g., Barnett & Lewis, 1994). The method used in this study involved Euclidean distance -- this is consistent with the (dis)similarity index used in this study for the cluster analysis. For the 1228 children, the Euclidean distance was calculated from the score vector for a given child to the score vector for each of the other 1227 children. For each child, the maximum of such distances was set aside. Thus, a distribution of 1228 maximum distances was determined. This distribution was examined to identify potential outliers. The maximum distances ranged from 85.1 to 157.6. Visual inspection suggested no gaps in the maximum distance distribution which led to the conclusion that there were no children who should be considered as outliers.

The final preliminary consideration made pertained to the index of similarity (or, dissimilarity) to use (see, e.g., Aldenderfer & Blashfield, 1984, pp. 16-33). In our case, we decided to use the popular index, Euclidean (squared) distance. This made sense to us because of the use of T scores, plus the advice advanced by Blashfield and Aldenderfer (1988, p. 460) and others who have studied the similarity issue. The Euclidian index is sensitive to profile elevation and dispersion (as well as profile shape) which were judged to be particularly relevant for assessing the similarity of children with respect to the behaviors considered.

## Cluster Analyses

### Cluster Method

As pointed out by a number of writers (e.g., Aldenderfer & Blashfield, 1984, pp. 33-62;

Kaufman & Rousseeuw, 1990), there is a fairly wide variety of methods one might use to identify groups/clusters/subtypes of children. The clustering method selected for use in the current study involved a two-step procedure: a Ward hierarchical analysis followed by an iterative cluster partitioning via a K-means analysis. The Ward method was chosen because of its overall cluster recovery ability and sensitivity to profile elevation and dispersion (Milligan & Cooper, 1987; Morey, Blashfield, & Skinner, 1983). Because of the behavior measures used in the current study, child profile elevation and dispersion were considered a potentially important determiner of cluster typology. A drawback of a Ward analysis is that once a child is assigned to a cluster, cluster membership cannot change. The cluster centroids obtained from the Ward analysis were used as "seeds" (i.e., starting points) in conducting a K-means analysis. The intent behind the use of a K-means analysis was to make possible some shifts in cluster membership of some children. Such membership shifts are accomplished in such a way that the cluster homogeneity of a Ward analysis is not appreciably sacrificed and may, in fact, be enhanced. Some empirical support for this analysis strategy is summarized by Milligan and Cooper (1987).

Initial Cluster Solution

With the Ward solution followed by the K-means analysis, the basic decision to be made pertains to the number of clusters to consider. The cubic clustering criterion obtainable via SAS CLUSTER was used for starters. A plot of this criterion versus number of clusters (as determined by an "elbow" in the plot) suggested between 4 and 11 clusters. Solutions (i.e., centroids) were determined for these cluster numbers, so that a substantive examination could be made. The final number of clusters was based on several rational considerations. Two considerations aided in determining the number of clusters to retain: (1) five comparable clusters appeared repeatedly in the 6- through 8-cluster solutions; and (2), a cluster was not retained if it was differentiated from others by only elevation or shape.

Cluster meaningfulness was determined using several rational criteria including cluster mean deviance from average (e.g., clusters with deviant T scale scores may reflect known patterns of psychopathology), gender distribution (e.g., gender breakdowns should be similar for less deviant groups, with greater male representation in pathological groups and in those marked by externalizing problems consistent with epidemiological research), similarity of profile shape to well recognized syndromes (e.g., a cluster with deviant T scale scores for the Depression and Anxiety scales would more likely be retained than one with deviant Anxiety and Conduct Problems scales because of the documented comorbidity of depressive and anxiety problems), predictable characteristics of the subtypes based on related research (e.g., Learning Problems elevations that are commonly associated with disruptive behavior problems), similarity to subtype dimensions that have been previously identified in the child psychopathology literature (e.g., deviant Hyperactivity versus Attention Problem scores resembling ADHD subtypes), size of cluster (e.g., the largest clusters should hover at about the normative mean for a nationally representative sample), and consistency with TRS prepublication research (e.g., a profile similar to that obtained for a diagnostic group that was sampled as part of the validation process).

Using the above criteria, it was judged that a seven cluster typology was most reasonable for the 1228 children. To obtain this seven-cluster solution, the number of iterations for the K-means was five. A substantive description of each of the seven clusters is given later in this section.

Cluster Validation

The claim has often been made in the literature by methodologists that some type of cluster typology is obtainable even with "random data" (the meaning of which is not always made clear; see Abelson, 1995, chp. 2). If so, then it behooves the researcher to somehow attempt to compare a cluster typology resulting from one data set with that from another relevant and appropriate data set. The term often used in making such a comparison is "validation," even though the term "reliable assessment" may be preferred by some. Whatever, it is desirable to present results that will give some assurance that the cluster typology interpreted approximates the "true typology," however that is interpreted (see, e.g., Milligan & Cooper, 1987, pp. 333-335). Another view of "validity" may be expressed in a question: Are the resulting clusters "real," or are they artifacts of the analysis methods used?

Before discussing three proposed validation methods, some comments pertaining to two data conditions are offered. One condition is that of multivariate normality. This condition is theoretically required for the first and third validation methods. The seven data sets for the initially proposed cluster typology were checked for normality by examining the seven chi-squared probability-plots -- obtained via the SAS OUTLIER macro (Friendly, 1991, p. 451). An examination of the resulting plots indicated some "skewness," but it was judged that the lack of normality was not too extensive.

The second data condition of relevance is the near-equality of the seven 14 x 14 covariance matrices. Such a comparison is very difficult to accomplish by simply "eyeballing" the seven 14 x 14 covariance matrices. At the same time, a statistical test with 14 outcome variables, seven groups, and N of 1228 is extremely powerful in a statistical sense. For the seven clusters in the current study, a transformation of the Box M criterion leads to an F (630, 366100) value of 5.93 with P = .0001. Many researchers would conclude that the seven corresponding population covariance matrices are not equal. And, therefore, the conclusion drawn would be that the consideration of the use of linear discriminant functions (i.e., linear composites) in the validation process is inappropriate. Perhaps so. One might argue, however, that linear discriminant functions may be of some descriptive (as opposed to inferential) value in the face of such statistical test results. The reasonableness of such an argument may be enhanced somewhat if one can assume that the children in the seven clusters constitute representative samples from the respective corresponding populations. For the current situation this assumption will be made, and we will proceed with the extraction of LDFs (for descriptive purposes).

What was done with the 1228 x 14 data set to address the validity-reliability issue was to do analyses on half-samples and compare the resulting cluster typologies. We randomly split the whole (N = 1228) data set into two m x 14 data sets. The splitting of the total sample into two half-samples was done three times to obtain three distinct pairs. The m for the half-samples ranged from 598 to 630. Each half-sample was clustered using the Ward analysis followed by a K-means analysis which was described earlier. The number of iterations for the K-means analyses across the three pairs ranged from 3 to 20. Comparisons of the cluster typologies for each of the three pairs of half-samples were made in three ways:

1. Comparison of group typologies. For each half-sample of each pair, a linear discriminant function (LDF) structure was determined using the SAS CANDISC procedure. The (canonical) structure considered for each half-sample is that determined by (error) correlations between LDF scores and scale scores (Huberty, 1994, p. 209). The structure r's for the first half

are correlated with the structure r's for the second half. Now, with seven groups (i.e., clusters) in each half-sample, it is possible to obtain six LDFs. Looking at the proportions of variance in the 14-scale system attributed to each LDF, it was concluded that at most three LDFs should be retained (Huberty, 1994, p. 214). The cumulative proportion of variance for three LDFs was at least 94% for each pair of half-samples.

The correlations between the corresponding structure r's for the three pairs of half-samples are reported in Table 4. Eight of the nine correlations are judged to be "high." What this indicates is that the separation (in 2-3 dimensions) of the clusters in one half-sample in a pair is comparable (in a correlative structure sense) to the separation in the other half-sample.

It is recognized that for a given half-sample, the seven-group covariance matrices may not be "in the same ballpark." The comparability of the covariance matrices was assessed by examining the patterns of the logarithms of the covariance matrix determinants across the two half-samples for each pair. The lack of comparability of the covariance matrices in one half was judged to be fairly similar to the lack of comparability in the other half for each of the three pairs of half-samples. Therefore, it seemed reasonable to compare the linear canonical structures of the two half-samples for each of the three pairs. [It should be noted that the LDFs were considered for purposes of half-sample comparability, not for substantive interpretation purposes.]

2. Cross-typology clustering. Another comparison of the cluster structure of each pair of half-samples was accomplished as follows:

(a) Use the final (Ward followed by K-means analysis) cluster means for the first half as a "seed" for assigning children from the second half via a single pass of a K-means analysis. [This is an adaptation of an approach discussed by McIntyre and Blashfield (1980).] This cross-typology clustering was applied only to clusters in one half of a given pair that were "matched" with clusters in the other half. For a given pair of half-samples, clusters were matched on the basis of substantive judgment (by examining each of the seven cluster centroids). There were five matched clusters for the first and second pairs that comprised about 70 percent of each half-sample. There were four matches for the third pair that comprised about 60 percent of each half-sample.

(b) Repeat (a) using the final cluster means for the second half as a "seed" for assigning children from the first half.

(c) For each of (a) and (b), develop a k x k table of "hits" (on the main diagonal) and "misses" for the seven clusters; k denotes the number of cluster matches.

(d) Determine whether each of the two k x k tables for each of the three pairs had proportions of total-group hits that were better than what may be expected by chance (Huberty, 1994, pp. 102-107); the set of prior probabilities used to obtain expected hit rates for each pair were estimated by the current writers.

(e) Assuming that the total-group hit rate was better than chance, calculate an "improvement over chance" statistic (I) value (Huberty, 1994, p. 107) for each of the k x k tables.

A summary of the hit rates for the cluster matches in the three pairs of half-samples is given in Table 5. It is obvious from the reported hit rates that the cross-typology clustering was accomplished with a fairly high degree of agreement. All across-cluster hit rates reported in Table 5 are higher than the corresponding hit rates expected by chance, and the six I values ranged from 61.9 to 95.9. Thus, there would be at least 61.9% fewer classification errors made using the proposed cross-typology clustering than if chance classification were used.

3. <u>Comparison of matched cluster centroid location</u>. The third and final method of comparing the cluster typologies of the matched clusters of the two half-samples within each pair involved plots of the centroids in the space of the two leading linear discriminant functions (LDFs). To repeat, for the first and second pairs there were five matched clusters, and for the third pair there were four matches. So that the LDF plots could be compared for the two half-samples, it was necessary to reverse the signs of the weights for one LDF (and thus reverse the sign of the LDF mean). [A set of LDF weights is only unique up to a constant of proportionality.] It was necessary to do this for one half-sample in each of the three pairs. Once the LDF weights were comparable, the centroids for the cluster matches were plotted in a two-dimensional LDF space. It was judged from the plots that the LDF centroids for the matched clusters were in very close proximity. For example, the five matched cluster LDF centroids for the second pair of half-samples are plotted in Figure 1.

It is concluded from the information yielded by conducting the three types of comparisons of the three pairs of half-samples that the initial seven-cluster solution is one that is not an artifact of our clustering method. [It is recognized that an alternative cluster method might generate an alternative solution.] It was thus concluded from the three comparisons made that the initial seven-cluster typology based on $N = 1228$ was "valid." That is, we were convinced that we obtained a legitimate typology which we should proceed to "interpret." In the next subsection, a rationale is provided for the definition of what we judge are meaningful (i.e., "real") clusters.

## <u>Cluster Typology</u>

The seven-cluster solution based on all 1228 children is presented in Table 6. The rationale for the interpretation of each cluster and applying its name is summarized next.

Cluster 1 is the largest of the seven clusters representing approximately one third of the national sample. It is labeled **Well Adapted** because of the significant elevations on the adaptive scales and the absence of behavior problems. The gender representation of this cluster is also predictable with twice as many girls as boys.

Cluster 2 is labeled **Average** because there are few deviations from a normative mean and the gender composition of the cluster is roughly 50/50. Clusters 1 and 2 combined make up over one-half of the students sampled (53%) suggesting that most children in this age range are free of problems and one third of them also possess strengths in study skills, social skills, leadership abilities, and they adapt well to changes in the environment.

Cluster 3 appears to represent what is commonly referred to as **Disruptive Behavioral Disorder** (Frick, et al., 1991). The mean scores for the externalizing scales for this cluster meet or surpass those for the samples of children with conduct disorder, behavior disorder, and attention deficit hyperactivity disorder (ADHD) that were collected as part of the TRS validation process (Reynolds & Kamphaus, 1992, p. 125). Moreover, this cluster is marked by significant adaptive behavior deficits and elevations on internalizing scales including Depression. The male dominance of this cluster is also consistent with expectations. The size or epidemiology of this group, comprising 8% of the sample, is not surprising given the frequency of occurrence of problems such as conduct disorder (Kazdin, 1995).

Cluster 4 is very similar to the profile obtained for a large learning disability sample with one exception (Reynolds & Kamphaus, 1992, p. 125). The cluster 4 members possess more significant deficits in adaptive skills. Because cluster 4 so closely mimics the sample of children with diagnosed learning disabilities, and is more severely impaired, a tentative label of **Learning**

**Disorder** seems appropriate.

Cluster 5, **Physical Complaints/Worry**, is marked by internalizing problems of a mild nature with somatic complaints being primary and symptoms of anxiety (chiefly worry and nervousness) secondary. Given the known epidemiology of internalizing problems, the greater female occurrence rate is predictable based on related research regarding internalizing problems. Given, however, that the problem scale elevations are small, and these childrens' adaptive skills are average, a nonclinical internalizing label is proposed. Only 6% of the sample is diagnosed. Additionally, this profile does not mimic any TRS validation profiles such as the one for depression (Reynolds & Kamphaus, 1992, p. 125).

Cluster 6 is clearly the most severely impaired of all, comprising 4% of the national sample. This cluster is dominated by males with diverse problems including psychotic thought processes (high Atypicality score) and impaired adaptive skills. Members of cluster 6 have more severe problems but they do resemble somewhat the validation sample of children who were diagnosed by school personnel as emotionally disturbed. Therefore, the label of **Severe Psychopathology** is proposed for this cluster.

Cluster 7 differs from cluster 3 in both shape and elevation. It is marked by mild scale elevations for only the varibles Aggression, Hyperactivity, and Adaptabilities. In many ways this profile looks like a subclinical form of disruptive behavior problems; we propose the label, **Mildly Disruptive**. This group of children may, however, be adjusting adequately in school as indicated by their adaptive skills scores. This cluster of children, along with the cluster 3 children may explain the high referral rate for boys suspected of ADHD. Together these two clusters make up 20% of the sample and perhaps the school age population.

Other findings lend credence to the seven cluster solution. As would be predicted, members of clusters 3, 4, and 6 are diagnosed at a higher rate than is typical of children in other clusters. Furthermore, the fact that females are at significantly lower risk for diagnosis (about half of special education samples) is also consistent with the composition of the clusters.

## Post-Typology Analyses

In most applications of clustering subjects in the behavioral sciences, a cluster solution is found using some cluster analysis strategy, and the resulting typology is discussed from a substantive point of view. This is all well and good, and needs to be done. It is proposed here that there are two additional sets of analyses of potential interest that might yield theoretical and practical information. One set of analyses pertains to the scale structure associated with the cluster typology, and the other pertains to the development of a prediction rule for associating a child with one (or more) of the seven clusters.

### Cluster Structure

Let us assume that the cluster typology determined is interesting, makes sense, and contributes something to the understanding of a collection of experimental units, which in the current situation was a sample of school children. It may be of interest, then, to study cluster differences in making an attempt to address the question: In what sense(s) are the clusters different? or, With respect to what unit attributes do the clusters differ? or, On what unit attributes do the clusters have an effect? [These are meant to be equivalent questions.] To address this common question, that which pertains to cluster structure, one can examine the linear discriminant functions (LDFs) associated with the cluster differences. Sample linear functions are only technically appropriately considered when it is reasonable to assume that the covariance matrices of the populations corresponding to the obtained clusters are approximately equal. This data condition, along with that pertaining to multivariate normality, was discussed in a previous section.

With seven clusters and 14 behavior scales, it is possible to extract six LDFs. By examining the proportions of variance of the system of 14 scales (which are, respectively, .72 , .14 , .08 , .05 , .01 , and .01), it was concluded that at most three LDFs should be given serious consideration (see Huberty, 1994, p. 214). The structure r's for the three leading LDFs are reported in Table 7. These structure r's may be utilized in labeling constructs that underlie cluster differences, and thus define a "cluster structure." Some additional interpretation of the resulting cluster typology may be obtained by viewing a plot of the cluster centroids in the space of the first two associated LDFs that, in this case, account for about 86% of the variability in the 14-variable system. Such a plot is given in Figure 2. From this plot, we see differences among cluster 1, cluster 2, clusters 5 and 7, cluster 4, cluster 3, and cluster 6. The other two-dimensional LDF plots (i.e., $LDF_1$ versus $LDF_3$ , and $LDF_2$ versus $LDF_3$) were also considered.

Joint examination of the structure r's and the three LDF plots led to the naming of the three LDFs: $LDF_1$ , "General Psychopathology"; $LDF_2$ , "Adaptive Skills"; and $LDF_3$ , "Affective Disorder." The General Psychopathology label seems appropriate for $LDF_1$ for several reasons, including the large amount of variance accounted for by this factor, the large number of clinical scales with sizeable structure r's, and the ordering of the seven clusters in the LDF space by severity (see Figure 2). $LDF_2$ is clearly marked by the adaptive scales of the BASC TRS suggesting that the adaptive scales label is easily applied. The third and smallest LDF, accounting for a mere 8% of the variance is the most difficult to label based on substantive related research. The tentative label of Affective Disorder is offered based on the theory that Somatization and Aggression, the scales that have the highest structure r's, assess contructs associated with depression and bipolar disorder (Emslie, Kennard, & Kowatch, 1995).

The interpretation of the cluster structure discussed above is based on the use of structure r's, correlations between each scale and each LDF. Another index for interpreting cluster structure using an LDF is the set of standardized LDF weights, as suggested by Harris (1993). As appealing as this index may be to some methodologists, we did not find with our results that it lead to nearly as meaningful constructs as did the use of structure r's; in fact, construct definitions based on the weights were virtually meaningless.

Prediction of Cluster Membership

The second follow-up analysis to consider involves the development of a rule for predicting membership in the seven clusters on the basis of the 14 behavior scale scores. Given the determined cluster typology, a question to be addressed is: How well can cluster membership be predicted for children on whom we have the 14 behavior scale measures? The intent of addressing this question is a very practical one. If we are able to fairly accurately predict cluster membership (i.e., identify a cluster with which a "new" child having a vector of 14 behavior scale measures may be associated), then educators may be in a better position to advocate some particular educational intervention strategies for the child.

Details regarding the development of a prediction rule will not be given here (see Huberty, 1994, chap. IV). It was decided that a normal-based linear classification rule would be used. The use of a normal-based linear rule is technically appropriate if it is reasonable to assume that the seven population distributions of 14-element score vectors are approximately multivariate normal, and that the seven group 14 x 14 covariance matrices are approximately equal. Discussions of these two data conditions were presented in a earlier section. It turns out that with N = 1228 and 14 variables -- a ratio of over 85:1 -- greater

confidence in classification results can be obtained with a linear rule than with a quadratic rule (Huberty, 1994, p. 260).

To complete the derivation of the rule to be used, estimates of prior probabilities of cluster membership had to be made. These estimates should reflect the relative population sizes. To arrive at reasonable estimates, six "experts" (excluding all current authors) in the field of child clinical psychology were consulted and asked to estimate the seven population sizes (in terms of proportions). An "averaging" of the six sets of proportions resulted in the following respective priors: .15, .45, .08, .10, .10, .02, and .10. These priors, then, were incorporated into the linear classification rule.

To assess the efficacy of the rule, an external analysis was used. This analysis involved developing a rule on one data set, and applying the rule to other data. The particular external analysis favored is a leave-one-out (L-0-0) analysis (see Huberty, 1994, pp. 88-90). With this analysis, one child is deleted and a rule is built on the remaining 1227 children; the rule is then applied to the deleted child. This process is repeated for the total set of 1228 children. It was in this manner that a number of "hits" was determined for each of the seven clusters as well as across all clusters.

The L-0-0 hit rate estimates are given on the main diagonal of the classification table in Table 8. It is obvious from the main diagonal entries in Table 8 that estimated the individual cluster hit rates are quite high and, thus, there is very limited overlap among the clusters. All of the 228 "average" children (in cluster 2) were correctly so identified. The misclassification rates for the other six clusters ranged from 10.7 (11/103 for cluster 3) to 23.4 (34/145 for cluster 7). All seven cluster hit rates are significantly higher than a chance hit rate, as is the across-cluster hit rate. The smallest improvement-over-chance I index value was 73.9 for cluster 7 -- that is, for cluster 7, about 73.9% fewer classification errors would be made using the derived linear rule than using chance. [See Huberty (1994, chap. VII) for a discussion of the statistical test and the I index.]

It turns out that the results of a quadratic L-0-0 classification analysis yielded cluster hit rates very similar to those for the linear analysis. The quadratic hit rates for the seven clusters were, respectively, 83.0, 92.5, 84.5, 78.5, 85.1, 80.8, and 82.1. The first three quadratic hit rates were a little lower than the corresponding linear hit rates, whereas the quadratic hit rates for clusters 5 and 7 were a little higher. A statistical comparison of the two sets of results may be made via a McNemar test (Huberty, 1994, pp. 108-110). It turned out in this case that all children correctly classified by the linear rule were also correctly classified by the quadratic rule. Because no "significant" improvement would be gained by using a quadratic rule rather than a linear rule, and because a linear rule is expected to be more stable over repeated sampling, and because a linear rule may be easier to apply with new children, a linear rule is clearly preferable for the current data set.

So much for the classification results from a cluster standpoint. Now let us discuss some classification results from an individual child standpoint. Even though it would be predicted that a child should be assigned to one particular cluster, might he/she also be identified with some other cluster nearly as well? To address this question we seek out those children who might be "indoubt" or "fence-rider" cases. These children would have two (or more) posterior probabilities of group membership that were "close." Two posterior probabilities were defined to be close if their absolute difference was less than .01. A rather stringent (i.e., low) difference was judged to be appropriate in this situation because of the relatively little cluster overlap. With this criterion, a total of only 21 fence-riders were found. In this situation, it would not be expected that such a small number would drastically affect any cluster hit estimate -- unless, of course the bulk of the 21 involved cluster 6 where n = 52, which was not the case. [In fact, none of the 21 fence-riders were associated with cluster 6.] Three children were "on the border" between cluster 2 (average) and cluster 5 (Physical Complaints/Worry), two of whom emanated from cluster 5 and were assigned to cluster 2. The point to be made with regard to fence-riders is simple: When assessing classification results, we should be aware of the possibility of units (e.g., children) who may belong to more than one cluster. In some (rare?) situations one may find units that have <u>three</u> posterior probabilities that are close in numerical value. The consideration of possible fence-riders is potentially important, also, (from a practical standpoint) when new units are being classified.

A rule to use with new children is in the form of a set of seven linear composites of the 14 scales. Weights for the linear composites -- called linear classification functions (LCFs) -- obtained from our 1228 children are given in Table 9. The intent in presenting Table 8 is to give the reader an idea of what a classification rule "looks" like. To apply such a rule in practice, more precise weights would be used; weights to four or five decimal places, not two. It should be noted that the prior probabilities are included in the calculation of the constants. So, to use a rule such as that given in Table 8 with a new child, one would apply the seven sets of weights to the vector of 14 scale scores (and adding the constants) and determine seven LCF scores. The child would be assigned to the cluster with which is associated the largest LCF score. If one LCF score is clearly the largest, then some confidence could be gained in selecting an appropriate intervention strategy for the child. If the two largest LCF scores are "close," then the collection of more information on the child may be desirable to decide on the intervention strategy to employ.

For example, consider the following vectors of 14 scale scores for five new children:

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 | S11 | S12 | S13 | S14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1229 | 46 | 50 | 49 | 47 | 59 | 48 | 48 | 57 | 37 | 57 | 31 | 46 | 36 | 64 |
| 1230 | 54 | 44 | 52 | 47 | 43 | 43 | 41 | 50 | 53 | 59 | 63 | 46 | 52 | 51 |
| 1231 | 46 | 63 | 73 | 54 | 59 | 59 | 80 | 73 | 55 | 61 | 55 | 59 | 39 | 51 |
| 1232 | 43 | 46 | 45 | 59 | 47 | 46 | 45 | 46 | 37 | 71 | 37 | 46 | 39 | 58 |
| 1233 | 43 | 55 | 72 | 55 | 51 | 61 | 49 | 57 | 39 | 65 | 44 | 59 | 39 | 51 |

Applying the five-place weights (and constant) we find the two largest LCF scores for the five new children to be:

| | | |
|---|---|---|
| 1229 | $LCF_4 = 332.52$ | $LCF_2 = 330.23$ |
| 1230 | $LCF_2 = 337.51$ | $LCF_1 = 337.08$ |
| 1231 | $LCF_3 = 422.33$ | $LCF_7 = 418.89$ |
| 1232 | $LCF_4 = 295.21$ | $LCF_2 = 294.30$ |
| 1233 | $LCF_5 = 353.88$ | $LCF_4 = 352.63$ |

The cluster assignments are indicated by the subscript on the larger LCF score. For example, child 1231 would clearly be associated with cluster 3 (Disruptive Behavior Disorder), while child 1233 would be assigned to cluster 5 (Physical Complaints/Worry), but less decisively.

If one has access to the original set of children on whom the rule was based, there is a straight-forward approach to assigning a new child. With this approach, one simply includes the new children's vectors of BASC scale scores in with the original set but with no cluster identification. The SAS DISCRIM procedure will calculate the seven posterior probabilities of cluster membership for each new child, values of which may be used in making a cluster assignment (see Huberty, 1994, pp. 112-113). For the five new children indicated above, the two largest (linear L-0-0) posterior probabilities are:

| | | |
|---|---|---|
| 1229 | $PP_4 = .90$ | $PP_2 = .10$ |
| 1230 | $PP_2 = .59$ | $PP_1 = .38$ |
| 1231 | $PP_3 = .97$ | $PP_7 = .03$ |
| 1232 | $PP_4 = .70$ | $PP_2 = .30$ |
| 1233 | $PP_5 = .53$ | $PP_4 = .41$ |

The cluster assignments are indicated by the subscript on the larger PP value. These assignments are the same as those based on the LCF scores. Cluster identifications for children 1229 and 1231 are fairly clear-cut, but not so for child 1233. It is clearly easier to identify potential fence-riders via the posterior probability values than via the LCF scores.

In a practical, real-life setting, it would be desirable to update the cluster typology and classification rule when a sizable number of new children are assessed via the BASC.

## Summary

As mentioned early in this paper, the intent was to report a subtyping of "normal" children using behavioral response measures, and to illustrate the conduct of a cluster analysis. The subtyping or cluster typology is summarized in Table 5 with a substantive description given in the text of this paper. Now a summary of suggested steps for a cluster analysis study is given:

1. Select study units (e.g., children)
2. Choose system of response variables

3. Decide on how to measure the response variables
4. Select similarity index
5. Select cluster method
6. Determine initial cluster typology
7. Provide some evidence of cluster validity
8. Interpret final cluster typology
9. Describe final cluster structure
10. Develop classification rule for new units

Comments on some of these steps will now be offered. Step 2 is dependent on the purpose of the study. For example, with respect to what unit (e.g., children) characteristics is the clustering of interest? Is the study being done to support or verify or falsify some theoretical position? Or, is there a practical implication or utility associated with the resulting cluster topology? The important aspect of step 2 is that careful consideration should be given to choosing an appropriate collection of response variables; a collection that "makes sense." In connection with step 3, it goes without saying, perhaps, that meaningfulness results of a cluster analysis depend upon the data used as input. The quality (in terms of validity and reliability) of the measurement methods used should be made clear. There are many choices to be made in steps 4 and 5. In this study we suggested a Ward analysis followed by a K-means analysis. In other situations -- that is, with other types of units and other response variables -- alternative analysis methods may be preferred. Step 7 may be approached in a number of ways. Three approaches were presented herein. Another possible "validation" method involves bootstrapping. A number of bootstrap samples (of, say, size 1228 each) could be selected (with replacement);

some type of comparison(s) of the resulting cluster typologies could be made. It is recognized that comparable bootstrap results may imply to some researchers that what is being assessed is reliability of a typology rather than validity of a typology. Whatever, it is urged that some evidence of validity/reliability be provided. Step 8 may be viewed as an effort to confirm some "theory," or simply as an effort to describe cluster differences in terms of the response variable system employed. Step 10 is a practical application of the obtained cluster typology. A prediction rule may be of some utility in a clinical setting where particular developmental interventions may need to be suggested.

## References


Abelson, R. P. (1995). Statistics as Principled Argument. Hillsdale, NJ: Erlbaum.

Adams, C. A., & Drabman, R. S. (1994). BASC: A critical review. Child Assessment News, 4, 1-5.

Aldenderfer, M. S., & Blashfield, R. K. (1984). Cluster Analysis. Beverly Hills, CA: Sage.

Barnett, V., & Lewis, T. (1995). Outliers in Statistical Data, (3rd ed.). New York: Wiley.

Blashfield, R. K., & Aldenderfer, M. S. (1988). Methods and problems of cluster analysis. In J. R. Nesselroade & R. B. Cattell (Eds.), Handbook of Multivariate Experimental Psychology, Second Edition. New York: Plenum.

DeLuca, J. W., Adams, K. M., & Rourke, B. P. (1991). Methodological and statistical issues in cluster analysis. In B. P. Rourke (Ed.), Neuropsychological Validation of Learning Disability Subtypes. New York: Guilford.

DeLuca, J. W., Rourke, B. P., & Del Dotto, J. E. (1991). Subtypes of arithmetic-disabled children: Cognitive and personality dimensions. In B. P. Rourke (Ed.), Neuropsychological validation of learning disability subtypes. New York: Guilford.

Edelbrock, C. (1979). Comparing the accuracy of hierarchial clustering algorithms: The problem of classifying everybody. Multivariate Behavioral Research, 14, 367-384.

Emslie, G. J., Kennard, B. D., & Kowatch, R. A. (1995). Affective disorder in children: Diagnosis and management. Journal of Child Neurology, 10, S42-S49.

Flanagan, R. (1995). A review of the behavior assessment system for children (BASC): Assessment consistent with the requirements of the individuals with Disabilities Education Act (IDEA). Journal of School Psychology, 33, 177-186.

Fletcher, J. M., Morris, R. D., & Francis, D. J. (1991). Methodological issues in the classification of attention-related disorders. Journal of Learning Disabilities, 24, 72-77.

Fletcher, J. M., & Satz, P. (1985). Cluster analysis and the search for learning disability subtypes. In B. P. Rourke (Ed.), Neuropsychology of Learning Disabilities. New York: Guilford.

Frick, P. J., Kamphaus, R. W., Lahey, B. B., Loeber, R., Christ, M. A. G., Hart, E. L., & Tannenbaum, L. E. (1991). Academic underachievement and the disruptive behavior disorders. Journal of Consulting and Clinical Psychology, 59, 289-294.

Friendly, M. (1991). SAS System for statistical graphics. Cary, NC: SAS Institute.

Fuerst, D. R., & Rourke, B. P. (1991). Validation of psychosocial subtypes of children with learning disabilities. In B. P. Rourke (Ed.), Neuropsychological Validation of Learning Disability Subtypes. New York: Guilford.

Glutting, J. J., McGrath, E. A., Kamphaus, R. W., & McDermott, P. A. (1992). Taxonomy and validity of subtest profiles on the Kaufman Assessment Battery for Children. Journal of Special Education, 26, 85-115.

Glutting, J. J., & McDermott, P. A. (1990). Patterns and prevalence of core profile types in the WPPSI standardization sample. School Psychology Review, 19, 471-491.

Harris, R. J. (1993). Multivariate analysis of variance. In L. K. Edwards (Ed.), Applied analysis of variance in behavioral science (pp. 255-296). New York: Marcel Dekker.

Hooper, S. R., & Willis, W. G. (1989). Learning disability subtyping. New York: Springer-Verlag.

Hoza, B. (1994). Review of the Behavior Assessment System for Children. Child Assessment News, 4, 8-10.

Huberty, C. J (1994). Applied Discriminant Analysis. New York: Wiley.

Jenkins, J. R., Pious, C. G., & Peterson, D. L. (1988). Categorical programs for remedial and handicapped students: Issues of validity. Exceptional Children, 55, 147-158.

Kazdin, A. E. (1995). Conduct disorders in childhood and adolescence. Thousand Oaks, CA: Sage.

Kamphaus, R. W., & Frick, P. J. (1996). Clinical assessment of child and adolescent personality and behavior. Needham Heights, MA: Allyn & Bacon.

Kamphaus, R. W., Huberty, C. J, & DiStefano, C. (1996). A typology of child behavior in U. S. classrooms. Unpublished manuscript.

Kaufman, L., & Rousseeuw, P. J. (1990). Finding groups in data. New York: Wiley.

Kline, R. B. (1994). New objective rating scales for child assessment. Journal of Psychoeducational Assessment, 12, 289-306.

Korhonen, T. T. (1991). Neuropsychological stability and prognosis of subgroups of children with learning disabilities. Journal of Learning Disabilities, 24, 48-57.

Little, R. J. A., & Rubin, D. B. (1987). Statistical analysis with missing data, New York: Wiley.

Martin, R. P. (1988). Assessment of personality and behavior problems: Infancy through adolescence. New York: Guilford.

McDermott, P. A., Glutting, J. J., Jones, J. N., & Noonan, J. V. (1989). Typology and prevailing composition of core profiles in the WAIS-R standardization sample. Psychological Assessment: Journal of Consulting and Clinical Psychology, 1, 118-125.

McIntyre, R. M., & Blashfield, R. K. (1980). A nearest-centroid technique for evaluating the minimum-variance clustering procedure. Multivariate Behavioral Research, 15, 225-238.

McKinney, J. D. (1989). Longitudinal research on the behavioral characteristics of children with learning disabilities. Journal of Learning Disabilities, 22, 141-150, 165.

Milligan, G. W., & Cooper, M. C. (1987). Methodology review: Clustering methods. Applied Psychological Measurement, 11, 329-354.

Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. Journal of Classification, 5, 181-204.

Morey, L. C., Blashfield, R. K., & Skinner, H. A. (1983). A comparison of cluster analysis techniques within a sequential validation framework. Multivariate Behavioral Research, 18, 309-329.

Morris, R. D. (1988). Classification of learning disabilities: Old problems and new approaches. Journal of Consulting and Clinical Psychology, 56, 789-794.

Morris, R. D., & Fletcher, J. M. (1988). Classification in neuropsychology: A theoretical framework and research paradigm. Journal of Clinical and Experimental Neuropsychology, 10, 640-658.

Reynolds, C. R., & Kamphaus, R. W. (1992). Behavior assessment system for children (BASC). Circle Pines, MN: AGS.

Reilly, M. (1993). Data analysis using hot deck multiple imputation, The Statistician, 42, 307-313.

Rourke, B. P. (1994). Neuropsychological Assessment of Children with Learning Disabilities. In G. R. Lyon (Ed.) Frames of reference for the assessment of learning disabilities, Baltimore: Brooks.

Sandoval, J., & Echandia, A. (1994). Behavior assessment system for children. Journal of School Psychology, 32, 419-425.

Speece, D. L., & D. H. Cooper (1991). Retreat, Regroup, or Advance? An Agenda for Empirical Classification Research in Learning Disabilities. In L. V. Feagans, E. J. Short, and L. J. Meltzer (Eds.), Subtypes of learning disabilities: Theoretical perspectives and research. Hillsdale, NJ: Erlbaum.

Speece, D. L. (1990). Methodological issues in cluster analysis: How clusters become real. In H. L. Swanson & B. Keogh (Eds.) Learning Disabilities: Theoretical and Research Issues. Hillsdale, NJ: Erlbaum.

Swanson, H. L., & Keogh, B. (Eds.) (1990). Learning Disabilities: Theoretical & Research Issues. Hillsdale, NJ: Erlbaum.

Watson, B. U., & Goldgar, D. E. (1988). Evaluation of a typology of reading disability. Journal of Clinical and Experimental Neuropsychology, 10, 432-450.

Williams, D. L., Gridley, B. E., Fitzhugh-Bell, K. (1992). Cluster analysis of children and adolescents with brain damage and learning disabilities using neuropsychological, psychoeducational, and sociobehavioral variables. Journal of Learning Disabilities, 25, 290-299.

Witt, J. (1994). Rating the ratings of raters: A critique of the Behavior Assessment System for Children. Child Assessment News, 4, 10-11.

Table 1

BASC teacher rating scales and descriptions

| Scale | Description |
|---|---|
| 1. Aggression | Tendency to act in a hostile manner (either verbal or physical) that is threatening to others |
| 2. Hyperactivity | Tendency to be overly active, rush through work or activities, and act without thinking |
| 3. Conduct Problems | Tendency to engage in antisocial and rule-breaking behavior, including destroying property |
| 4. Anxiety | Tendency to be nervous, fearful, or worried about real or imagined problems |
| 5. Depression | Feelings of unhappiness, sadness, and stress that may result in an inability to carry out everyday activities (neurovegetative symptoms) or may bring on thoughts of suicide |
| 6. Somatization | Tendency to be overly sensitive to and complain about relatively minor physical problems and discomforts |
| 7. Attention Problems | Tendency to be easily distracted and unable to concentrate more than momentarily |
| 8. Learning Problems | Presence of academic difficulties, particularly in understanding or completing schoolwork |
| 9. Atypicality | Tendency to behave in ways that are immature, considered "odd," or commonly associated with psychosis (such as experiencing visual or auditory hallucinations) |
| 10. Withdrawal | Tendency to evade others to avoid social contact |
| 11. Adaptability | Ability to adapt readily to changes in the environment |
| 12. Leadership | Skills associated with accomplishing academic, social, or community goals, including, in particular, the ability to work well with others |
| 13. Social Skills | Skills necessary for interacting successfully with peers and adults in home, school, and community settings |
| 14. Study Skills | Skills conducive to strong academic performance, including organizational skills and good study habits |

Adapted from Reynolds and Kamphaus (1992) with permission.

Table 2

BASC teacher rating scales internal consistency coefficients for scales and composites for ages 6 through 11

| Composite or Scale | Number of Items | Ages 6-7 | Ages 8-11 |
|---|---|---|---|
| Externalizing Problems | | .93 | .95 |
| Aggression | 14 | .93 | .95 |
| Hyperactivity | 13 | .92 | .93 |
| Conduct Problems | 10 | .62 | .77 |
| Internalizing Problems | | .90 | .91 |
| Anxiety | 8 | .76 | .79 |
| Depression | 10 | .83 | .87 |
| Somatization | 8 | .78 | .77 |
| School Problems | | .93 | .95 |
| Attention Problems | 8 | .89 | .93 |
| Learning Problems | 9 | .84 | .90 |
| Atypicality | 14 | .84 | .84 |
| Withdrawal | 8 | .80 | .79 |
| Adaptive Skills | | .96 | .97 |
| Adaptability | 6 | .74 | .83 |
| Leadership | 9 | .90 | .89 |
| Social Skills | 12 | .93 | .92 |
| Study Skills | 12 | .92 | .93 |
| | 141 | | |

Table 3

Correlations among the 14 scales (N = 1228)

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Aggression | 1.00 | | | | | | | | | | | | | |
| 2. | Hyperactivity | .82 | 1.00 | | | | | | | | | | | | |
| 3 | Conduct Problems | .77 | .68 | 1.00 | | | | | | | | | | | |
| 4 | Anxiety | .35 | .37 | .30 | 1.00 | | | | | | | | | | |
| 5. | Depression | .67 | .58 | .57 | .64 | 1.00 | | | | | | | | | |
| 6. | Somatization | .26 | .27 | .26 | .44 | .46 | 1.00 | | | | | | | | |
| 7 | Attention Problems | .58 | .72 | .57 | .41 | .54 | .21 | 1.00 | | | | | | | |
| 8 | Learning Problems | .44 | .59 | .49 | .40 | .43 | .23 | .82 | 1.00 | | | | | | |
| 9 | Atypicality | .57 | .65 | .59 | .47 | .61 | .27 | .65 | .55 | 1.00 | | | | | |
| 10 | Withdrawal | .27 | .26 | .35 | .48 | .58 | .28 | .45 | .43 | .49 | 1.00 | | | | |
| 11 | Adaptability | -.61 | -.61 | -.56 | -.43 | -.65 | -.26 | -.74 | -.59 | -.60 | -.51 | 1.00 | | | |
| 12 | Leadership | -.23 | -.29 | -.32 | -.21 | -.29 | -.07 | -.63 | -.58 | -.34 | -.41 | .60 | 1.00 | | |
| 13 | Social Skills | -.33 | -.34 | -.37 | -.09 | -.28 | -.04 | -.60 | -.45 | -.35 | -.39 | .64 | .78 | 1.00 | |
| 14 | Study Skills | -.43 | -.53 | -.46 | -.25 | -.40 | -.12 | -.82 | -.73 | -.47 | -.40 | .73 | .80 | .78 | 1.00 |

Table 4

Correlations between corresponding structure r's

| | Half-Sample Pair | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1st LDF | -.99 | .99 | .99 |
| 2nd LDF | .74 | .83 | .98 |
| 3rd LDF | -.84 | -.08 | .75 |

Table 5

Hit rates and index values for cross-typology clustering

| Cluster Match | Half-Sample Pair 1 | 2 | 3 |
|---|---|---|---|
| 1 | 100.0 (75.8) | 100.0 (97.6) | 100.0 (64.9) |
| 2 | 55.6 (63.8) | 88.6 (99.0) | 49.6 (100.0) |
| 3 | 68.7 (70.3) | 96.8 (98.6) | 88.0 (70.3) |
| 4 | 74.6 (92.6) | 91.5 (96.0) | 51.4 (100.0) |
| 5 | 100.0 (93.5) | 100.0 (87.9) | - |
| Total | 76.0 (74.0) | 95.4 (97.0) | 74.3 (77.7) |

Note. Two hit rates for each cluster match for each pair are given; one for the classification rule based on one half-sample that is applied to the other half-sample, and the one in parentheses for the opposite.

Ta
De ve information for the seven clusters (N = 1228)

| | Clusters | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1. Well Adapted | 2. Average | 3. Disruptive Behavior Disorder | 4. Learning Disorder | 5. Physical Complaints/ Worry | 6. Severe Psychopath-ology | 7. Mildly Disruptive |
| | n = 417 (34%) | n = 228 (19%) | n = 103 (8%) | n = 149 (12%) | n = 134 (11%) | n = 52 (4%) | n = 145 (12%) |
| | 39% male | 48% male | 78% male | 60% male | 40% male | 67% male | 70% male |
| | 4.2% Dx | 4.8% Dx | 19.6% Dx | 13.4% Dx | 5.8% Dx | 17.6% Dx | 8.5% Dx |
| Scale | | | | | | | |
| Aggression | 44.0 | 54.3 | **67.8** | 49.2 | 49.6 | **69.6** | 57.7 |
| Hyperactivity | 43.5 | 44.6 | **66.3** | 52.3 | 49.6 | **69.9** | 57.5 |
| Conduct | 45.3 | 45.6 | **65.4** | 51.3 | 47.6 | **71.3** | 52.7 |
| Anxiety | 45.9 | 44.8 | 54.4 | 52.3 | 58.4 | **70.6** | 47.3 |
| Depression | 44.5 | 44.6 | **61.0** | 51.8 | 55.3 | **76.4** | 50.3 |
| Somatization | 46.6 | 45.2 | 53.6 | 48.9 | **65.0** | **61.8** | 47.4 |
| Attention | 41.0 | 49.2 | **63.4** | **60.8** | 49.2 | **68.3** | 52.5 |
| Learning | 42.3 | 49.3 | **62.9** | **61.1** | 50.6 | **65.6** | 49.7 |
| Atypicality | 45.1 | 46.2 | 58.9 | 55.1 | 49.4 | **80.8** | 50.3 |
| Withdrawal | 45.1 | 47.2 | 55.0 | **59.4** | 53.8 | **69.4** | 45.2 |
| Adaptability | 58.9 | 50.1 | **37.3** | 41.1 | 48.2 | **32.5** | 46.6 |
| Leadership | 59.0 | 43.4 | 41.8 | **38.8** | 50.0 | 41.6 | 50.7 |
| Social Skills | 58.8 | 44.3 | 41.2 | **39.7** | 51.9 | 42.3 | 47.4 |
| Study Skills | 60.0 | 46.4 | **38.0** | **38.4** | 51.1 | **38.5** | 47.9 |

Note. Values that deviate at least one standard deviation from the T mean of 50 are in boldface. Dx : previously diagnosed with a behavioral, emotional, or academic problem.

Table 7

Structure r's for the three leading LDFs

| Scale | LDF 1 | LDF 2 | LDF 3 |
|---|---|---|---|
| Aggression | **.51** | **.42** | **-.59** |
| Hyperactivity | **.49** | .24 | **-.44** |
| Conduct | **.41** | .26 | -.31 |
| Anxiety | .25 | .29 | **.45** |
| Depression | **.44** | **.40** | .22 |
| Somatization | .17 | .33 | **.54** |
| Attention | **.59** | -.30 | -.02 |
| Learning | **.41** | -.23 | .11 |
| Atypicality | **.45** | .26 | .14 |
| Withdrawal | .28 | .00 | **.45** |
| Adaptability | **-.53** | .20 | -.02 |
| Leadership | -.34 | **.63** | -.20 |
| Social Skills | -.33 | **.56** | .02 |
| Study Skills | **-.52** | **.64** | -.01 |

Note. The dominating structure r's for each LDF are in boldface.

Table 8

L-0-0 linear classification results

| | | Predicted Cluster | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| | 1 | 370 (88.7) | 44 | 0 | 0 | 2 | 0 | 1 | 417 |
| | 2 | 0 | 228 (100) | 0 | 0 | 0 | 0 | 0 | 228 |
| | 3 | 0 | 0 | 92 (89.3) | 4 | 1 | 3 | 3 | 103 |
| Actual Cluster | 4 | 0 | 19 | 3 | 117 (78.5) | 4 | 0 | 6 | 149 |
| | 5 | 4 | 19 | 1 | 1 | 107 (80.0) | 0 | 2 | 134 |
| | 6 | 0 | 0 | 7 | 1 | 2 | 42 (80.8) | 0 | 52 |
| | 7 | 2 | 27 | 1 | 2 | 2 | 0 | 111 (76.6) | 145 |
| | | 376 | 337 | 104 | 125 | 118 | 45 | 123 | 1228 |

Note. Cluster hit rates are given in parentheses. The across-cluster hit rate is 1067/1228 ≐ 86.9%.

Table 9

Linear classification function weights

| Scale | Cluster 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Aggression | 1.13 | 1.11 | 1.66 | 1.26 | 1.30 | 1.55 | 1.46 |
| Hyperactivity | -0.29 | -.27 | -.06 | -.23 | -.22 | -.09 | -.09 |
| Conduct | 0.74 | .70 | .89 | .74 | .65 | 1.02 | .71 |
| Anxiety | 0.27 | .24 | .30 | .29 | .39 | .47 | .27 |
| Depression | 0.76 | .78 | .92 | .79 | .83 | 1.21 | .78 |
| Somatization | 0.51 | .50 | .57 | .54 | .81 | .64 | .51 |
| Attention | 2.80 | 2.88 | 3.00 | 3.07 | 2.86 | 3.05 | 2.92 |
| Learning | 0.52 | .56 | .72 | .69 | .58 | .67 | .54 |
| Atypicality | 0.69 | .68 | .78 | .77 | .70 | 1.33 | .68 |
| Withdrawal | 0.85 | .84 | .98 | 1.04 | .94 | 1.08 | .86 |
| Adaptability | 2.35 | 2.24 | 2.07 | 2.08 | 2.20 | 2.15 | 2.16 |
| Leadership | 0.50 | .32 | .35 | .35 | .39 | .35 | .45 |
| Social Skills | -0.08 | -.20 | -.17 | -.21 | -.11 | -.16 | -.18 |
| Study Skills | 2.70 | 2.50 | 2.37 | 2.41 | 2.54 | 2.37 | 2.50 |
| (Constant) | -337.80 | -306.50 | -392.64 | -345.12 | -357.45 | -475.56 | -343.05 |

Figure 1

Plot in LDF space of matched clusters for the second pair of half-samples

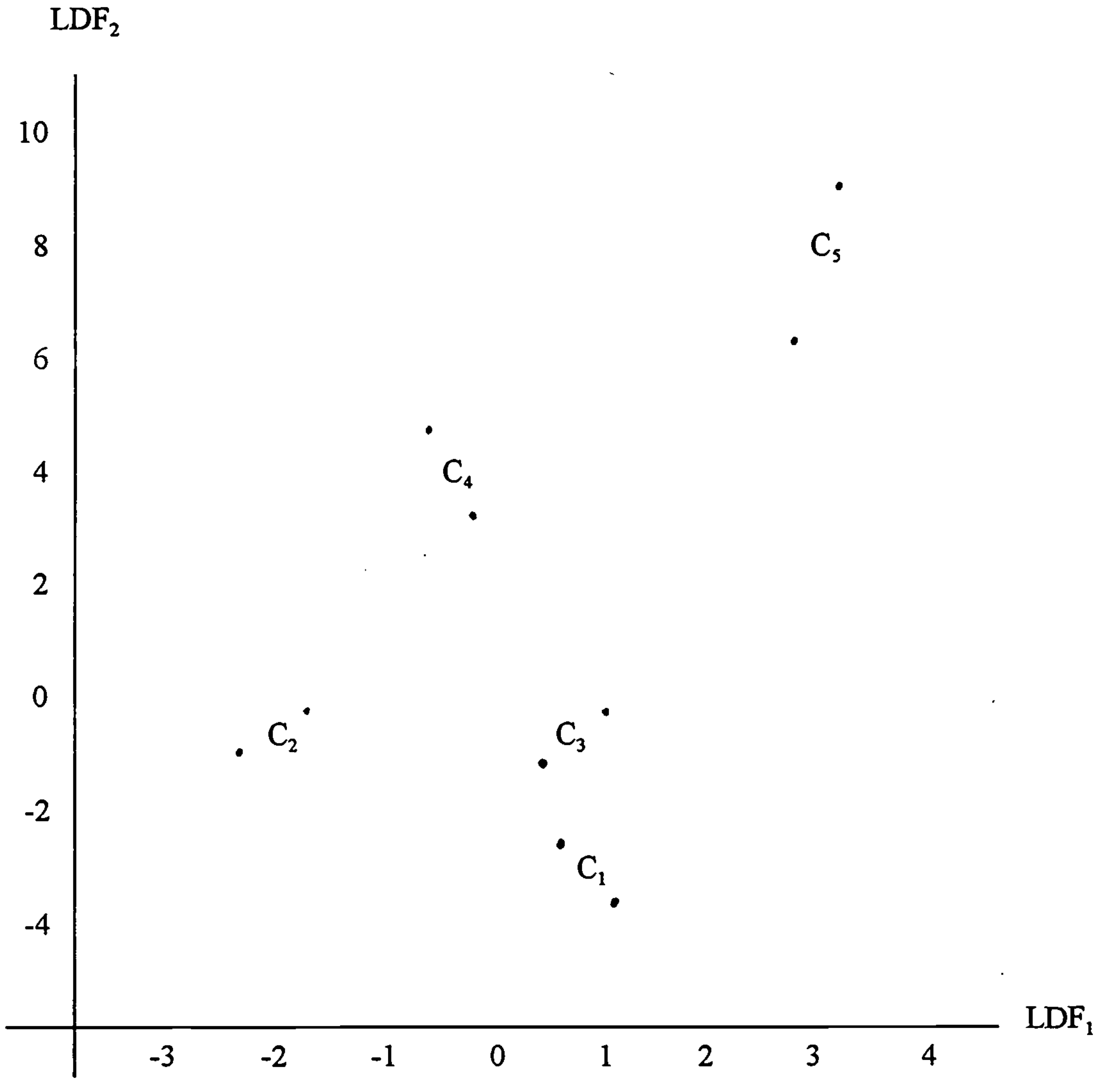
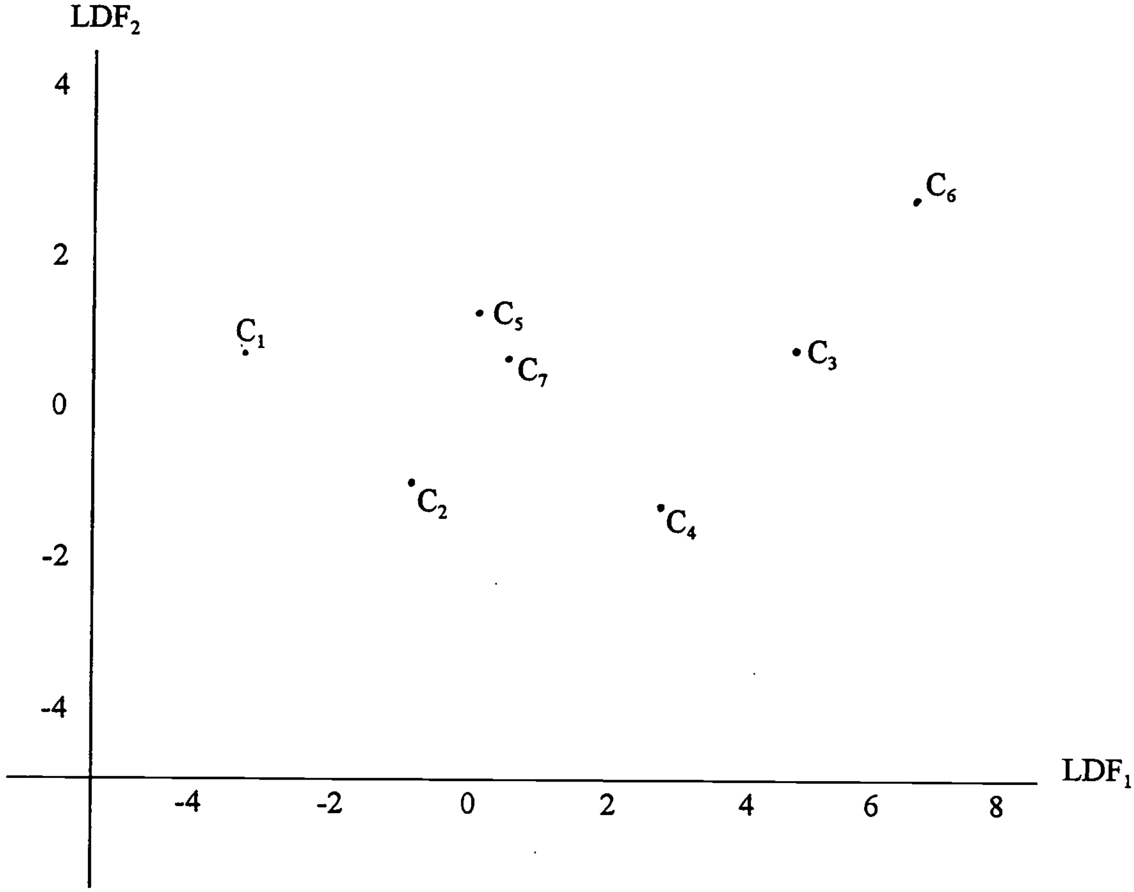


Figure 2

Plot of cluster centroids in LDF space



AERA April 8-12, 1996

TM025859

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

ERIC®

# REPRODUCTION RELEASE

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| Title: Behavioral Clustering of School Children | |
| Author(s): C. J Huberty, C. DiStefano, & R. W. Kamphaus | |
| Corporate Source: University of Georgia | Publication Date: presented at AERA 4/96 |

## II. REPRODUCTION RELEASE:

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

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

| | Sample sticker to be affixed to document | Sample sticker to be affixed to document | |
|---|---|---|---|
| [X] **Check here** Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction | "PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)." | "PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)." | [ ] **or here** Permitting reproduction in other than paper copy. |
| | **Level 1** | **Level 2** | |

## Sign Here, Please

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

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

| | |
|---|---|
| Signature: C J Huberty | Position: Professor and Head |
| Printed Name: Carl J Huberty | Organization: University of Georgia |
| Address: 325 Aderhold Hall Athens, GA 30602 | Telephone Number: (706 ) 542- 4110 |
| | Date: 5/15/96 |

CUA

**THE CATHOLIC UNIVERSITY OF AMERICA**
*Department of Education, O'Boyle Hall*
*Washington, DC 20064*
*202 319-5120*

February 27, 1996

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a written copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (23)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:

AERA 1996/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://tikkun.ed.asu.edu/aera/). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

---

[1]If you are an AERA chair or discussant, please save this form for future use.