Journal of Abnormal Child Psychology, Vol. 8, No. 4, 1980, pp. 441–470

A Typology of Child Behavior Profile Patterns: Distribution and Correlates for Disturbed Children Aged 6—16[1]

Craig Edelbrock[2]
Boys Town Center for the Study of Youth Development

Thomas M. Achenbach
Department of Psychiatry, University of Vermont

This article reports the construction of a typology of behavior problem patterns, as scored on the Child Behavior Profile. Hierarchical cluster analysis was used to identify reliable profile patterns that characterize clinically referred boys and girls aged 6-11 and 12-16. Procedures for classifying the profiles of individual children were developed and good agreement was obtained between classifications based on ratings by mothers and a clinician. The distribution of profile patterns was determined for each age and sex group. Children classified by different patterns were found to differ significantly in age, race, school performance, and amount and quality of participation in activities and social relationships. Issues in the application of cluster analysis to behavioral data were analyzed and various options for the classification of individual cases were compared.

The lack of an objective and reliable taxonomy of children's behavior disorders has been a major handicap to training, research, and communication pertaining to child psychopathology. Classification systems such as that contained in the second edition of the Diagnostic and Statistical Manual of the American Psychiatric Association ("DSM-II"; American Psychiatric Association, 1968) and

Manuscript received in final form April 18, 1980.
[1] The authors are indebted to Faith Dickerson, Claire Horowitz, and Brian McLaughlin for their assistance in this research. The authors would like to thank Mary Pat Roy and Dorothy Runte for their help in the preparation of the manuscript.
[2] Address all correspondence to Craig Edelbrock, Research Division, Boys Town Center for the Study of Youth Development, Boys Town, Nebraska 68010.

0091-0627/80/1200-0441 $03.00/0

the one proposed by the Group for the Advancement of Psychiatry (GAP, 1966) consist of narrative descriptions of disorders that are not operationally defined and yield mediocre agreement among diagnosticians (Beitchman, Dielman, Landis, Benson, & Kemp, 1978; Freeman, 1971; Tarter, Templer, & Hardy, 1975). It is also questionable whether these systems are useful in differentiating among disturbed children with respect to etiology, prognosis, and differential response to treatment. The DSM-III (American Psychiatric Association, 1980) employs more explicit diagnostic criteria, but preliminary studies indicate no better reliability than the DSM-II for children's disorders (Mattison, Cantwell, Russell, & Will, 1979; Mezzich & Mezzich, Note 1).

The need for a useful, well-differentiated, and reliable taxonomy of child psychopathology has stimulated numerous efforts to derive syndromes empirically from behavior problems reported by parents, teachers, and clinicians. Although several syndromes have been replicated in diverse samples of disturbed children (cf. Achenbach & Edelbrock, 1978), these syndromes have seldom been translated into procedures for discriminating among individual children. Moreover, when this has been attempted, it has generally been limited to the identification of children having a single syndrome such as hyperkinesis (Conners, 1969) or has been based on two global syndromes variously labeled Internalizing versus Externalizing (Achenbach, 1966), Personality Problem versus Conduct Problem (Peterson, 1961), or Inhibition versus Aggression (Miller, 1967). Most efforts to validate empirically derived taxonomies have involved determining the correlates of these global syndromes (for a review of such efforts see Quay, 1979).

To provide an empirically based system for making more differentiated discriminations among disturbed children, we have developed the Child Behavior Profile (CBP). The CBP consists of behavior problem and social competence scales scored from the Child Behavior Checklist, which comprises 118 behavior problems and 20 social competence items designed to be reported by parents or parent surrogates. To reflect age and sex differences in the prevalence and patterning of behaviors, separate editions of the CBP have been developed and standardized for each sex at ages 4-5, 6-11, and 12-16.

The CBP includes three a priori social competence scales designed to reflect school performance and involvement in activities and social relationships, plus behavior problem scales that have been derived separately for each edition through factor analysis of checklists filled out by parents of children referred for mental health services. Nine reliable behavior problem factors were obtained for boys aged 6-11, boys aged 12-16, and girls aged 6-11, whereas eight reliable factors were obtained for girls aged 12-16 (Achenbach, 1978; Achenbach & Edelbrock, 1979). Second-order factor analyses have shown that the first-order, narrow-band behavior problem scales for each edition form two broad-band groupings, which have been labeled Internalizing and Externalizing.

Norms have been constructed for the scales of each edition of the CBP by computing normalized *T* scores from checklists filled out by 1,300 randomly selected parents of normal children.

The CBP can be used to assess children in terms of the empirically derived narrow-band and broad-band syndromes of behavior problems. In order to develop a *classification* system for discriminating among children, however, it is necessary to translate scores on these syndromes into categories of individuals. One way to derive differentiated categories of disturbed children is to cluster-analyze their profiles of scores on the narrow-band syndromes. Cluster analysis thus serves as a "multivariate heuristic" for identifying *patterns of reported problems that characterize groups of disturbed children.* In this article, such empirically derived patterns will be called *profile types.*

The identification of profile types is not an end in itself, since the value of the resulting classification depends on its relations to other variables. Owing to the lack of a definitive taxonomy of child psychopathology, empirically derived classifications have not generally been validated against established diagnostic categories. Instead, other correlates of profile types have been sought to provide a more complete picture of how groups of disturbed children differ. In previous cluster analyses of children's behavior problems, for example, groups of children representing different profile types were found to differ in IQ, academic achievement, parent education, race, sex, age of onset of the disorder, and mental health referrals (Eisenberg, Gersten, Langner, McCarthy, & Simcha-Fagan, 1976; Prior, Boulton, Gajazago, & Perry, 1975; Spivack, Swift, & Prewitt, 1971).

The value of these taxonomies for discriminating among disturbed children, however, is likely to be limited by the subject samples employed: Eisenberg et al. and Spivack et al. used primarily normal, nonreferred children, while Prior et al. used only psychotic children. Moreover, correlates of the clusters derived in these studies may primarily reflect differences in demographic characteristics. For example, Spivack et al. used cluster analysis to identify groups of children manifesting different patterns of teacher-reported classroom behavior. These groups were subsequently shown to differ on such variables as academic achievement and IQ. They also differed markedly in race and level of parent education. The differences in academic achievement and IQ may therefore reflect the demographic differences rather than being independent correlates of the cluster types. Similar questions arise in the taxonomy of welfare children developed by Eisenberg et al. (1976).

The purpose of the present study was to develop a classification system based on the CBP through which groups of disturbed children having similar patterns of reported behavior problems can be identified. The specific goals were (a) to identify reliable profile types for each age and sex group; (b) to develop procedures for classifying the CBPs of individual children according

to the profile types; (c) to assess the interrater reliability of the classification system; (d) to determine the distribution of the profile types among clinically referred children; and (e) to identify correlates of the profile types, including differences in demographic variables and social competence scores.

METHOD

Subjects

Data on 2,683 children were analyzed in this study, including 1,050 boys aged 6-11, 633 boys aged 12-16, 500 girls aged 6-11, and 500 girls aged 12-16. CBP scores for each child were computed from checklists filled out by the child's parent at intake into one of 30 East Coast mental heatlh facilities. These facilities included child guidance clinics, community mental health centers, health maintenance organizations, and private practices. The racial composition of the entire sample was 78.4% white and 21.6% black. Other racial groups were excluded from the analyses because they constituted less than 1% of the children seen in these mental health facilities. Socioeconomic status (SES) was assessed with Hollingshead's (Note 2) 7-step scale for breadwinner's occupation; it averaged 3.7 (SD = 1.7) for the total sample.

Cluster Analysis of Profiles

Several problems arise when using cluster analysis in behavioral research. A major problem is simply the choice of which clustering algorithm and measure of profile similarity to use. More than 100 clustering algorithms have been developed and different methods can produce different results when applied to the same data. Likewise, various measures of profile similarity are available for use in clustering, including Euclidean distance and other distance metrics, the product-moment correlation, intraclass correlations, and information statistics (cf. Anderberg, 1973; Everitt, 1974). Several Monte Carlo Studies have recently been addressed to the evaluation of clustering algorithms (Blashfield, 1976; Edelbrock, 1979; Edelbrock & McLaughlin, 1980). Although the hierarchical methods have been found to be more accurate and reliable than nonhierarchical methods, no one algorithm seems uniformly best for all applications. Instead, subsets of algorithms that are likely to perform well will with certain types of data have been identified.

A second problem in using cluster analysis is artifactual findings. Clustering algorithms can create homogeneous groups even when applied to random data. Some clusters may therefore be statistical artifacts rather than representing reliable profile types. One way to reduce such artifactual findings is to replicate

the cluster analysis on different samples and retain only those profile types that are obtained in two or more samples. Clustering algorithms are also very sensitive to the standardization and scaling of the profile data to which they are applied. Differences in scale parameters alone may contribute to the creation of artifactual clusters, for example.

A third problem when using hierarchical clustering methods is the choice of level in the hierarchy from which to draw clusters. Hierarchical methods do not produce a discrete number of clusters, but rather a hierarchical arrangement of individuals and groups. Low levels in the hierarchy comprise small, homogeneous, and well-differentiated clusters; at higher levels, these small clusters are combined into larger, less differentiated, and more heterogeneous clusters. The problem is whether to classify children according to many specific, low-level profile types or a few global, higher level profile types. Since the cluster analyses reveal the hierarchical relations among subjects and groups, however, it is possible to construct hierarchical taxonomies wherein subjects can be simultaneously classified according to both low-level and high-level profile types.

A final problem involves the classification of new subjects. An important step in constructing any taxonomy is the development of methods for classifying new subjects who were not included in the sample from which the taxonomy was derived. One procedure for classifying new subjects is to compute the degree to which each child's profile of scores is similar to the profile types obtained in the cluster analysis. Each child can then be classified according to the profile type that his or her profile most resembles.

Due to the diversity of etiologies, environmental contingencies, and ways of expressing psychopathology, as well as differences among raters, the profiles of some children will not resemble any of the empirically derived profile types. Should such children be classified? The value of any classification system is related to its *coverage,* that is, the proportion of individuals it can classify. Yet, in some situations, it may be better to classify fewer individuals into more reliable and homogeneous groups than to try to classify everybody (cf. Edelbrock, 1979). This complex issue of coverage will be discussed in more detail later.

Identifying Profile Types. To identify profile types, separate hierarchical cluster analyses were performed on the narrow-band problem scales for each age and sex group. Initial cluster analyses utilizing both the behavior problem and social competence scales indicated that the social competence scales contributed little to the identification of differentiated profile types. One reason for this is that the social competence scales are negatively correlated with the behavior problem scales. This "built-in" pattern of high behavior problem and low competence scores interferes with the detection of distinctive profile patterns. In developing our cluster-based classification system, we therefore used only the behavior problem portion of the CBP.

In the cluster analyses, standard scores based on samples of clinically referred children were used, rather than standard scores based on normal children.

In many applications of the CBP, it is valuable to use scores standardized on samples of normal children, but, according to these norms, clinically referred children tend to have high scores on all scales. In addition, scale means and standard deviations for normals are quite different from those for referred children. Scoring referred children according to scores standardized on normals can produce a "built-in" pattern of scores that can bias the cluster analyses. In order to overcome such biases and construct taxonomies that differentiate maximally among clinically referred children, it is necessary to standardize scores within clinical samples. It is important to emphasize that raw scale sums are calculated in the same manner whether the sums are transformed into standard scores based on clinical or normal samples. Although their distribution parameters differ, scores standardized on referred children are almost perfectly colinear with scores standardized on normal children.

A centroid clustering algorithm was employed. This clustering method was developed by Sokal and Michener (1958) and has also been called the "weighted pair group method" (Sokal & Rohlf, 1962). The measure of profile similarity was the intraclass correlation (ICC), which is obtained by calculating the proportion of variance shared by two profiles (Bartko, 1976; Haggard, 1958). It reflects similarity in both the *pattern* and the *elevation* of profiles. Although most previous applications of cluster analysis to behavioral data have used Euclidean distance as the similarity measure (cf. Blashfield, 1976), the centroid algorithm using the ICC was found to be among the most accurate of several methods in Monte Carlo studies using data similar to those provided by the CBP (Edelbrock & McLaughlin, 1980). Previous cluster analyses of the CBP data used Q-correlations obtained by calculating the product-moment correlation between two subjects' scores across the scales of the profile (Edelbrock & Achenbach, Note 3). This reflects similarity in profile patterns but not in elevation. Although the majority of profile types found via Q-correlations were replicated using ICCs, the ICCs resulted in greater differentiation among profile types.

The centroid clustering algorithm proceeds by first calculating an ICC between each possible pair of profiles in the sample. Next, the two profiles having the highest ICC are located and combined into a cluster. These two profiles are then replaced by their *centroid,* which is the *profile* created by calculating the average of the two subjects' scores on each scale. On the next step, the centroid of the newly created cluster is treated just like the profile of a single subject and the ICCs between all possible pairs of profiles are recomputed. In each cycle, the two profiles having the highest ICC are located, combined into a cluster, and replaced by their centroid. Whenever an individual profile or cluster is combined with another cluster, the centroid is recomputed by calculating the average of the scores of all members of the cluster on each of the scales in the profile. As cycles proceed, larger and larger clusters are formed and combined in a hierarchical manner. The result is a hierarchical clustering of all profiles, in which groups

of subjects having similar profiles and the hierarchical relationships among these groups can be identified.

For each age and sex group, profile types that replicated across two samples of 250 profiles each were retained. These samples were randomly drawn without replacement from the total pool of subjects of each age and sex. The criterion for replication was a significant ($p < .05$) ICC between a cluster centroid obtained in the other sample. A significant ICC between cluster centroids obtained in different samples indicated that the pattern and elevation of scores on the CBP was shared by a group of children in each sample, and hence that the profile type was reliable.

Classification of Individual Profiles

To classify individual children according to the reliable profile types, the ICCs between each child's CBP and the centroids of the profile types were calculated. Each child was then classified according to the profile type with which his or her CBP correlated most highly. A minimum ICC required for classification could be specified so that children whose CBPs were not very similar to any of the profile types could be left unclassified. By changing this minimum cutoff ICC, the coverage of the classification can be varied. That is, the use of a high cutoff point results in a small proportion of subjects being classified into homogeneous groups that represent relatively "pure types." Conversely, the use of a low cutoff point results in the classification of a higher proportion of subjects into larger and more heterogeneous groups. For some purposes, such as research on a focused treatment, a high cutoff point may be desirable since it produces very homogeneous groups for study. For other purposes, such as epidemiological surveys, a lower cutoff point may be desirable since it increases the proportion of children that can be classified.

We did not try to classify children having a total score of 25 or less on the behavior problem portion of the checklist because a change of only 1 or 2 points on certain scales can change their pattern of scores on the profile. Hence, their profile patterns are likely to be unreliable for purposes of classification. We also excluded children with total scores of 100 or more on the checklist because they tend to have extreme scores on all scales. The result is that their profile patterns are determined primarily by differences in the standardization of the scales, rather than by the relative concentration of behavior problems.

RESULTS

Using the hierarchical clustering procedure, six reliable profile types were obtained for boys aged 6-11 and boys aged 12-16, whereas seven reliable profile

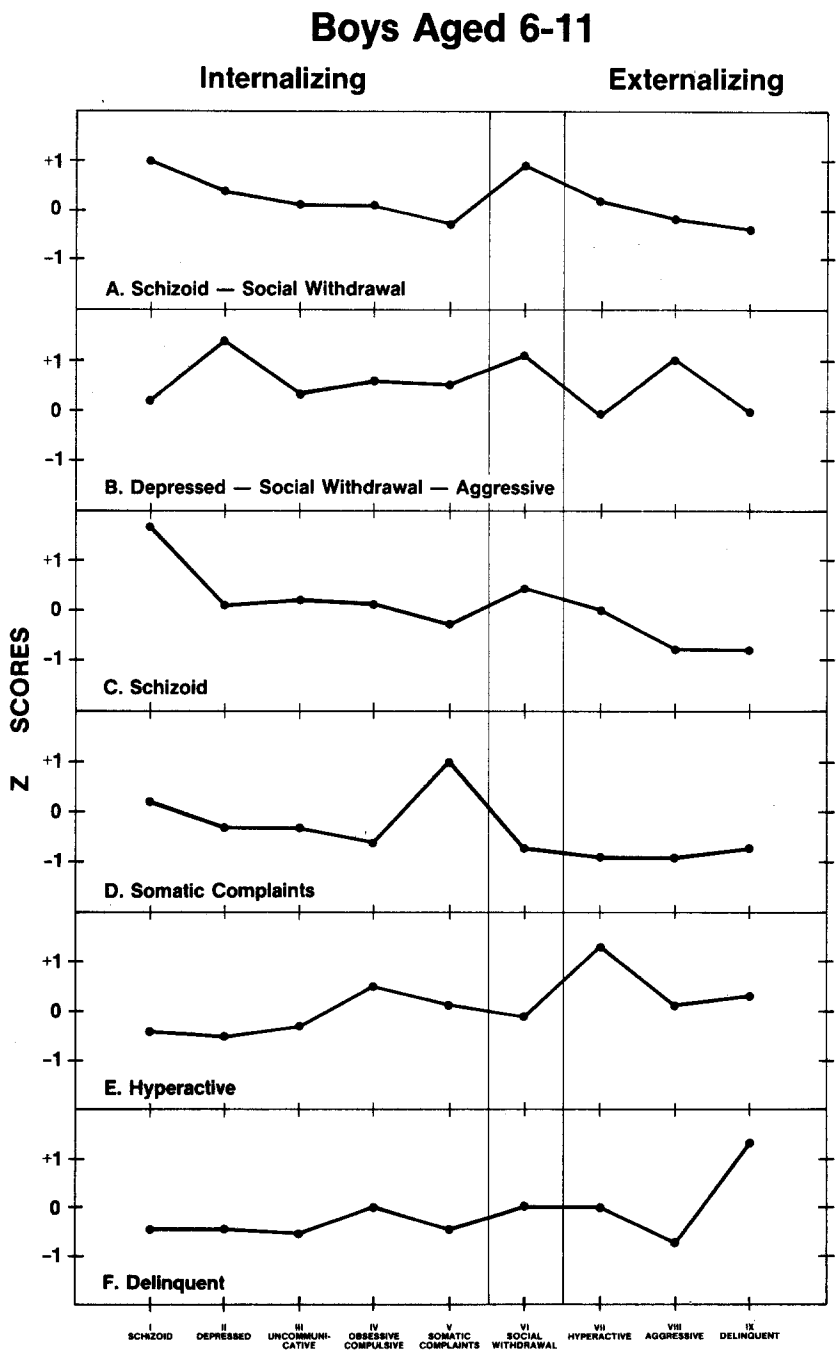

Fig. 1. Child Behavior Profile types identified for boys aged 6-11.

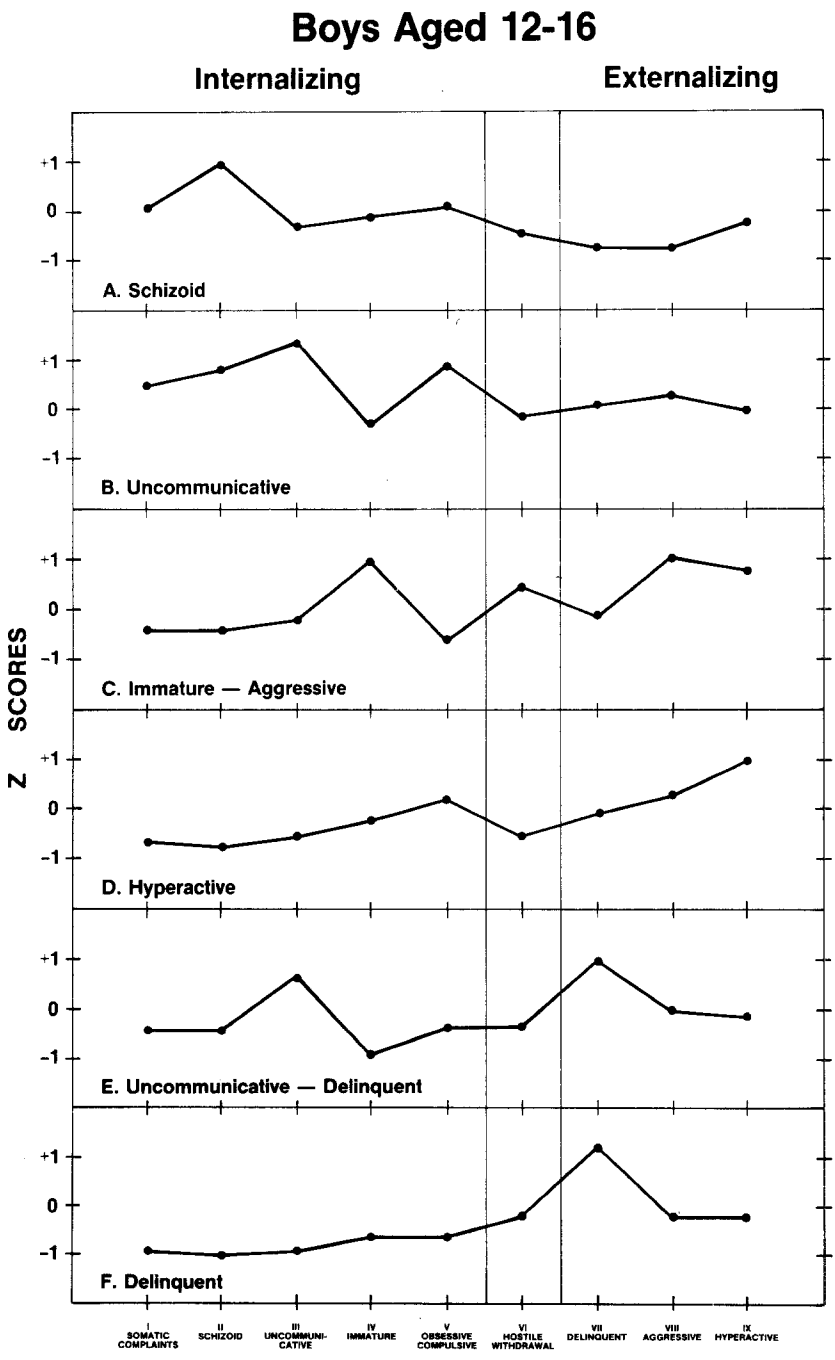

Fig. 2. Child Behavior Profile types identified for boys aged 12-16.

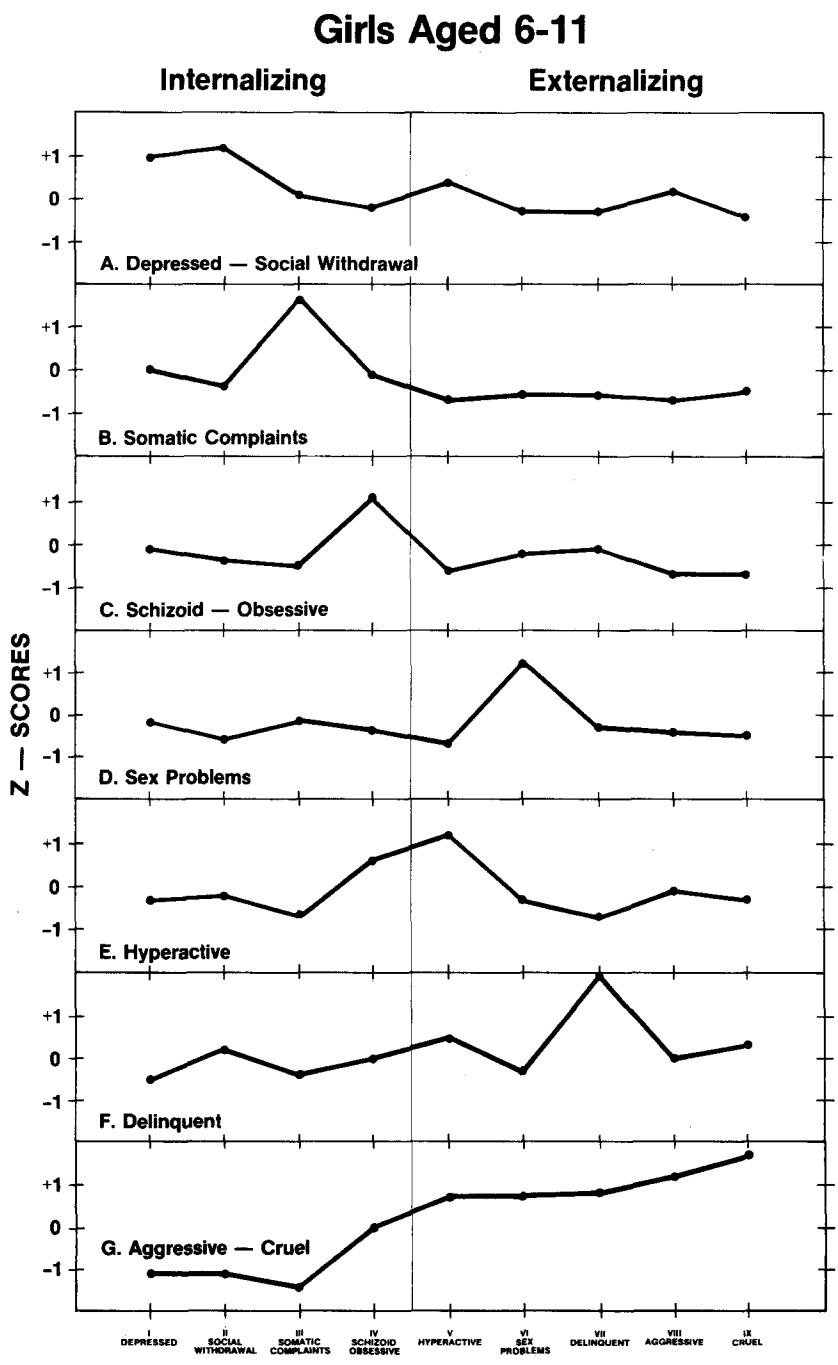

Fig. 3. Child Behavior Profile types identified for girls aged 6-11.

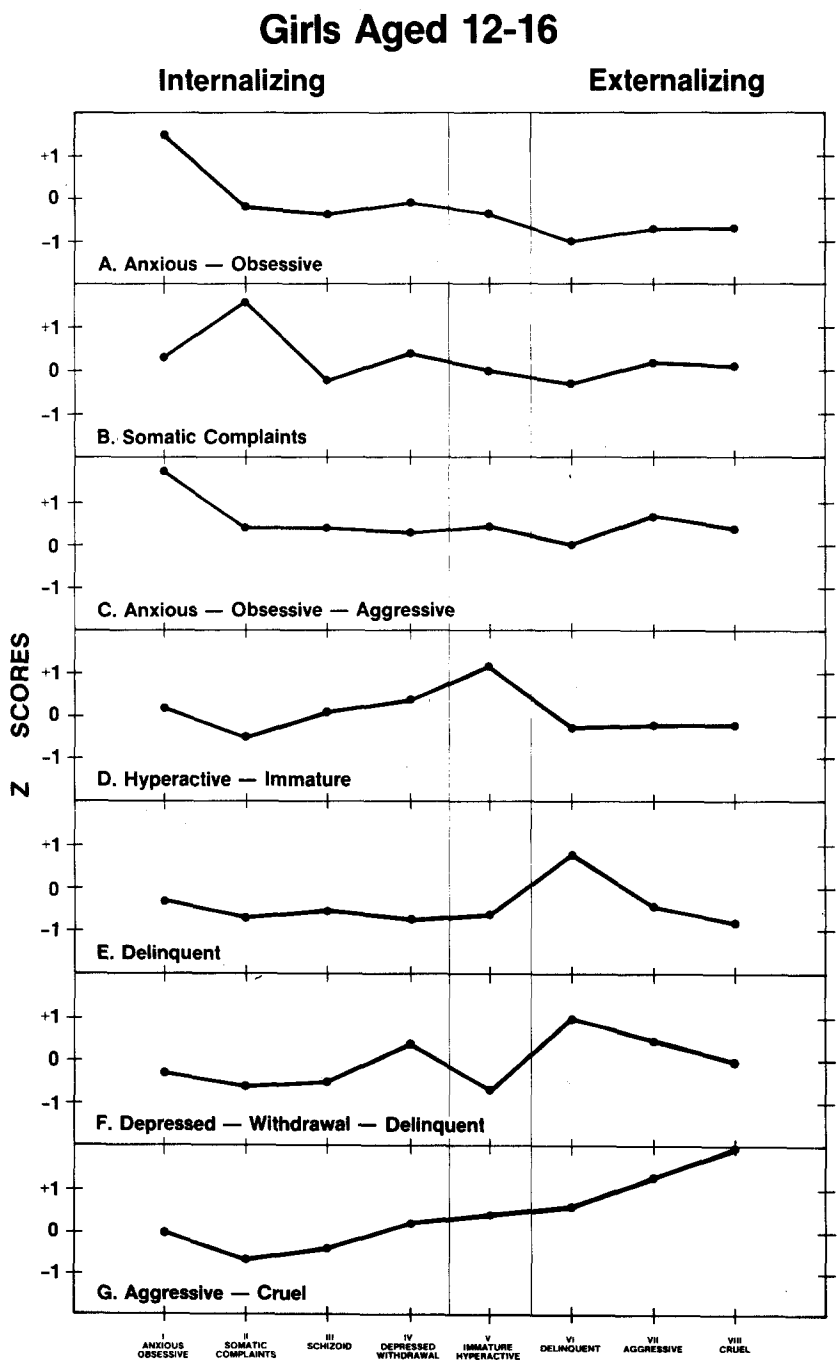

Fig. 4. Child Behavior Profile types identified for girls aged 12-16.

types were obtained for girls 6-11 and girls 12-16. Each of these profile types met the criterion of a significant ($p < .05$) ICC between a cluster centroid obtained in one sample and a cluster centroid obtained in the other sample. The names and profile patterns of the profile types are shown in Figures 1-4. These names simply reflect the distinguishing high points of the profile types, but it is important to note that each profile type is defined by its entire pattern and elevation of scores on the CBP, rather than by its high points alone.

The hierarchical relations among profile types were found to replicate across samples and are shown in Figures 5-8. (Following the initial cluster analyses, 65 girls of each age range were excluded from subsequent analyses because of missing social competence data. This is why $N = 435$ in Figures 7 and 8.) For boys aged 6-11, the Schizoid-Social Withdrawal cluster and the Depressed-Social Withdrawal-Aggressive cluster combined with each other to form a homogeneous cluster whose centroid had high scores on the Schizoid, Depressed, and Social Withdrawal scales. For older boys, the Delinquent cluster and the Uncommunicative-Delinquent cluster combined to form a cluster with a high score on the Delinquent scale. Two such hierarchical relationships were

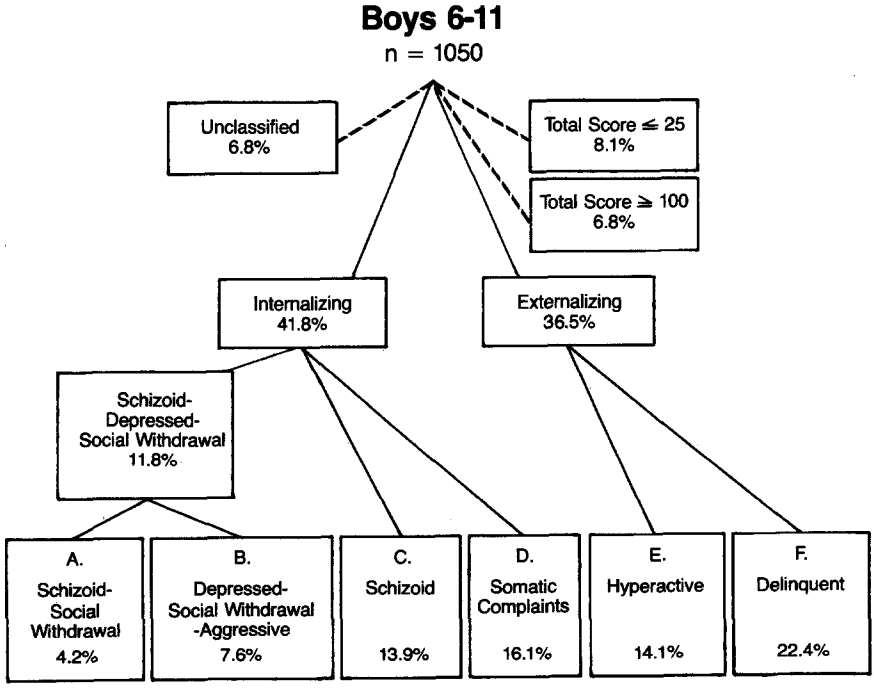

Fig. 5. A hierarchical taxonomy of Child Behavior Profile patterns for boys aged 6-11.

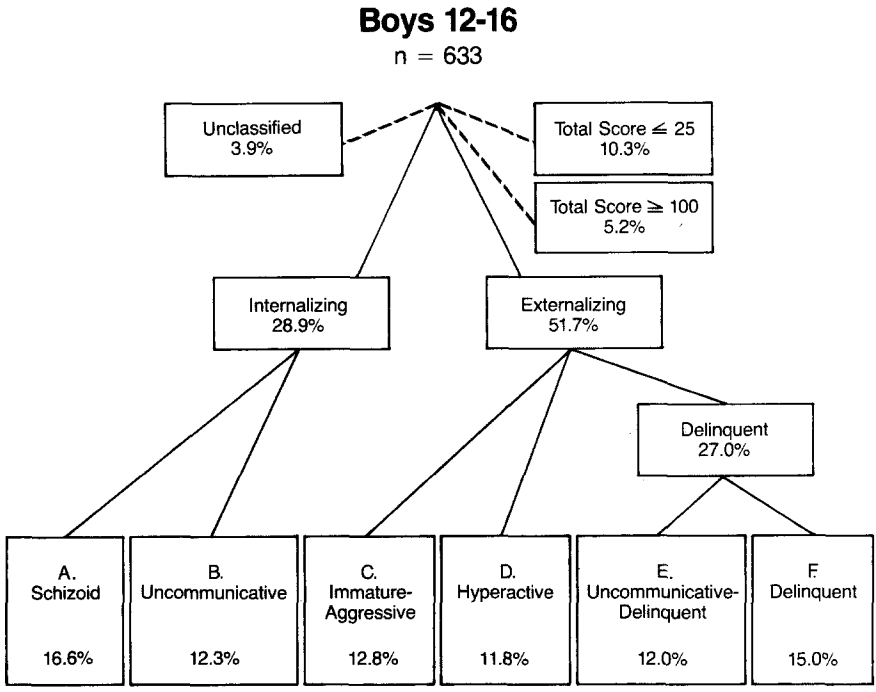

Fig. 6. A hierarchical taxonomy of Child Behavior Profile patterns for boys aged 12-16.

found for older girls. The Anxious-Obsessive cluster and the Somatic Complaints cluster combined to form a cluster with high scores on both the Anxious-Obsessive scale and the Somatic Complaints scale; and the Delinquent cluster combined with the Depressed Withdrawal-Delinquent cluster to form a cluster with a high score on the Delinquent scale.

In addition, at higher levels in the hierarchies, certain profile types were combined into a cluster that represented an Internalizing pattern of scores, characterized by high scores on the Internalizing scales and low scores on the Externalizing scales. Other profile types were combined into an Externalizing cluster, with high scores on the Externalizing scales and low scores on the Internalizing scales. For girls aged 6-11 and girls aged 12-16, mixed profile types were found that did not reliably combine with either the Internalizing or the Externalizing clusters. These profile types are labeled as "Mixed" in Figures 7 and 8. For younger girls, the Mixed profile type had a high score on the Sex Problems scale, which is an Externalizing scale, but also had high scores on other Internalizing and Externalizing scales. For older girls, the Mixed profile type had a high score on the Immature-Hyperactive scale, which has moderate loadings on both the second-order Internalizing and Externalizing factors.

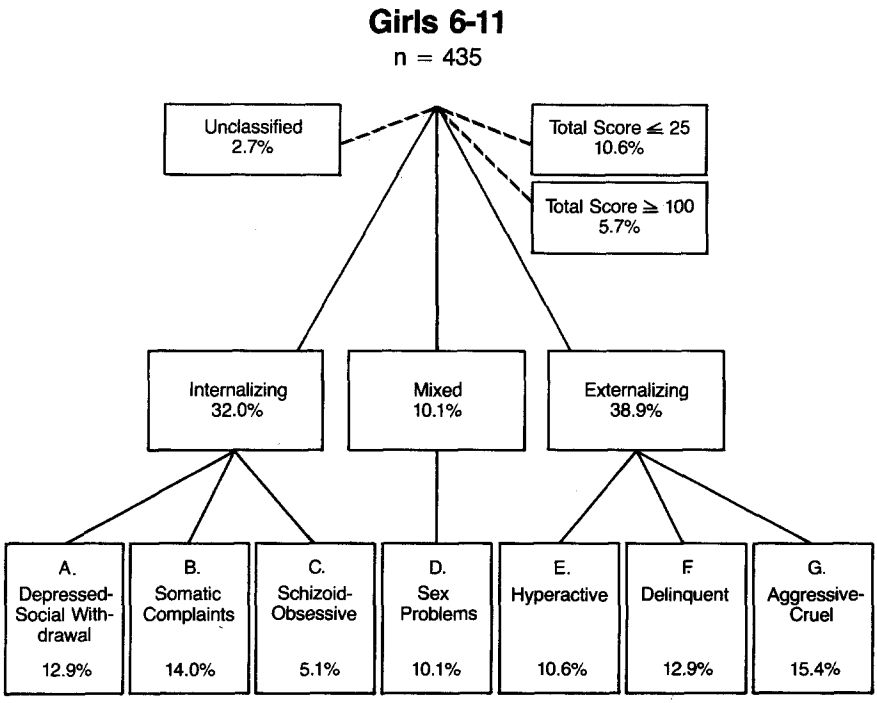

Fig. 7. A hierarchical taxonomy of Child Behavior Profile patterns for girls aged 6-11.

Distribution of Profile Types

To determine the distribution and correlates of the profile types among clinically referred children, the CBPs of all children with data on both the behavior problem and social competence portions of the checklist were classified according to their similarity to the profile types obtained for their age and sex group.

In order to classify as many children as possible, a minimum cutoff ICC of .00 was used. Hence, in order to be classified, a child's CBP had to have an ICC greater than .00 with the centroid of at least one profile type. To determine if such a low cutoff ICC produced groups that were similar to the profile types, the z-transformed average of the ICC of each group member's CBP with the centroid of the profile type was calculated. This is a measure of the degree to which the groups resulting from the classification of individual profiles were similar to the profile types obtained in the cluster analyses. The average of these ICCs for the 28 profile types was .47, which corresponds to a product-moment r of .80 (z = 1.10, $p < .01$). This indicates that although the groups contained some

members who were minimally similar to their profile type, most group members were highly similar to their profile type.

Figures 5-8 show the percent of cases classified according to each profile type and the Internalizing, Externalizing, and Mixed profile patterns. The "unclassified" groups in the figure are cases whose CBP had no ICC greater than .00 with the centroid of any profile type. The figures also show the percent of cases who were not candidates for classification because their total scores on the checklist were either $\leqslant 25$ or $\geqslant 100$.

Correlates of Profile Types

We wished to determine if classifications according to the profile types are related to age, race, or SES, and how groups of disturbed children having similar patterns of reported problems differ in adaptive competencies. Because differences in adaptive competencies may be confounded with demographic characteristics, we controlled for demographic differences when assessing the social competence correlates of the profile types.

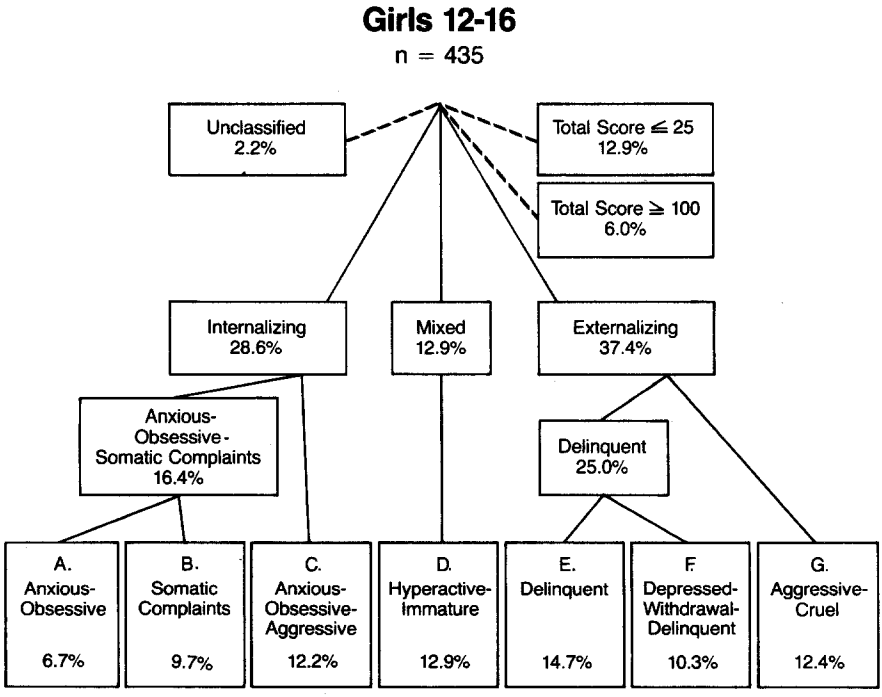

Fig. 8. A hierarchical taxonomy of Child Behavior Profile patterns for girls aged 12-16.

The social competence variables do not necessarily represent independent validity criteria, because both the behavior problem and social competence data were provided by the same respondents. To some extent, there are also a priori relations between the behavior problem and social competence scales because such scales tap similar behaviors in opposite ways. For example, the Hyperactivity syndrome, which includes items such as "poor schoolwork," "can't concentrate," and "confused," would be expected to correlate negatively with ratings of school performance. Nevertheless, social competence correlates extend the meaning of the profile types by explicating the relations between patterns of reported problems on one hand and adaptative competencies on the other. Validation of these taxonomies involves determining differences among profile types on independent criteria related to etiology, prognosis, treatment response, cognitive functioning, personality characteristics, etc.

To determine the demographic correlates of the profile types, one-way analyses of variance (ANOVAs) using profile type as the classification variable were performed separately for each age and sex group. The dependent variables for these ANOVAs were age and SES. Differences in racial distribution were first tested by an overall χ^2 analysis across all profile types, followed by z tests (Bruning & Kintz, 1968, pp. 197-198) to determine significant deviations from the proportion of blacks and whites expected on the basis of the pooled sample. To control for demographic differences, one-way analyses of covariance (ANCOVAs) were used to determine the social competence correlates of the profile types. These ANCOVAs were performed separately for each age and sex group, using profile type as the classification variable and scores on the three social competence scales as the dependent variables. Where significant age, race, or SES differences among profile types were found, the significant demographic variables were included as covariates in the ANCOVAs. Differences among the profile types on the social competence scales were thus adjusted for differences in demographic characteristics. For the ANOVAs and ANCOVAs, significant differences were further analyzed with the modified least significant difference test (Winer, 1971, pp. 199-201).

One-way ANOVAs were run for each age and sex group to determine if the three unclassified groups differed significantly from the pool of children classified according to the profile types. The dependent variables were age and SES. Differences between the unclassified groups and classified children on race were determined by χ^2 analyses, followed by specific z-tests as described previously. For the social competence scales, one-way ANCOVAs were used to compare the unclassified groups with classified children, using significant demographic variables as covariates.

Tables I-IV show mean values of the dependent variables for the profile types and the unclassified groups, as well as the significant differences. Unadjusted mean values are shown for the demographic variables, whereas mean values adjusted for any significant demographic covariates are shown for the

Table I. Demographic and Social Competence Correlates of the Profile Types for 1,050 Boys Aged 6-11[a]

Dependent variables	Comparisons among profile types			Profile types[b]						Internalizing vs. Externalizing[d]	Unclassified groups[e]		
				Internalizing		Externalizing							
	Significance	Effect size	LSD[c]	A	B	C	D	E	F		ICC <0	Score ≥100	Score ≤25
N	–	–	–	44	80	146	169	148	235	–	72	71	85
Demographic characteristics													
Age	n.s.	.07	–	8.5	8.9	8.3	8.6	8.4	8.5	n.s.	8.5	8.2	8.6
Race[f]	$p < .001$	.13	–	75.6	93.2[+]	82.1	74.5	68.7[+]	74.6	Int > Ext**	78.1	85.4	75.0
SES	n.s.	.06	–	3.5	3.8	3.8	4.0	3.9	3.9	n.s.	3.6	4.0	3.9
Social competence													
Activities	n.s.	.08	–	5.7	6.7	6.3	6.6	6.0	6.3	n.s.	6.6	6.4	6.3
Social	$p < .001$	.24	.6	4.0	4.5	4.7	5.5	4.6	4.6	Int > Ext*	5.4*	3.9*	6.4*
School	$p < .001$	.26	.4	3.3	3.6	3.6	3.9	2.9	3.4	Int > Ext***	3.9	3.2	4.0*

[a] Age (years), race (% white), SES (Hollingshead's 7-step scale). Mean values for social competence scales based on raw scale sums – range: Activities (0-12), Social (0-12), School (0-6).

[b] Profile types are (A) Schizoid-Social Withdrawal, (B) Depressed-Social Withdrawal-Aggressive, (C) Schizoid, (D) Somatic Complaints, (E) Hyperactive, (F) Delinquent.

[c] Modified least significant difference (Winer, 1971) for comparisons among profile types. Not applicable to differences in race. Mean differences that meet or exceed the LSD are significant ($p < .05$).

[d] For comparisons between Internalizing and Externalizing groups, * $p < .05$, ** $p < .01$, *** $p < .001$.

[e] Asterisk indicates groups that differ significantly ($p < .05$) from children classified according to profile types (pooled).

[f] For race, [+] indicates the profile types in which the proportion of whites differed significantly ($p < .05$) from the proportion expected on the basis of the pooled sample.

Table II. Demographic and Social Competence Correlates of the Profile Types for 633 Boys Aged 12-16[a]

Dependent variables	Comparison among profile types			Profile types[b]						Internalizing vs. Externalizing[d]	Unclassified groups[e]		
				Internalizing		Externalizing							
	Significance	Effect size	LSD[c]	A	B	C	D	E	F		ICC <0	Score ≥ 100	Score ≤ 25
N	–	–	–	105	78	81	75	76	95	–	25	33	65
Demographic characteristics													
Age	$p < .001$	.28	.6	13.2	13.8	13.0	13.6	14.1	13.8	n.s.	13.7	13.5	13.7
Race	n.s.	.14	–	72.0	81.4	87.8	71.9	83.1	80.0	n.s.	88.0	71.0	81.5
SES	n.s.	.08	–	3.6	3.5	3.2	3.5	3.5	3.5	n.s.	2.8	4.6*	3.5
Social competence													
Activities	$p < .05$	.16	.8	6.8	5.8	6.3	6.7	6.2	6.3	n.s.	7.5*	5.6	6.4
Social	$p < .001$	.26	.9	5.5	4.4	4.3	5.5	5.2	5.3	n.s.	5.2	4.5	6.2*
School	$p < .01$	.19	.4	3.5	3.2	2.8	3.0	3.0	3.1	Int > Ext***	4.0*	3.0	4.0*

[a] Age (years), race (% white), SES (Hollingshead's 7-step scale). Mean values for social competence scales based on raw scale sums – range: Activities (0-12), Social (0-12), School (0-6).

[b] Profile types are (A) Schizoid, (B) Uncommunicative, (C) Immature-Aggressive, (D) Hyperactive, (E) Uncommunicative-Delinquent, and (F) Delinquent.

[c-e] See notes for Table I.

Table III. Demographic and Social Competence Correlates of the Profile Types for 435 Girls Aged 6-11[a]

Dependent variables	Comparisons among profile types			Profile types[b]							Internalizing vs. Externalizing[d]	Unclassified groups[e]		
				Internalizing			Mixed	Externalizing						
	Significance	Effect size	LSD[c]	A	B	C	D	E	F	G		ICC < 0	Score ⩾ 100	Score ⩽ 25
N	–	–	–	56	61	22	44	46	56	67	–	12	25	46
Demographic characteristics														
Age	$p < .05$	.21	.8	9.1	8.5	8.5	8.5	8.3	8.5	8.0	Int > Ext**	8.2	8.3	9.1
Race[f]	$p < .05$	.22	–	78.8	79.2	94.7	76.7	95.0+	64.7+	70.7	n.s.	90.9	69.6	80.6
SES	n.s.	.12	–	4.0	3.8	4.2	3.8	4.0	4.2	4.4	n.s.	3.2	5.2*	4.1
Social competence														
Activities	$p < .05$	.16	.8	6.0	6.7	6.4	7.0	6.2	5.9	6.2	n.s.	8.0*	5.9	7.3
Social	$p < .001$	.26	.9	4.3	5.1	5.4	5.3	4.5	4.4	4.0	Int > Ext*	5.4	4.3	6.1*
School	$p < .001$	.35	.5	3.7	4.1	4.6	4.3	3.1	3.7	3.5	Int > Ext**	3.8	3.4	4.5*

[a] Age (years), race (% white), SES (Hollingshead's 7-step scale). Mean values for social competence scales based on raw scale sums -- range: Activities (0-12), Social (0-12), School (0-6).

[b] Profile types are (A) Depressed-Social Withdrawal, (B) Somatic Complaints, (C) Schizoid-Obsessive, (D) Sex Problems, (E) Hyperactive, (F) Delinquent, (G) Aggressive-Cruel.

[c-f] See notes for Table I.

Table IV. Demographic and Social Competence Correlates of the Profile Types for 435 Girls Aged 12-16[a]

Dependent variables	Comparisons among profile types			Profile types[b]							Internalizing vs. Externalizing[d]	Unclassified groups[e]		
				Internalizing			Mixed	Externalizing						
	Significance	Effect size	LSD[c]	A	B	C	D	E	F	G		ICC < 0	Score ⩾100	Score ⩽25
N	–	–	–	29	42	56	56	69	45	54	–	22	26	36
Demographic characteristics														
Age	$p < .05$	.24	.7	13.6	14.3	14.1	13.8	14.6	14.3	13.8	n.s.	14.3	14.1	14.0
Race	n.s.	.14	–	82.8	79.5	92.0	73.6	79.7	79.5	79.6	n.s.	80.9	76.9	69.4
SES	n.s.	.18	–	3.4	3.8	3.7	3.5	2.9	3.2	3.6	n.s.	4.1	4.6*	3.8
Social competence														
Activities	n.s.	.14	–	6.5	6.1	6.1	6.4	6.1	6.1	5.4	n.s.	5.2*	5.8	7.4*
Social	$p < .001$	.30	1.0	4.9	5.2	5.5	4.7	5.8	5.1	4.1	n.s.	5.6	3.8*	6.6*
School	$p < .001$	.25	.6	4.0	3.7	4.3	3.8	3.7	3.5	3.3	Int > Ext**	4.3	3.3	4.8*

[a]Age (years), race (% white), SES (Hollingshead's 7-step scale). Mean values for social competence scales based on raw scale sums – range: Activities (0-12), Social (0-12), School (0-6).

[b]Profile types are (A) Anxious-Obsessive, (B) Somatic Complaints, (C) Anxious-Obsessive-Aggressive, (D) Hyperactive-Immature, (E) Delinquent, (F) Depressed-Withdrawal-Delinquent, (G) Aggressive-Cruel.

[c-e]See notes for Table I.

social competence variables. These tables also indicate the significance and direction of differences between the Internalizing and Externalizing groups on the dependent variables with the social competence variables adjusted for significant demographic covariates as described previously.

Effect Size. Previous studies in this area have only reported the levels of statistical significance for their findings (Eisenberg et al., 1976; Spivak et al., 1971). Due to the power afforded by large sample sizes, a high level of statistical significance can be achieved when only a small proportion of variance in the dependent variable has been explained. Conversely, the low statistical power afforded by small sample sizes may preclude significant results even if effects are large. We therefore computed the sizes of effects that classification by profile type had on the dependent variables shown in Tables I-IV.

For the continuous variables, effect size was measured by Cohen's f (Cohen, 1977, pp. 338-339). Each f value is an index of the proportion of total (pooled) variance in the dependent variable accounted for by the independent variable of profile type. Cohen designates f values of .10, .25, and .40 as representing small, medium, and large effects, respectively. For race, effect size was measured by the *contingency coefficient* (c). One reason for choosing f and c is that they have comparable scales for assessing effect sizes in the analyses of continuous variables (f) and frequency data (c).

Demographic Correlates. As shown in Tables I-IV, some profile types were found to differ significantly on the demographic variables, but the effect sizes were small and the findings were not consistent across age and sex groups. Within three age and sex groups, there were significant age differences among profile types, but these differences were not consistently due to a particular profile pattern or the Internalizing/Externalizing dichotomy. For example, among girls aged 6-11, the Depressed-Social Withdrawal group was significantly older than other groups, whereas among girls aged 12-16, the Delinquent group was significantly older than other groups.

Significant racial differences were found among profile types only for boys and girls aged 6-11, and these were also inconsistent. Among boys, for example, blacks were significantly overrepresented in the Hyperactive group and underrepresented in the Schizoid group, whereas among girls, balcks were underrepresented in the Hyperactive group and overrepresented in the Delinquent group. No significant SES differences between profile types or Internalizing/Externalizing groups were found within any of the age and sex groups. The unclassified groups did not differ significantly from classified children on age or race. The groups having total behavior problem scores $\geqslant 100$ were of significantly lower SES than classified children, among all age and sex groups except boys aged 6-11.

Social Competence Correlates. Significant differences among profile types on the social competence scales showed larger effect sizes and were more consistent across age and sex groups. On the School scale, for example, highly signi-

ficant differences among profile types were found within all age and sex groups. Among boys and girls aged 6-11, the Hyperactive groups obtained significantly lower School scores than some other groups, whereas among girls of both age ranges, the Aggressive-Cruel groups also scored significantly lower than other groups. For girls aged 6-11, the Hyperactive and Aggressive-Cruel groups were significantly lower than other groups, but not significantly different from each other on the School scale.

Significant differences were also found within all age and sex groups on the Social scale. Among girls of both age ranges, the Aggressive-Cruel groups scored significantly lower than other groups, and among boys and girls aged 6-11, Externalizers scored significantly lower than Internalizers. For the Activity scale, boys aged 12-16 classified as Uncommunicative scored significantly lower than other groups, as did girls aged 6-11 in the Delinquent and Depressed-Social Withdrawal categories. Internalizers and Externalizers did not differ significantly on the Activity scale in any of the samples.

As shown in Tables I-IV, several significant differences were detected between the unclassified groups and classified children. Overall, for those differences that reached significance, children with total behavior problem scores $\geqslant 100$ obtained lower social competence ratings, whereas children with total behavior problem scores $\leqslant 25$ obtained higher social competence ratings than classified children. For those children whose profile patterns did not resemble any profile types (ICC < 0), significant differences reflected social competence scores that were higher than those obtained for classified children. An exception to this was that among girls aged 12-16, this unclassified group obtained significantly lower activities ratings than classified girls.

The Issue of Coverage

In our taxonomy, the proportion of children classified is determined by the minimum cutoff ICC one chooses. A low cutoff point results in the classification of a high proportion of children. A high cutoff point results in the classification of fewer children, but each group is more homogeneous with respect to profile pattern. Yet a high cutoff point also reduces the generalizability of findings to the clinical population as a whole, since only children who are highly similar to the profile types will be classified. The correlates reported in Tables I-IV were based on a cutoff point of .00, which results in a high level of coverage and findings that can be generalized to a large proportion of clinically referred children. There are several considerations besides generalizability, however, in choosing a cutoff point for classification.

Proportion and Similarity of Subjects Classified. It is important to consider what proportion of children can be classified using different cutoff points and how similar the members of the resulting groups will be to the profile types. To

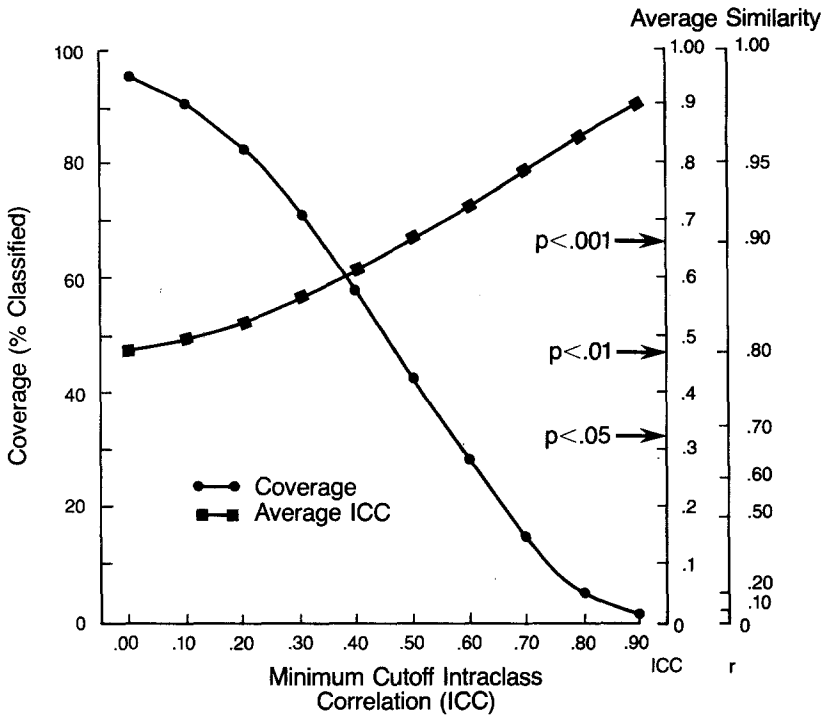

Fig. 9. The effect of varying the cutoff point for classification on the coverage and homogeneity of the resulting groups.

assess these relationships, several classifications were made for each age and sex group, using cutoff points ranging from .00 to .90, at intervals of .10. We excluded children with total behavior problem scores of $\leqslant 25$ or $\geqslant 100$, for reasons stated earlier. The results are shown in Figure 9, which depicts the relations between the cutoff point used, the percentage of children classified (left axis), the *z*-transformed average of the ICCs between the group members and their profile types (right axis), and the statistical significance of these average ICCs. Since they were almost identical, results for all four age and sex groups are averaged in the curves shown. To offer some perspective on the magnitude of the average ICCs, the relation between the ICC and the more familiar product-moment *r* is shown on the right axis.

For any cutoff point used, the resulting level of coverage, the average similarity of the group members to the profile types, and the corresponding level of significance of the average similarity can be determined. A cutoff point of .00, for example, classified 93% of the children; their average similarity to the profile types was ICC = .47 (r = .80, $p < .01$). A cutoff point of .50, by

contrast, classified 42% of the children, with an average similarity of ICC = .67 (r = .91, $p < .001$). These curves indicate that too high a cutoff point (e.g., > .70) may result in groups too small to permit statistical comparisons, although this obviously depends on the sample size. The members of the resulting groups, however, would represent relatively "pure types," which may be valuable for certain research purposes. Because researchers do not have unlimited samples of clinically referred children from which to select, it may be necessary to use a lower cutoff point to obtain groups large enough for study. This will result in groups whose members are less similar to their group's centroid. Yet it is important to note that the average similarity of the members to their group's centroid will remain high, even though some members' similarities to the profile types may be as low as the cutoff point.

Detection of Differences Among Profile Types. It is also important to determine how different cutoff points affect detection of differences among profile types. Statistical power in this situation is a complex function of the cutoff point used, the sizes of the resulting groups, and the sizes of the differences being analyzed. Our large samples provided high statistical power, particularly for detecting medium-to-large effects. However, an analysis of the statistical power of tests performed at cutoff points ranging from .00 to .90 showed that the ability to detect significant differences due to a *small* effect size increased rapdily as the cutoff point was raised from .00 to .20, and peaked between .20 and .30. With sample sizes like ours, power remains at a plateau between cutoff points of .30 and .70 and decreases above .70, due to the decline in the size of the groups. This indicates that the use of different cutoff points substantially affects the ability to detect significant correlates of the profile types.

To determine if higher cutoff points would result in the discovery of different correlates than those reported in Tables I–IV, the demographic and social competence correlates of the profile types were analyzed for each age and sex group using cutoff points of .30 and .50. These cutoff points cover the range in which statistical power is maximized and result in the classification of approximately 70% and 40% of the samples, respectively. Increasing the cutoff point had the effect of only moderately accentuating the differences among the profile types. Effect sizes, levels of statistical significance, and homogeneities of the groups were improved but the differences reported in Tables I-IV were unchanged. That is, the same groups remained significantly higher or lower than others on the dependent variables. For these data, cutoff points of .30 and .50 thus failed to reveal any different correlates from these of a cutoff point of .00.

Optimizing Coverage and Power. The power analyses discussed above and the coverage function shown in Figure 9 suggest that a cutoff point in the range of .20 to .30 would maximize the ability to detect significant differences among groups but still classify a high proportion of children. Hence, for most applications this would be the recommended range for the cutoff point. A cutoff point

of .20 may sound low when compared to the scale of the product-moment r, but as a measure of shared variance between profiles, the ICC is more comparable to r^2 than to r. As shown on the right axis of Figure 9, the range of .20 to .30 for the ICC would correspond roughly to product-moment rs of .50 to .65. It is also important to stress that this recommended range is only a guideline for choosing a cutoff point. The optimal cutoff point for a particular study will depend on the sample size, the actual distribution of children according to the profile types, and the sizes of the differences being analyzed.

Reliability of Classification

To determine the reliability of classifying children according to the profile types, we computed the agreement between the classification of children's CBPs obtained from their mothers and a clinician. One reason for considering the agreement between mother and clinician is that, of all possible informants, these are the most likely to be available in clinical and research settings. Mothers' CBPs were scored from checklists they completed at intake of the child into an outpatient mental health clinic. The clinician's CBPs were scored from checklists she completed on the basis of information obtained from each mother in an intake interview, firsthand observation and interview of the child, and information from fathers and teachers. Children with total scores of $\leqslant 25$ or $\geqslant 100$ on *either* the mother's or the clinician's checklist were excluded.

Table V shows the agreement obtained within each age and sex group for the profile types and the more global Internalizing and Externalizing dichotomy. Because there is no perfect index of interrater reliability for categorical classification, three measures of agreement are reported: percent agreement, the statistic kappa (Cohen, 1960), and the ratio of obtained kappa to the maximum possible kappa. Percent agreement is the most intuitively appealing index, but it can be misleading when children are not equally distributed among the profile types. That is, differences in group size increase the probability of chance agreements.

Table V. Agreement Between Profile Classifications Derived from Mothers' and Clinician's Ratings

		Profile types[a]			Internalizing/Externalizing		
Group	*N*	% Agreement	K	K/Kmax	% Agreement	K	K/Kmax
Boys 6-11	20	70%	.583	.811	95%	.900	1.00
Boys 12-16	23	70%	.555	.897	74%	.303	1.00
Girls 6-11	4	75%	.667	1.00	75%	.500	1.00
Girls 12-16	20	80%	.746	1.00	90%	.759	1.00
Average of all groups		74%	.638	.927	83%	.616	1.00

[a] K = kappa (Cohen, 1960). K/Kmax = ratio of kappa to maximum possible kappa.

The statistic kappa corrects for chance agreements, in that a kappa value of zero corresponds to the level of agreement expected by chance rather than corresponding to no agreements. A kappa value of 1.00 corresponds to perfect agreement between raters but can only be achieved when the distribution of children according to the profile types is the same for both raters, which was not the case for our data. One convention has therefore been to calculate the maximum possible kappa given the marginal distributions of the classifications and to express agreement as the ratio of the obtained kappa to the maximum possible kappa. This ratio can in turn be misleading since a value of 1.00 does not necessarily correspond to no disagreements between raters, but to the maximum possible level of agreement given the marginal distributions.

The results shown in Table V indicate good agreement between mothers and the clinician in classifying children's CBPs. Across all groups, agreement averaged 83% for the Internalizing/Externalizing dichotomy and 74% for the lower level profile types. Within each age and sex group, the kappa values indicate substantially higher than chance agreement, and in several instances the maximum possible agreement, given the marginal distributions, was achieved.

DISCUSSION

In order to reflect age and sex differences among disturbed children, we derived separate taxonomies for boys and girls aged 6-11 and 12-16. These taxonomies are similar in that global Internalizing versus Externalizing patterns were found in all four groups. Some lower level profile types, representing more specific patterns of reported problems, were also similar across groups. Profile types reflecting primarily delinquent behavior, for example, were found for all four groups, although more differentiation among the delinquent profiles was obtained for 12- to 16-year-olds than for 6- to 11-year-old (see Figures 5-8). A profile pattern representing primarily hyperative behavior and one respresenting somatic complaints were also found in three out of four groups.

Certain profile types were found for both age groups of one sex, such as the Aggressive-Cruel type for girls and the Schizoid type for boys. Age-specific profile types were also found. Younger children of both sexes exhibited profile patterns showing concentrations of problems in the areas of depression and social withdrawal, with the boys' version including an aggressive component. Other profile types, such as the Sex Problems pattern for younger girls, had no counterparts among other groups.

Although previous studies have not used the same measures or procedures for classifying children, the social competence correlates we found agree with other comparisons of Internalizing (Inhibited, Shy-Anxious, Personality Problem) and Externalizing (Aggressive, Acting-Out, and Conduct Problem) children. Significant differences in social competence ratings favored Internalizers over

Externalizers, particularly in the areas of school performance and social relations. These differences are consistent with previous findings that children classified as Internalizers have better school performance, are more popular with teachers and peers, and have fewer social problems than children classified as Externalizers (Achenbach, 1966; Achenabch & Lewis, 1971; Hafner, Quast, & Shea, 1975; Lorion & Cowen, 1976; Lorion, Cowen, & Caldwell, 1974; Rolf, 1972).

A major advantage of the taxonomies developed here is that they offer a fine-grained differentiation among children in addition to the global Internalizing versus Externalizing dichotomy. This, in turn, permits more precise comparisons among more homogeneous groups of disturbed children. For example, 6- to 11-year-old boys classified as Internalizers obtained higher ratings of school performance than those classified as Externalizers. This was due, however, to the extremely low scores of Externalizers having the Hyperactive profile type. Boys manifesting the Delinquent profile type, which is also an Externalizing pattern, obtained ratings of school performance as high as boys in the Internalizing groups. Focusing only on the Internalizing/Externalizing dichotomy may thus obscure the fact that some differences characterize certain subsets, but not all Internalizers or Externalizers.

Some differences may also be obscured by depending solely on the global Internalizing/Externalizing classification. For instance, among girls aged 12-16, we found no significant differences between Internalizers and Externalizers on the Social scale. However, highly significant differences were found among specific profile types. Among the Externalizing profile types, the Aggressive-Cruel group scored lower than any other group, whereas the Delinquent group scored higher than any other group. In some situations it may therefore be necessary to use the lower level profile types in order to detect significant differences among groups of disturbed children. Ultimately, the choice of the Internalizing/Externalizing dichotomy or the lower level profile types to classify children will depend on the goals of classification, the sample size, the size of the effects involved, and the actual distribution of children according to the profile types.

Demographic variables have often been a factor in cluster analyses of children's behavior. In the taxonomies developed by Eisenberg et al. (1976) and Spivack et al. (1971), for example, cluster types differed markedly in race and SES. Hence, when assessing other correlates of these taxonomies, the relative contributions of differences in behavioral patterns and differences in demographic characteristics could not be separated. In our study, demographic variables had small effects and were inconsistent across age and sex groups. Nevertheless, it was important to control for such differences when assessing the social competence correlates of the profile types. By controlling for such possible confounding effects, we could determine which differences in adaptive competencies were associated with profile patterns per se, rather than reflecting age, race, or SES differences.

The long-term goals of this research are, of course, to predict more than social competencies. Our taxonomies can be used to group disturbed children in order to compare them on variables such as etiology, prognosis, course, and differential response to treatment. To facilitate clinical and research applications, computer- and hand-scored procedures have been developed for scoring the profiles of individual children and classifying them according to the profile types reported here. These classification procedures have been designed to be maximally sensitive to differences among groups. Hence, children can be classified according to the Internalizing/Externalizing dichotomy or more specific profile patterns, depending on the user's needs. The coverage of the classification, which determines the size and homogeneity of the resulting groups and the power of the statistical comparisons, can also be varied by choosing different cutoff points for classification.

In these taxonomies, a cutoff point of zero did not result in the classification of all children. Within each age and sex group, a small proportion of profiles did not resemble any of the profile types and some children had behavior problem scores too high or too low to permit classification according to profile patterns. Although these "unclassified" groups do not represent specific profile patterns, they are empirically defined groups that may have distinctive correlates. Moreover, informed decisions are still required when such "unclassifiable" children are referred for services. These groups, therefore, should not be ignored in future research.

The taxonomies reported here were based on parents' ratings of their children's behavior. Although the agreement between classifications based on parent and clinician ratings was good, it is not necessary to depend on a single source for either classifying children or assessing changes in their behavior. There is much to recommend parents as informants, but other informants are also important, including teachers, clinicians, and trained observers, We are therefore developing companion rating procedures for these different informants. Although their different perspectives, biases, contexts, and effects on children may limit agreement among them, we need to captialize on whatever valid discriminative power each type of informant can add to our knowledge of children's problems and adaptive competencies.

REFERENCE NOTES

1. Mezzich, A. C., & Mezzich, J. E. *Diagnostic reliability of childhood and adolescence behavior disorders.* Paper presented at the meetings of the American Psychological Association, New York, September 1979.
2. Hollingshead, A. B. *Two-factor index of social position.* Unpublished manuscript (available from the author, Department of Sociology, Yale University, New Haven, Connecticut 06520).
3. Edelbrock, C. S., & Achenbach, T. M. *Child Behavior Profile patterns of children referred for clinical services.* Paper presented at the 86th annual meeting of the American Psychological Association, Toronto, Canada, August 1978.

REFERENCES

Achenbach, T. M. The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs,* 1966, *80* (7, Whole No. 615).

Achenbach, T. M. The Child Behavior Profile: I. Boys aged 6-11. *Journal of Consulting and Clinical Psychology,* 1978, *46,* 478-488.

Achenbach, T. M., & Edelbrock, C. S. The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin,* 1978, *85,* 1275-1301.

Achenbach, T. M., & Edelbrock, C. S. The Child Behavior Profile: II. Boys aged 12-16 and girls aged 6-11 and 12-16. *Journal of Consulting and Clinical Psychology,* 1979, *47,* 223-233.

Achenbach, T. M., & Lewis, M. A proposed model for clinical research and its application to encopresis and enuresis. *Journal of the American Academy of Child Psychiatry,* 1971, *10,* 535-554.

American Psychiatric Association. *Diagnostic and Statistical Manual* (2nd ed.). Washington, D.C.: Author, 1968.

American Psychiatric Association. *Diagnostic and Statistical Manual* (3rd ed.). Washington, D.C.: Author, 1980.

Anderberg, M. R. *Cluster analysis for applications.* New York: Academic Press, 1973.

Bartko, J. J. On various intraclass correlation reliability coefficients. *Psychological Bulletin,* 1976, *83,* 762-765.

Beitchman, J. H., Dielman, T. E., Landis, J. R., Benson, R. M., & Kemp, P. L. Reliability of the Group for the Advancement of Psychiatry diagnostic categories in child psychopathology. *Archives of General Psychiatry,* 1978, *35,* 1461-1466.

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

Bruning, J. L., & Kintz, B. L. *Computational handbook of statistics.* Glenview, Illinois: Scott, Foresman, 1968.

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

Cohen, J. *Statistical power analysis for the behaviorial sciences* (Rev. ed.). New York: Academic Press, 1977.

Conners, C. K. A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry,* 1969, *126,* 884-888.

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

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

Eisenberg, J. G., Gersten, J. C., Langner, T. S., McCarthy, E. D., & Simcha-Fagan, O. A behavioral classification of welfare children from survey data. *American Journal of Orthopsychiatry,* 1976, *46,* 447-463.

Everitt, B. *Cluster analysis.* London: Heinemann Educational Books, 1974.

Freeman, M. A reliability study of psychiatric diagnosis in childhood and adolescence. *Journal of Child Psychology and Psychiatry,* 1971, *12,* 43-54.

Group for the Advancement of Psychiatry. *Psychopathological disorders in childhood: Theoretical considerations and a proposed classification (Report No. 62).* New York: Author, 1966.

Hafner, A. J., Quast, W., & Shea, M. J. The adult adjustment of one thousand psychiatric patients: Initial findings from a twenty-five year follow-up. In R. O. Wirt, G. Winokur, & M. Roff (Eds.), *Life history in psychopathology* (Vol. 4). Minneapolis: University of Minnesota Press, 1975.

Haggard, E. A. *Intraclass correlation and the analysis of variance.* New York: Dryden, 1958.

Lorion, R. P., & Cowen, E. L. Comparison of two outcome groups in a school based mental health project. *American Journal of Community Psychology,* 1976, *4,* 65-73.

Lorion, R. P., Cowen, E. L., & Caldwell, R. A. Problem types of children referred to a school based mental health program. *Journal of Consulting and Clinical Psychology,* 1974, *42,* 491-496.

Mattison, R., Cantwell, D. P., Russell, A. T., & Will, L. A comparison of DSM-II and DSM-III in the diagnosis of childhood psychiatric disorders. *Archives of General Psychiatry*, 1979, *36*, 1217-1222.

Miller, L. C. Louisville Behavior Checklist for males 6-12 years of age. *Psychological Reports*, 1967, *21*, 885-896.

Peterson, D. R. Behavior problems of middle childhood. *Journal of Consulting Psychology*, 1961, *25*, 205-209.

Prior, M., Boulton, D., Gajzago, C., & Perry, D. The classification of childhood psychosis by numerical taxonomy. *Journal of Child Psychology and Psychiatry*, 1975, *16*, 321-330.

Quay, H. C. Classification. In H. C. Quay & J. S. Werry (Eds.), *Psychopathological disorders of childhood* (2nd ed.). New York: Wiley, 1979.

Rolf, J. E. The social and academic competence of children vulnerable to schizophrenia and other behavior pathologies. *Journal of Abnormal Psychology*, 1972, *80*, 225-243.

Sokal, R. R., & Michener, C. D. A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin*, 1958, *38*, 1409-1438.

Sokal, R. R., & Rohlf, F. J. The comparison of dendograms by objective methods. *Taxon*, 1962, *11*, 33-40.

Spivack, G., Swift, M., & Prewitt, J. Syndromes of disturbed classroom behavior: A behavior diagnostic system for elementary school. *Journal of Special Education*, 1971, *5*, 269-292.

Tarter, R. E., Templer, E. K., & Hardy, C. Reliability of psychiatric diagnosis. *Diseases of the Nervous System*, 1975, *36*, 30-31.

Winer, B. J. *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill, 1971.