

Convergence Between Statistically Derived Behavior Problem Syndromes and Child Psychiatric Diagnoses

Craig Edelbrock^{1,2} and Anthony J. Costello¹

The relations between scores on statistically derived behavior problem syndromes and DSM-III diagnoses were examined for 270 clinically referred children aged 6 through 16. Each child's parent completed the Child Behavior Checklist (CBCL) and was administered the NIMH Diagnostic Interview Schedule for Children (DISC)—a structured interview covering DSM-III diagnostic criteria. Numerous behavior problems scales scored from the CBCL were significantly related to one or more diagnoses. The strongest relations were between scores on the Hyperactive, Delinquent, and Depressed scales and diagnoses of Attention Deficit Disorder, Conduct Disorder, and Depression/Dysthymia, respectively. This convergence supports the validity of some syndromal constructs common to both assessment paradigms.

The assessment and classification of child psychopathology has been guided by two major paradigms: the *medical* and the *psychometric*. The medical paradigm, exemplified by the American Psychiatric Association's *Diagnostic and Statistical Manual* (DSM-III; American Psychiatric Association, 1980) is founded on the assumption that psychopathological disorders are disease entities. Clinical interviews, physical examinations, and laboratory tests are used to categorize the patient by type of organic pathology. Even when an

Manuscript received in final form September 2, 1987.

This research was supported in part by NIMH grant no. MH37372 and by NIMH contract no. RFP-DB-81-0027. The first author is supported by a Faculty Scholar's Award from the William T. Grant Foundation. The authors wish to thank Tom Achenbach and Russ Barkley for their valuable comments on an earlier draft.

¹Department of Psychiatry and Behavioral Science, University of Massachusetts Medical School, Worcester, Massachusetts 01605.

²Address all correspondence to Dr. Craig Edelbrock, Department of Psychiatry and Behavioral Science, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, Massachusetts 01605.

organic etiology cannot be identified, an underlying defect or deficit is often postulated to account for observed signs and symptoms.

Within the psychometric paradigm, there is less concern with identifying organic etiologies and more concern with describing phenotypic variations. Standardized rating scales, personality inventories, and psychological tests are used to quantify individual differences in behavior, personality, and intelligence. Psychopathology can be viewed as a quantitative deviation from normal, rather than as a discrete disorder.

There has been considerable progress within each paradigm in the past few years. The current edition of the DSM (DSM-III), for example, includes more disorders of infancy, childhood, and adolescence and more explicit diagnostic criteria than its predecessors. Within the psychometric paradigm, numerous assessment materials and procedures for children and adolescents have been recently developed or revised (see Achenbach & Edelbrock, 1984, for a review). Unfortunately, there have been few empirical efforts addressing the relations between the two paradigms. Instead, there are two worlds of research and clinical practice: one employing categorical diagnoses, the other employing quantitative measurements.

Synergistic research combining both paradigms seems valuable for several reasons. First, neither approach has offered definitive achievements. Synergistic research could highlight the strengths and weaknesses of each approach and illuminate areas for future work within each paradigm. Second, finding significant convergence between the medical and psychometric approaches would bolster the validity of constructs common to both. Determining the diagnostic correlates of quantitative measures would not only contribute to their construct validation but would increase their clinical and research utility. Determining how diagnostic subgroups differ on quantitative measures would yield parallel benefits. Third, child mental health professionals, whether they are involved in professional training, clinical service, or research, must be conversant with both paradigms. Better understanding of the relations between psychometric measures and psychiatric diagnoses would facilitate the translation of information across paradigms and improve communication among professionals differing in background, training, and theoretical orientation.

The goal of this study was to determine the diagnostic correlates of a behavioral measure that is widely used in the field of child mental health—the Child Behavior Profile (Achenbach & Edelbrock, 1983). Achenbach (1980) has discussed the correspondence between statistically derived behavior problem syndromes and DSM-III diagnoses. However, these approximate relations were based on descriptive similarities between the syndromes and the diagnostic criteria. At the time, tests of the relations between statistically derived syndromes and DSM-III diagnoses had not been made. The present study

provides an empirical test of the convergence between statistically derived syndromes and psychiatric diagnoses in a large sample of clinically referred children.

METHOD

Subjects

Subjects were 270 children aged 6–16 who had been referred for inpatient or outpatient mental health services. Details regarding this sample and the purposes of this multifaceted project have been provided elsewhere (Costello, Edelbrock, Dulcan, Kalas, & Klaric, 1984). Briefly, parents of 447 referred children were asked to participate in a diagnostic study and 379 consented. Of those, 316 parents and their children participated in at least one portion of the study. The current sample includes the 270 subjects whose parents completed both the diagnostic interview and the CBCL. This clinic sample was 68% male and 32% female; 60% were white, 39% were black, and 1% represented other groups. The sample included the entire range of socioeconomic status (SES), but the lower classes were moderately over-represented.

Measures

Child Behavior Profile. The Child Behavior Checklist (CBCL) is an omnibus measure of child behavior problems and competencies designed to be completed by parents (see Achenbach & Edelbrock, 1983). It is scored on the Child Behavior Profile, which includes problem scales derived from factor analyses of CBCLs completed on large sample of clinically referred children. Separate editions of the profile have been developed and standardized for boys and girls at ages 4–5, 6–11, and 12–16. Each edition comprises eight or nine factor-based problem scales representing fairly circumscribed “narrow-band” syndromes (e.g., Aggressive, Delinquent, Hyperactive). Two global “broad-band” syndromes, labeled Internalizing and Externalizing, have also been developed for each edition via second-order factor analysis of the narrow-band scales. The reliability, validity, and normative standardization of the profile have been discussed in detail by Achenbach and Edelbrock (1983).

DSM Diagnoses. One difficulty with the DSM is that it specifies *what* to assess but not *how* to assess it. In the past few years, structured interview schedules have been developed as a means of operationalizing DSM criteria. There are several interview schedules covering childhood disorders, differ-

ing in adherence to the DSM, diagnostic coverage, degree of structure, reliability, and amount of clinical training and inference required for administration and scoring (see Edelbrock & Costello, 1987, for a comparative review). In this study, the NIMH Diagnostic Interview Schedule for Children (DISC) was used as an operationalization of the DSM-III.

The DISC is a structured diagnostic interview that covers almost all Axis I diagnoses applied to children and adolescents. The history, development, reliability, and validity of the DISC have been described in detail elsewhere (Costello et al., 1984; Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985; Edelbrock, Costello, Dulcan, Conover, & Kalas, 1986). Briefly, the DISC was designed as a research tool, is highly structured, and is suitable for administration by either clinically trained or lay interviewers. Parallel versions have been developed for parents (DISC-P) and children (DISC-C). The focus here was on the DISC-P interview with the parent. Although the DISC-C may be useful with adolescents, it has low reliability and questionable validity with young children (see Edelbrock et al., 1985, 1986; Costello, Edelbrock, & Costello, 1985).

Short-term test-retest reliabilities of symptom scores derived from the DISC-P have been moderately high (average $r = .76$, range: .44-.90). Validity of the DISC-P has been supported by convergent relations with numerous behavioral and psychiatric measures (see Costello et al., 1984) and by the ability to discriminate significantly between matched samples of pediatric and psychiatric referrals (Costello et al., 1985).

Procedure

During their first clinic appointment, each parent completed both the CBCL and the DISC-P during the course of a comprehensive child psychiatric evaluation that covered demographic and background information, medical and developmental history, current social, behavioral, and psychiatric status, school history, and current academic performance. Parents with difficulties reading the CBCL were administered the measure orally by a clerical worker. The DISC interviews were administered individually to each parent by a clinician or a lay interviewer trained in the use of the DISC. Previous analyses have shown no significant differences between clinicians and lay interviewers in the prevalence or severity of child psychiatric symptoms assessed by the DISC (Costello et al., 1984).

The CBCLs were scored according to scales and norms appropriate for each child's sex and age. Scores on scales having counterparts in two or more sex and age groups were combined in the analyses according to guidelines provided by Achenbach and Edelbrock (1983, pp. 142-143). The DISC interview with the parent was analyzed using diagnostic algorithms that oper-

ationalize DSM-III criteria for Axis I child psychiatric disorders. The algorithms determine if DSM-III criteria for (a) the number and severity of symptoms, (b) duration, and (c) associated impairments were fulfilled for each diagnosis. If the criteria were fulfilled, the diagnosis was coded present; if not, the diagnosis was coded absent.

RESULTS

Point-biserial correlations between each Axis I diagnosis derived from the DISC interviews (coded 0 = absent, 1 = present), and *T* scores for the narrow-band behavior problem scales of the profile were computed. Since numerous correlations were tested for significance, the alpha level was raised to $p < .01$ in order to reduce Type I errors. Table I summarizes the significant correlations for each profile scale. For example, scores on the Aggressive scale were significantly correlated with the diagnosis of Conduct Disorder. Scores on the Anxious-Obsessive scale (which is scored only for girls aged

Table I. Summary of Relations Between Child Behavior Profile Scales and DSM-III Diagnoses^a

Scale	DSM-III diagnosis
Aggressive ^{1,2,3,4}	Conduct disorder
Anxious-obsessive ⁴	Overanxious disorder
Cruel ^{3,4}	Conduct disorder
Depressed withdrawal ⁴	Major depression Dysthymia
Delinquent ^{1,2,3,4}	Conduct disorder
Depressed ^{1,3}	Major depression Dysthymia
Hostile withdrawal ²	Attention deficit disorder Oppositional disorder
Hyperactive ^{1,2,3}	Attention deficit disorder
Immature ²	Attention deficit disorder Separation anxiety
Immature-hyperactive ⁴	Attention deficit disorder
Schizoid ^{2,4}	Overanxious disorder
Schizoid or anxious ¹	Separation anxiety Overanxious disorder
Social withdrawal ^{1,3}	Avoidant disorder Major depression Dysthymia
Somatic complaints ^{1,2,3,4}	Overanxious disorder
Uncommunicative ^{1,2}	Overanxious disorder Major depression Dysthymia

^aSuperscripts indicate sex and age groups for which the scale is scored: 1 = boys aged 6-11, 2 = boys aged 12-16, 3 = girls aged 6-11, 4 = girls aged 12-16.

12-16) were significantly correlated with the diagnosis of overanxious Disorder, and so on. Some scales were significantly correlated with two or more diagnoses. Scores on the Immature scale, for instance, were significantly related to the diagnosis of Attention Deficit Disorder and Separation Anxiety.

All of the relations summarized in Table I were positive in direction. Thus, increasing scores for a given scale were associated with a higher probability of receiving the listed diagnosis. The relations are bidirectional and can be interpreted in two ways: as diagnostic correlates of the behavioral rating or as behavioral correlates of the diagnoses. For example, children with high scores on the Aggressive scale had a significantly higher probability of being diagnosed Conduct Disorder than children with low scores on the Aggressive scale. Alternatively, children diagnosed as having a Conduct Disorder had significantly higher scores on the Aggressive scale than children *not* diagnosed as having a Conduct Disorder.

Internalizing/Externalizing

Table II lists the point-biserial correlations between *T* scores for the broad-band Internalizing and Externalizing scales and each Axis I diagnosis. As shown in Table II, diagnoses of Conduct Disorder, Oppositional Disorder, and Attention Deficit Disorder correlated significantly with Externalizing scores but not with Internalizing scores. Diagnoses of Separation Anxiety, Avoidant Disorder, Overanxious Disorder, Simple Phobia, and

Table II. Relations Between Internalizing and Externalizing Scores and DSM-III Diagnoses^a

Diagnosis	Internalizing	Externalizing
Attention deficit disorder	.14	.41 ^c
with Hyperactivity	.09	.31 ^c
without Hyperactivity	.08	.19 ^b
Conduct disorder	.05	.34 ^c
Aggressive	.10	.38 ^c
Nonaggressive	.00	.21 ^c
Oppositional disorder	.19 ^b	.42 ^c
Separation anxiety	.22 ^c	.07
Avoidant disorder	.17 ^b	-.01
Overanxious disorder	.35 ^c	.09
Major depression	.31 ^c	.14 ^b
Dysthymia	.43 ^c	.22 ^c
Social phobia	.21 ^c	.01
Simple phobia	.33 ^c	.08

^aTable entries are point-biserial correlations between each diagnosis (coded 0, 1) and *T* scores for Internalizing and Externalizing (*n* = 270).

^b*p* < .01.

^c*p* < .001.

Social Phobia correlated significantly with Internalizing scores but not with Externalizing scores. Diagnoses of both Depression and Dysthymia correlated significantly with Internalizing and Externalizing scores, although the correlations with Externalizing scores were somewhat lower.

Specific Diagnoses

Three diagnoses were analyzed in detail: Attention Deficit Disorder, Conduct Disorder, and Depression/Dysthymia. This seemed worthwhile because the diagnostic criteria for these disorders closely parallel specific profile scales (Hyperactive, Delinquent, and Depressed, respectively), and there were sufficient numbers of children with each diagnosis to permit fine-grained analyses. To examine these relations more closely, *T* scores for each scale were divided into eight ranges: *T* = 55, 56–60, 61–65, 66–70, 71–75, 76–80, 81–85, 86–100. The relationship between these *T*-score intervals and the target diagnosis (coded 0 or 1) was tested for linear trend using regression analysis (see Cohen & Cohen, 1975, pp. 229–230, for an extensive discussion of regression analysis with dichotomous variables).

Attention Deficit Disorder. There was a very strong linear association, $F(1, 6) = 56.9, p < .001$, between scores on the Hyperactive scale and the diagnosis of Attention Deficit Disorder (ADD). Figure 1 shows the percentage of children in each *T*-score interval who fulfilled diagnostic criteria for ADD and the best-fitting line derived from the regression analysis. As shown in Figure 1, as scores increased, successively higher proportion of children

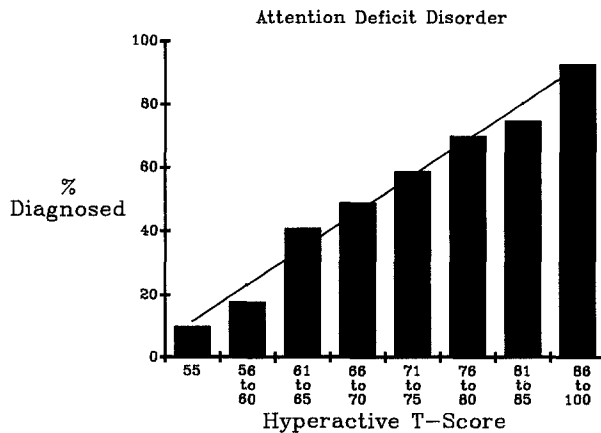


Fig. 1. The relationship between *T* scores on the Hyperactive scale and the diagnosis of Attention Deficit Disorder.

received the diagnosis. For example, only 10% (3/29) of the children with scores of 55 were diagnosed ADD, compared with 18% (4/22) of those with *T* scores of 56 to 60, and so on. At the other extreme, 93% (14/15) of those with *T* scores of 86 to 100 obtained the diagnosis.

These results suggest that there is no "diagnostic threshold" for ADD below which children do not receive the diagnosis and above which they do. Instead, the relationship between *T* scores and diagnoses is linear: Increasing scores were related to incrementally higher probabilities of obtaining the diagnosis.³ The cutoff point for determining whether a given children is "disturbed" or not is somewhat arbitrary but can be guided by normative data, as in the case of the profile, or by clinical wisdom, as in the case of the DSM. For example, *T* scores greater than 70 on the Hyperactive scale exceed the 98th percentile in the normative sample and are taken to represent clinically significant hyperactivity. In contrast, the DSM requires that a minimum number of diagnostic criteria, determined by committee consensus, are fulfilled to warrant the diagnosis of ADD. Surprisingly, there was some correspondence between these radically different approaches to classification. As shown in Figure 1, the majority of children with *T* scores greater than 70 were diagnosed ADD, whereas the majority of children with *T* scores of 70 or below were not.

Conduct Disorder. There was a very strong linear association, similar to that depicted in Figure 1, between scores on the Delinquent scale and the diagnosis of Conduct Disorder, $F(1, 6) = 65.5, p < .001$. Only 7% (1/14) of the subject with *T* scores of 55 were diagnosed Conduct Disorder, compared with 85% (11/13) of those having *T* scores of 86–100. Overall, *T* scores greater than 70 were associated with a probability of Conduct Disorder diagnosis, but the majority of subjects (63%) having *T* scores in the range of 66–70 also received the diagnosis. Scores on the Delinquent scale were also strongly related to the diagnosis of both Aggressive and Nonaggressive subtypes of Conduct Disorder and did not differentiate among these subtypes. However, scores on the Aggressive scale and the Cruel scale (which is scored only for girls) were more strongly associated with Aggressive Conduct Disorder than Nonaggressive Conduct Disorder and may therefore provide a basis for differentiating among subgroups. Specifically, both Aggressive and Nonaggressive subtypes tend to obtain high scores on the Delinquent scale,

³Cohen and Cohen (1975, p. 229) have pointed out that interval size and interval inequality in the independent variable can effect the shape of observed relationships with other variables. The issue here was whether the use of smaller *T*-score intervals and/or equal *T*-score intervals would reveal a curvilinear relationship with diagnosis suggestive of a diagnostic threshold. However, subsequent analyses revealed linear relationships, regardless of how *T*-score intervals were chosen.

but only the Aggressive subtype tends to obtain high scores on the Aggressive and/or Cruel scales.

Child Depression. Two DSM-III diagnoses are related to childhood depression: Major Depression and Dysthymia. Only nine children met diagnostic criteria for Major Depression, but all of them had *T* scores greater than 70 on the Depressed scale. All four children having *T* scores greater than 80 on the Depressed scale were diagnosed as Depressed.

Diagnostic criteria for Dysthymia parallel those for Major Depression, but the former requires fewer and less severe symptoms of shorter minimum duration. More children fulfilled criteria for Dysthymia than Major Depression, permitting fine-grained analysis. Also, all nine children who fulfilled diagnostic criteria for Major Depression also fulfilled criteria for Dysthymia, so this subgroup is better termed *Depression/Dysthymia*.

There was a significant linear association, similar to that in Figure 1, between scores on the Depressed scale and the diagnosis of Depression/Dysthymia ($F(1, 6) = 35.5, p < .001$). This suggests that high scores on the Depressed scale are related to diagnosable childhood depression/dysthymia. However, many children diagnosed as Dysthymic scored low on the Depressed scale.

Mean Scores of Diagnostic Groups

An alternative statistical approach is to examine mean differences in scale scores between diagnostic groups. In order to explore variation above the diagnostic thresholds set by the DSM, we distinguished between cases that fulfilled minimum symptom criteria for diagnosis (i.e., "mild" disorder) and those that exceeded the minimum (i.e., "severe" disorder). Each diagnosis was treated as having three "levels": None, Mild, and Severe. Such distinctions are not made in the DSM-III, but they were used here as an exploratory research technique. Figure 2 shows mean differences between the levels of three diagnoses (Conduct Disorder, ADD, and Depression/Dysthymia) and *T* scores on the most similar scales (Delinquent, Hyperactive, and Depressed, respectively).

As shown in Figure 2, cases grouped according to level of diagnoses had progressively higher mean *T* scores on the corresponding scales. For each diagnosis, (a) the mild and severe group obtained significantly higher mean *T* scores than the group failing to fulfill diagnostic criteria ($p < .05$ by *t* tests), (b) the mild and severe groups obtained mean *T* scores above 70, whereas the cases not fulfilling diagnostic criteria obtained mean *T* scores below 70, and (c) the severe group obtained higher mean *T* scores than the

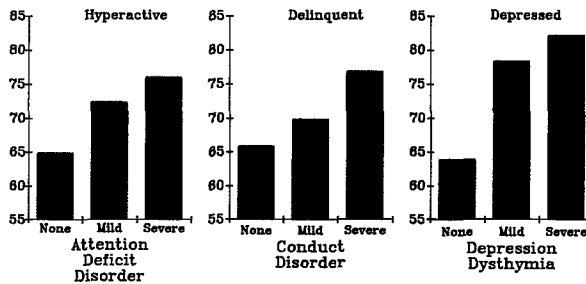


Fig. 2. Mean differences in Child Behavior Profile scales for DSM diagnostic groups.

mild group (although this was statistically significant only for Conduct Disorder).

DISCUSSION

These results indicate significant relations between scores on the Child Behavior Profile and child psychiatric diagnoses derived from structured interviews. These relations suggest substantial convergence between two different approaches to assessing child psychopathology, and they lend support to the validity of certain constructs common to both. Both the ratings and the diagnoses were derived from parents, so the degree of convergence may be inflated by shared method variance. However, a recent meta-analysis (Achenbach, McConaughy & Howell, 1987) indicates that correlations are generally low between different informants' ratings of children's emotional/behavioral problems. It would appear difficult, therefore, to detect convergence across assessment paradigms and informants simultaneously.

There are also fundamental differences between the nosological and psychometric approaches, and they partition the phenomena of child psychopathology differently. Several profile scales were significantly related to two or more diagnoses, and several diagnoses were related to two or more scales. Although many profile scales have significant diagnostic correlates, most scales do not correspond to one diagnosis. Conversely, children who fulfill diagnostic criteria for a given disorder may manifest emotional and behavioral problems in several other areas.

Three scales had clear-cut diagnostic correlates. Scores on the Hyperactive scales were strongly related to the diagnosis of ADD. The Delinquent scale is an index of general Conduct Disorder, whereas the Aggressive and Cruel scales were more strongly related to Aggressive than to Nonaggressive subtypes of Conduct Disorder. High scores on the Depressed scale were as-

sociated with diagnosed Major Depression. Many children who fulfilled diagnostic criteria for Dysthymia, however, obtained low scores on the Depressed scale, suggesting that the diagnostic criteria for Dysthymia are too lenient (at least for children).

These results also show how DSM-III diagnoses align with the broad-band Internalizing and Externalizing syndromes. Not surprisingly, ADD, Conduct Disorder, and Oppositional Disorder aligned with the Externalizing syndrome, whereas Separation Anxiety, Avoidant Disorder, Overanxious Disorder, Simple Phobia, and Social Phobia aligned with the Internalizing syndrome. One unexpected finding was that diagnoses of Depression and Dysthymia correlated significantly with both Internalizing and Externalizing scores. This suggests that a substantial proportion of depressed children may also manifest externalizing problems, such as Conduct Disorder. Previous cluster analyses of Profile scores for girls aged 12–16 have identified a profile type defined by high scores on both the Delinquent and Depressed-Withdrawal scales (Edelbrock & Achenbach, 1980). Other investigators have also found a high prevalence of depression among conduct-disorder children (Puig-Antich, 1982). In the DSM-III, disorders of infancy, childhood, and adolescence were not grouped into broader categories. In the proposed revision of the DSM (DSM-III-R; American Psychiatric Association, 1987), however, ADD, Conduct Disorder, and Oppositional Disorder are grouped together as Disruptive Behavior Disorders.” The current results offer some empirical support for this higher-order grouping of diagnoses.

The changes proposed in the DSM-III-R probably would not alter the nature of the convergence between diagnoses and profile scales. For diagnoses such as Attention Deficit/Hyperactivity Disorder and Conduct Disorder, the new criteria are more similar in content to the corresponding profile scales than the previous criteria. Also, the “Chinese menu” approach of employing multiple symptom clusters for a given disorder has been abandoned. The new approach has been to provide for each disorder a single symptom list with one diagnostic threshold – an approach very similar to using quantitative scale scores with a single cutoff point. These changes might result in a higher *degree* of convergence between certain diagnoses and their corresponding profile scales – although such relations must be tested empirically.

Cross-paradigm studies of this type raise fundamental questions regarding the nature of child psychopathology and the appropriateness of different assessment approaches. Scores on the Hyperactive, Delinquent, and Depressed scales were linearly related to their corresponding diagnoses. A diagnostic threshold was not apparent. Rather, increasing scores on each scale paralleled incrementally higher probabilities of obtaining the corresponding diagnosis. These findings appear to argue for a continuous distribution of such problem behaviors rather than the existence of discrete disorders that are either present or absent.

Categorical approach such as the DSM impose a fairly arbitrary cutoff point for determining whether a child has a given disorder or not. This entails two major limitations: (a) Potentially important information regarding how far above or below the cutoff point a child falls is ignored, and (b) diagnostic subgroups can be very heterogeneous. Children who fulfill at least minimum diagnostic criteria, for example, still differ in the number and severity of symptoms they manifest. Although such children are grouped under one diagnostic label, the number and severity of symptoms *beyond the diagnostic cutoff point* may be predictive of course, prognosis, treatment response, and outcome. Conversely, children who fall just short of the cutoff point are similar to those fulfilling minimum diagnostic criteria, but this potentially important similarity is obscured by the present/absent categorization.

Despite these limitations, categorical approaches may prove superior to quantitative-descriptive approaches in some areas of child psychopathology. Achenbach (1980) has pointed out that psychometric approaches are not well suited to the assessment of rare psychopathologies. Unfortunately, some rare disorders are not covered by the DISC and others were too infrequent in this clinic sample to permit statistical analysis. For example, the relationship between the Somatic Complaints scale of the profile and the diagnosis of Somatization Disorder could not be determined because that diagnosis is not addressed by the DISC. Several other possible relations discussed by Achenbach (1980) could not be explored owing to inadequate numbers of children receiving certain diagnoses. These included the relations between (a) the Schizoid scale and diagnoses of Schizoid Personality Disorder and Schizophrenia, (b) the Obsessive-Compulsive scale and the diagnosis of Obsessive-Compulsive Disorder, (c) the Sex Problems scale and the diagnosis of Gender Identity Disorder and (d) the Uncommunicative scale and the diagnosis of Elective Mutism. Further research on selected samples of children manifesting these disorders is necessary to determine correspondence to empirically derived syndromes.

REFERENCES

- Achenbach, T. M. (1980). DSM-III in light of empirical research on the classification of child psychopathology. *Journal of the American Academy of Child Psychiatry*, *19*, 395-412.
- Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington, Vermont: University of Vermont.
- Achenbach, T. M., & Edelbrock, C. (1984). Psychopathology of childhood. *Annual Review of Psychology*, *35*, 227-256.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.

- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, D.C.: Author.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, D.C.: Author.
- Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, New Jersey: Erlbaum.
- Costello, E. J., Edelbrock, C., & Costello, A. J. (1985). Validity of the NIMH Diagnostic Interview Schedule for Children: A comparison between psychiatric and pediatric referrals. *Journal of Abnormal Child Psychology*, *13*, 579-595.
- Costello, A. J., Edelbrock, C., Dulcan, M. K., Kalas, R., & Klaric, S. H. (1984). *Development and testing of the NIMH Diagnostic Interview Schedule for Children on a clinical population: Final Report*. (Contract No. RFP-DB-81-002). Rockville, Maryland: Center for Epidemiologic Studies, National Institute of Mental Health.
- Edelbrock, C., & Achenbach, T. M. (1980). A typology of Child Behavior Profile patterns: Distribution and correlates for disturbed children aged 6-16. *Journal of Abnormal Child Psychology*, *8*, 441-470.
- Edelbrock, C., & Costello, A. J. (1987). A review of structured diagnostic interview schedules for children. In M. Rutter, H. Tuma, & I. Lann (Eds.), *Assessment and diagnosis in child and adolescent psychopathology*. New York: Guilford Press.
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kalas, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry*, *27*, 181-190.
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development*, *56*, 265-275.
- Puig-Antich, J. (1982). Major depression and conduct disorder in prepuberty. *Journal of the American Academy of Child Psychiatry*, *21*, 118-128.