*Journal of Abnormal Child Psychology, Vol. 23, No. 4, 1995*

# Community-Based Multiple-Gate Screening of Children at Risk for Conduct Disorder

**Gerald J. August,[1,2] George M. Realmuto,[1] Ross D. Crosby,[1] and Angus W. MacDonald III[1]**

*The present study employed a multiple-gate screening procedure to identify children at risk for the development of conduct disorder. Measures of cross-setting disruptive behavior and parent discipline practices were administered in sequential fashion to screen a population of 7,231 children attending suburban elementary schools. Convergent validity of the respective gating measures was confirmed by significant correlations with adjustment constructs. Analyses of covariance performed between positive screens, negative screens, and low-risk comparison children on adjustment constructs at each gate supported the discriminative validity of the gating procedure. Hierarchical regression analyses demonstrated that the gating measures were predictive of diagnostic ratings of attention deficit hyperactivity disorder and oppositional defiant disorder that were obtained 18 months following the screening. A stepwise logistic regression analysis indicated that the best predictors of high-risk group membership were variables related to family process, including poor family communication and involvement, poor maternal coping skills, and an external parent locus of control.*

An important issue in Conduct Disorder (CD) prevention research is defining the methods by which subjects in the population are selected for intervention trials. Several difficulties exist with the design of community-based interventions that target "at-risk" populations. These include determining the factors that constitute risk, using cost-effective assessment procedures to measure these risk

This study was supported by grant MH-46584 from the National Institute of Mental Health.
[1]Department of Psychiatry, Division of Child & Adolescent Psychiatry, University of Minnesota, Minneapolis, MN 55455.
[2]Address all correspondence and requests for reprints to Gerald J. August, Ph.D., Box 95 Mayo Building, Department of Psychiatry, University of Minnesota, Minneapolis, Minnesota 55455.



0091-0627/95/0800-0521$07.50/0 

factors, and overcoming the practical, ethical, and data analytical problems associated with false positive prediction error in subject selection.

The difficulties involved in identifying participants for prevention programs have prompted innovative approaches to risk screening. For example, Loeber, Dishion, and Patterson (1984) designed an elegant multiple-gate screening procedure to improve the detection of children at risk for delinquency. The procedure involved a stepwise series of assessments in which the least costly assessment was administered first to the larger population. More expensive assessments were subsequently administered to only those who screened positive at earlier gates. Each successive assessment was designed to narrow down the population so as to reduce prediction error to an acceptable level and to minimize the cost of large-scale screening.

The present study is based on data that were gathered as part of the Minnesota Competence Enhancement Project. The project is a longitudinal study designed to assess the efficacy of a school-based early detection and multicomponent intervention program for suburban children at risk for the development of CD and related problem outcomes. Large scale interventions for disruptive children in suburban populations are not common, yet violence, substance abuse and school dropout among adolescents are rapidly growing problems as they continue to be in urban communities (Dryfoos, 1990). The pressing need to study these problems and the anticipated high geographical stability of a suburban population were thought to be positive sample attributes that would allow accurate assessment of the effectiveness of the intervention over time.

Conduct disorder has been reported to affect approximately 1.5% to 5.5% of children in the general population (Costello, 1989), with rates approaching 10 percent in urban communities (Offord *et al.*, 1987). Because the population studied was suburban and of predominantly middle socioeconomic status (SES), the base rate of this outcome was expected to be lower than rates reported in surveys of urban, socially disorganized communities. Low base rate presented the problem of increased false positive predictor error. False prediction can result in the ineffective use of resources, personal and community bias regarding stigma, and iatrogenic effects of the intervention itself (Muehrer & Koretz, 1992). To this end, a multiple-gate screening device, adapted from that of Loeber *et al.* (1984), was used to select moderate- and high-risk children for participation in an experimental intervention trial. The device first screened for a premorbid risk factor, evidence of cross-setting disruptive behavior, and then screened for a mediator variable suspected of "catalyzing" the pathological effects of such disruptive behavior, namely, unskilled parent discipline practices. In the present study we present data to demonstrate cross-sectional relations between our screening criteria and concurrent levels of psychological adjustment, emerging psychopathology, and selected family process variables in a middle-SES suburban community.

## METHOD

### *Subjects*

The population screened included 7,231 children in grades one through four in 22 suburban elementary schools of five independent school districts located in the outer ring of the Minneapolis metropolitan area. The socioeconomic status of these communities spanned levels I to V of the Hollingshead (1975) SES classification, but were predominantly middle class (levels II and III). The mean percent of students receiving free or reduced-price lunches across all the schools was 13.2% with a range of 1% to 34%. More than 95% of the students were Caucasian.

### *Procedure for Assessment of Risk*

A three-gate procedure was employed to identify high-risk children. Students with IQs less than 80, documented diagnoses of pervasive developmental disorders, and severe emotional disorders (e.g., psychosis) were excluded. Students previously diagnosed with disruptive behavioral disorders (e.g., attention deficit hyperactivity disorder, or ADHD) or students currently on psychotropic medications were included in the screening. Because this was a large-scale prevention study, it was necessary to choose a screening device that identified children at a relatively young age, prior to the onset of serious conduct problems, and that was easily administered and convenient for mass screening.

*Gate 1: Teacher Screening.* The Hyperactivity Index (HI-T) taken from the Revised Conners Teacher Rating Scale (CTRS-R; Goyette, Conners, & Ulrich, 1978) was used to screen the entire population. The HI-T is also known as the Conners Abbreviated Symptom Questionnaire (ASQ). This index was chosen because it includes the 10 most frequently endorsed items by teachers in rating hyperactive and disruptive behaviors in young children. Test–rest reliabilities ranging from .91 to .98 have been found over 1-week intervals and .89 over a 2-week interval (see Barkley, 1988).

The mean and standard deviation (*SD*) of the HI-T for the entire population ($N$ = 7,231) was .46 and .65. A cutoff score of 1.6 (for combined sexes) was chosen because of an observed discontinuity in the skew at this point in the distribution and because the score is within the clinical range in most studies of childhood hyperactivity. This score defined a threshold 1.75 standard deviation units above the normative mean identifying 642 (8.9%) children.

Research staff contacted the legal guardian of each child screened positive by Gate 1 to obtain consent for participation in a second gating procedure and possible entry into a school-based intervention. If consent was not given, the child was dropped from further procedures. Twenty-three percent ($n$ = 148) were unable or unwilling to consent for Gate 2. Children who consented for further participation ($N$ = 494) were compared with those who refused consent ($n$ = 148). No significant group differences were found on the HI-T, Conduct Problems, or Inattention/Passivity subscales of the CTRS-R.

*Gate 2: Parent Screening.* The 10-item Hyperactivity Index (HI-P) of the Revised Conners Parent Rating Scale (CPRS-R; Goyette et al., 1978) was completed as part of a larger telephone interview with the child's primary caretaker (usually the mother). Since only positive screens participated at Gate 2, normative data from this measure for the entire original population were not available. Therefore, the cutoff score for positive screen was derived from normative data on the HI-P reported by Goyette *et al.* for combined sexes for children 9 to 11 years old. This suggested a cutoff score of 1.3 on the HI-P which, similarly to the HI-T, established a threshold + 1.75 standard deviation units from the mean. In sum, the two-gate screening procedure identified any child scoring 1.75 standard deviation units above the normative means on both the HI-T and HI-P. Of 494 children screened at Gate 2, 400 (81%) screened positive for cross-setting disruptive behavior and were designated as *cases.* Cases represented 5.5% of the original population surveyed. The ratio of boys to girls was 4.1:1 (322:78). Following Gate 2 screening, but preceding further risk assessment and formal intervention (approximately 6 months later), there were both random and systematic losses of cases. The largest contribution to the systematic losses of cases was boundary changes in one school district that resulted in the transfer of cases from study schools to nonparticipating schools. Dropout also resulted from family relocation out of the participating school districts, parent noncompliance, and child refusal. A total of 82 (20%) of the 400 cases were lost (dropouts), while 318 cases were rostered for continued participation. The ratio of boys to girls in the remaining group was 4:1 (256:62). Cases ($n$ = 318) and dropouts ($n$ = 82) were compared on socioeconomic status, family size, single-parent status, parents' ages, and children's ages with no significant differences found. HI-T (Gate 1) and HI-P (Gate 2) scores did not significantly discriminate groups, however, the dropout group was significantly less deviant on the CTRS-R Total score.

*Gate 3: Parent Behavioral Management Practices.* To further increase the precision of high-risk identification among cases, a measure of parent

behavior management practices was added as a third gate. A parent-report, Likert-type scale that assessed behavioral management practices frequently targeted in the clinical practice of parent training was constructed for this study (Behavioral Management Self Assessment: BMSA). A set of 15 items adapted from items appearing on the Parental Practices Scale developed by Strayhorn and Weidman (1988) was assembled. A Cronbach alpha coefficient of .81 was obtained for this scale (see Appendix).

The 6-month stability of the BMSA was assessed for a group of negative-screened, low-risk families ($N = 91$). The Pearson correlation of scores obtained at two times was $r = .71$ ($p < .001$). Because cases experienced a part of the experimental intervention during the interval between the two assessments, stability was assessed for the entire sample (cases and low-risk comparisons) using partial correlation techniques to control for the effects of the intervention. The Pearson correlation for this adjusted measure of stability was $r = .74$, $p < .001$. Sociodemographic indices (e.g., SES, family size, single parenthood) did not correlate with BMSA scores. On the other hand, concurrent measures of family functioning correlated significantly with the BMSA. For example, $r$'s ranging from .34 to .55, were obtained on the Cohesion, Conflict and Expressiveness scales of the Family Environment Scale (Moos & Moos, 1981).

From a sample of 318 cases (children who screened positive for cross-setting disruptive behavior) parent BMSAs were returned by 188 (59%). Differences between returners of the BMSA and nonreturners were compared with individual $t$-tests (continuous data) and chi square analyses (categorical data). The groups did not differ in teacher or parent ratings of child's disruptive behavior or academic achievement. However, nonreturners were more likely to be single parents with children who had lower IQs. Scores on the BMSA ranged from 15 (skillful) to 75 (deficient). A grand total of 296 scales were returned (these included 188 from the cases and 108 from a random sample of low-risk comparison children). The distribution of scores across the sample ranged from 19 to 61, with a mean of 34.21 and a standard deviation of 8.14. To identify those at high risk, a score 1 standard deviation above the mean (42) was selected as the cutoff. Of the 188 cases who completed the BMSA and whose children screened positive at Gates 1 and 2, 42 (22%) were classified as high-risk subjects. This number represented .6% of the original population that was surveyed. Table I presents the summary of the results of the three-gate screening procedure.

*Procedure for Identification of Low Risk Comparison Subjects*

A sample of low-risk comparison students was selected from the pool of students who screened negative at Gate 1. Inclusion criteria were (i) score on HI-T less than 1.1 (1 *SD* above mean), (ii) no history of psychotropic medication use, and (iii) no prior evaluation for behavior problems. A 10% sample of negative screens was randomly generated from each of the 22 schools. Next a stratified random sampling procedure was used to generate a final sample of low-risk children that was proportionally equivalent to the cases in terms of school, grade, and gender representation. A total of 193 (3.1:1 boys/girls) subjects were selected, of whom 144 (74%) (2.8:1 boys/girls) consented to participate further. Multiple *t*-tests were performed on selected variables to discern differences between consenting comparison families ($n$ = 144) and nonconsenting comparison families ($n$ = 49). No meaningful differences were observed.

*Measures*

*Descriptive.* Sociodemographic data were obtained from a biographical questionnaire. Ratings for family SES were computed on the basis of educational and occupational levels described in the Hollingshead (1975) Four-Factor Index of Social Status. A modified algorithm was used in which the maximum level of two working parents was used. The Kaufmann Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 1990) was administered to the children to derive an IQ composite score.

*Child Adjustmen.* Adjustment was defined in terms of five constructs that were assessed from measures administered as part of a comprehensive *preintervention* assessment battery. Child self-report questionnaires and tests of academic achievement were administered individually at school. Parent questionnaires were completed at a formal information meeting for those who attended and via the mail for those who did not attend. Teacher questionnaires were delivered in packets and completed on personal time.

(i) Self-concept. The Total score from the Piers-Harris Self-Concept Scale — Revised (Piers, 1984) was used. This is an 80-item child, self-report inventory with a mean score of 50 and a standard deviation of 10. High scores suggest a positive self-evaluation. Cronbach alpha coefficients for the Total score range from .73 to 81. Test–retest reliability coefficients range from .42 (with an interval of 8 months) to .96 (with an interval of 1 month) (Piers, 1984).

(ii) Problem behaviors. The Total Scale of the Problem Behaviors domain of the Social Skills Rating System (Gresham & Elliott, 1990) was

**Table I.** Number and Percent of Children Identified Across Successive Screening Gates[a]

| Grades | Total screened *N* | Gate 1 HI - T ≥ 1.6 *N* | *%* | Gate 2 HI - P ≥ 1.3 *N* | *%* | Gate 3 BMSA ≥ 42 *N* | *%* |
|---|---|---|---|---|---|---|---|
| 1 | 2129 | 199 | 9.3 | 137 | 6.4 | 16 | 0.8 |
| 2 | 2082 | 171 | 8.2 | 106 | 5.1 | 16 | 0.4 |
| 3 | 2070 | 200 | 9.7 | 114 | 5.5 | 12 | 0.6 |
| 4 | 950[b] | 72 | 7.6 | 43 | 4.5 | 6 | 0.6 |
| Total | 7231 | 642 | 8.9 | 400 | 5.5 | 42 | 0.6 |

[a] HI - T = Hyperactivity Index rated by teachers; HI - P = Hyperactivity Index rated by parents; BMSA = Behavioral Management Self Assessment Scale.

[b] Following approval of the study but prior to the onset of any study activities, eight schools included in one of the districts altered their grade structure from K-6 to K-4. In these schools it was not reasonable to screen fourth-grade students since they would be unavailable to complete participation in the two-year intervention trial.

used to assess a broad range of behaviors (externalizing, internalizing, and hyperactivity) which might interfere with social development. A 3-point rating system (*never, sometimes, very often*) is used to rate the severity of each behavior. Responses were analyzed as raw scores, with higher scores indicating greater severity. The measure was completed by the child's mother (occasionally the father when the mother was absent from home). Internal consistency for parent report of elementary school girls was .86 and for boys was .87; test–retest reliability over a 4-week interval was .84 (Gresham & Elliott, 1990).

(iii) Social skills. The Total Scale of the Social Skills domain of the Social Skills Rating System (Gresham & Elliott, 1990) was used to assess child social behaviors and prosocial skills that can affect interpersonal relationships, peer acceptance, and academic performance. This measure was completed by the child's mother (occasionally the father). A similar scoring format to that described for the Problem Behaviors scale was utilized. Alpha coefficients for the Total Scale were .85 for elementary school girls and .88 for boys while the temporal stability (4-week interval) was .85 (Gresham & Elliot, 1990).

(iv) School-related social adjustment. The Walker-McConnell Scale of Social Competence and School Adjustment (Walker & McConnell, 1988) was completed by each child's teacher to measure adaptive sociobehavioral competencies. The Total score for the scale was used as the

dependent variable. Raw scores were converted into standard scores with a mean of 100 and standard deviation of 15. Internal consistency coefficients for children grades 1 through 5 ranged from .96 to .97, while test–retest reliability for students in grades 2 and 4 was .92 over a 3-week interval (see Walker & McConnell, 1988).

(v) Academic achievement. A composite index of academic achievement was computed from the mean *T*-scores of the Wide Range Achievement Test — Revised (WRAT-R) Spelling and Arithmetic scales (Jastak & Wilkinson, 1984) plus the Broad Reading Scale from the Woodcock Johnson Tests of Achievement (Woodcock & Johnson, 1989).

### *Dimensions of Disruptive Behavior*

The Behavioral Assessment System for Children (BASC; Reynolds & Kamphaus, 1992) was completed by each child's teacher (a different teacher from the one who completed screening and adjustment measures) and parent as part of a comprehensive preintervention assessment. The BASC is a multidimensional inventory that measures various aspects of behavior and personality. For the present study, only the data from the Aggression, Conduct Problems, Attention Problems, and Hyperactivity subscales are reported. Normative scores are provided in the form of *T*-scores with a mean of 50 and a standard deviation of 10. Coefficient alpha reliabilities for the four scales range from the middle .70s to the middle .90s across both the Teacher Report Scale (TRS) and the Parent Report Scale (PRS). Test–retest reliabilities for both forms, conducted over intervals ranging from 2 to 8 weeks are excellent, ranging from the middle .80s to the middle .90s.

### *Diagnostic Ratings of Disruptive Behavior Disorders Symptoms*

The Disruptive Behaviors Disorders (DBD) Rating Scale (Pelham, Gnagy, Greenslade, & Milich, 1992) was completed by a new teacher for each child approximately 18 months following the initial screening. This rating scale includes 36 items that coincide with the DSM-III-R diagnostic criteria for the disruptive behavior disorders (American Psychiatric Association, 1987) attention deficit hyperactivity disorder, oppositional defiant disorder, and conduct disorder. A Likert-type response format is used with each item rated on a 4-point scale ranging from 0 (not at all) to 3 (much). This allows categorical diagnoses to be converted into dimension scores. A majority of the teachers were unable to respond to the CD items, so this diagnosis was excluded. In sum, the DBD Rating Scale yielded three dependent variables, ADHD, ODD, and a disruptive symptoms composite

score that was calculated on the basis of summing *T*-Score conversions of the ADHD and ODD scales. Internal consistency coefficients computed for ADHD, ODD, and CD by the test developers were .96, .95, and .75, respectively (Pelham *et al.*, 1992).

### *Family Psychosocial Characteristics*

The Parent Personality Profile of the BASC (Kamphaus & Reynolds, personal communication) was completed by each parent as part of a comprehensive preintervention assessment. Only the data on mother report are presented here. The following subscales were analyzed to assess various dimensions of parent functioning: (i) Communication and Involvement, (ii) Confident Coping, and (iii) Depression. All scales are scored true = 1, false = 0. High scores indicate positive functioning on the Communication/Involvement and Confident Coping scales and negative functioning on the Depression scale. Coefficient alpha reliabilities are .85 for Communication and Involvement, .80 for Confident Coping, and .88 for Depression. The Parent Locus of Control Scale (PLOC; Campis, Lyman, & Prentice-Dunn, 1986) was also completed by the child's mother. The internal consistency coefficient ($r = .92$) and the test-retest reliability over a 16-day interval ($r = .83$) for the Total score are excellent (Roberts & Rowe-Hallbert, 1992). High scores on the PLOC indicate an external locus of control, an orientation that is correlated with generalized feelings of incompetence and lack of control.

## RESULTS

### *Convergent Validity of the Multiple-Gate Screening Device*

Product-moment correlations were computed among the three- gate screening measures. The Gate 1, teacher-rated Hyperactivity Index (HI-T) showed a relatively strong relationship with the Gate 2, parent-rated Hyperactivity Index (HI-P), $r = .65$, $p < .000$. HI-T was moderately related to the Gate 3, parent discipline practices (BMSA), $r = .26$, $p < .001$, as was HI-P, $r = .41$, $p < .001$. Next, measures from four different sources (child, parent, teacher, achievement testing) assessing five different constructs (self-concept, problem behaviors, social skills, school adjustment, and academic achievement) were used to determine whether the gates independently identified the children with poor adjustment scores. Different sources were used in this procedure to minimize the confounding influence of "halo" effects as much as possible. The analyses included the 494 sub-

jects who screened positive at Gate 1 plus the 144 subjects selected as low-risk comparisons. Subjects were disqualified only on those analyses where there were missing data, so the number of subjects in a given analysis varies. The results presented in Table II show that the screening measures correlated significantly and in the expected directions with the adjustment constructs (all $p < .001$).

### *Discriminant Validity I: Group Comparisons*

In order to examine the discriminative validity of the multiple-gating devise, comparisons were performed at each sequential gate between three groups of subjects: (i) those who screened positive at that gate, (ii) those who screened negative at that gate, and (iii) a comparison group of low-risk subjects who were randomly selected from the larger pool of Gate 1 negative screens. Since at Gate 1 negative screens and low risk comparisons would constitute groups selected from essentially the same subject pool, positive screens are compared only to low-risk comparison subjects at that gate. As a consequence of the sequential process involved in subject recruitment, the groups of positive and negative screened subjects are reconstituted at each gate and consequently have different numbers of subjects (only positive screens advance to subsequent gates). Comparisons were performed using one-way analyses of covariance (ANCOVAs) controlling for gender, age, SES, and IQ (Table III). Cell entries reflect adjusted means + standard deviations. As can be seen, Gate 1 positive screens were significantly more impaired than low-risk comparisons on all constructs.

At Gate 2, *post hoc* tests were performed to assess group differences following an overall significant ANCOVA. Gate 2 positive screens were significantly more impaired than low-risk comparison subjects on all constructs. Gate 2 negative screens were significantly more impaired than low-risk comparisons on all constructs except academic achievement. Gate 2 positive and negative screens differed from each other on Problem Behaviors and Social Skills. At Gate 3, both positive and negative screens were significantly more impaired than low-risk comparisons on all constructs. Similar to the results at Gate 2, positive screens at Gate 3 were significantly more impaired than negative screens on Problem Behaviors and Social Skills. Additionally, Gate 3 positive screens reported significantly lower self-concept ratings than did the negative screens.

**Table II.** Intercorrelations Among Screening Measures and Adjustment Constructs[a]

| Gate | | Self-concept[b] | | Problem behaviors[c] | | Social skills[d] | | School adjustment[e] | | Academic achievement[f] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *r* | *n* | *r* | *n* | *r* | *n* | *r* | *n* | *r* | *n* |
| 1 | HI – T | –.35 | 532 | .60 | 383 | –.48 | 385 | –.67 | 516 | –.27 | 637 |
| 2 | HI – P | –.34 | 531 | .75 | 383 | –.64 | 385 | –.57 | 516 | –.32 | 636 |
| 3 | BMSA | –.25 | 321 | .50 | 336 | –.55 | 338 | –.31 | 324 | –.20 | 341 |

[a] $p < .001$ for all correlations. HI – T = Hyperactivity Index rated by teachers; HI – P = Hyperactivity Index rated by parents; BMSA = Behavioral Management Self Assessment Scale.

[b] Piers-Harris Self-Concept Scale.

[c] Social Skills Rating System.

[d] Social Skills Rating System.

[e] Walker-McConnell Scale of Social Competence and School Adjustment.

[f] Wide Range Achievement Test — Revised and Woodcock-Johnson Tests of Achievement.

**Table III.** Comparisons of Positive Screens, Negative Screens, and Low-Risk Comparisons on Adjustment Constructs at Each Gate

| Gate | Group | | Self-concept[a] | | Problem behaviors[b] | | Social skills[c] | | School adjustment[d] | | Academic achievement[e] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *M* | *(SD)* | *M* | *(SD)* | *M* | *(SD)* | *M* | *(SD)* | *M* | *(SD)* |
| 1 | + Screen | 494 | 53.8 | (12.3) | 20.4 | (6.4) | 44.3 | (11.1) | 79.1 | (12.7) | 50.2 | (9.7) |
| | Low risk | 144 | 61.6 | (11.3) | 11.6 | (6.9) | 55.1 | (10.6) | 99.5 | (15.6) | 52.7 | (10.0) |
| 2 | + Screen | 318 | 53.8 | (12.2) | 21.5 | (6.2) | 43.1 | (11.1) | 78.9 | (12..8) | 50.1 | (9.9) |
| | - Screen | 94 | 53.6 | (13.2) | 15.4 | (4.5) | 50.1 | (8.8) | 80.4 | (12.0) | 50.8 | (7.9) |
| | Low risk | 144 | 61.6 | (11.3) | 11.6 | (6.9) | 55.1 | (10.6) | 99.5 | (15.6) | 52.7 | (10.0) |
| | *Post hoc:* | | +, -<LR | | LR<-<+ | | +<-<LR | | +, -<LR | | +<LR | |
| 3 | + Screen (high risk) | 42 | 50.2 | (14.1) | 26.3 | (5.0) | 34.9 | (9.9) | 75.3 | (10.7) | 48.1 | (9.5) |
| | - Screen (moderate risk) | 146 | 54.7 | (11.8) | 20.4 | (5.9) | 45.0 | (10.6 | 78.9 | (12.2) | 49.5 | (10.0) |
| | Low risk | 144 | 61.6 | (11.8) | 11.6 | (6.9) | 55.1 | (10.6) | 99.5 | (15.6) | 52.7 | (10.0) |
| | *Post hoc:* | | +<-<LR | | LR<-<+ | | +<-<LR | | +, -<LR | | +, -<LR | |

*Note:* All *F*-tests significant at least $p < .006$, and for *post hoc* analysis, $p < .05$; + indicates the group that screened positive at this gate; - indicates the group that screened negative at this gate; LR indicates the low-risk comparison group (Hyperactivity Index rated by teachers <1.1).

[a] Piers-Harris Self-Concept Scale.

[b] Social Skills Rating System.

[d] Social Skills Rating System.

[c] Social Skills Rating System.

[d] Walker-McConnell Scale of Social Competence and School Adjustment.

[e] Wide Range Achievement Test — Revised and Woodcock-Johnson Tests of Achievement.

*Discriminant Validity II: Stepwise Regression*

In an effort to provide further evidence of discriminant validity for the multiple-gate screening procedure, a stepwise regression analysis was performed using screening measures to predict the adjustment constructs. These stepwise analyses allow the determination of which screening measures contribute significantly to the regression equation. The analysis first entered that screening measure, if any, which best predicted the criterion measure. Screening measures were entered at subsequent stages only if they provided a significant contribution over and above those variables entered on previous steps. All subjects with data on adjustment measures (both screened subjects and low-risk comparison subjects) were included in the analyses. Table IV provides a summary of these analyses. As can be seen, all three screening measures were predictive of child's self-concept, problem behaviors, and social skills. However, the order of importance was different for self-concept and the two parent-report measures (Problem Behaviors and Social Skills). Both HI-T and HI-P, but not the BMSA, were predictive of the school adjustment score, and only HI-P was predictive of the child's academic achievement composite.

*Predictive Validity: Diagnostic Ratings of Disruptive Behavior Disorders Symptoms*

To examine an early index of the predictive validity of the multiple-gating procedure, hierarchical regression analysis procedures were used to predict teacher-rated psychiatric symptoms approximately 18 months following the initial screening. In the absence of CD outcome data, we assessed putative developmental antecedents of CD, namely, attention deficit hyperactivity disorder symptoms and oppositional defiant disorder symptoms. Ratings on DSM-III-R criteria for ADHD and ODD were obtained from the DBD Rating Scale. A Disruptive Behavior Disorders Composite (DBD-C) was also computed by adding *T*-score conversions of ADHD and ODD. To control for the initial intervention activities which had begun at the time of this assessment, intervention condition was expressed as a "dummy" variable. On consecutive steps, gating measures were forced into the equation in the order each gate was performed.

The results of this procedure are summarized in Table V. These results suggest that the first gate (HI-T) was significantly predictive of subsequent diagnoses of ADHD and ODD as well as the DBD-C. The

**Table IV.** Stepwise Regression of Screening Measures on Adjustment Constructs

| Step | Variable[a] | $R^2$ | $R^2$ change | $F^b$ | $p^b$ |
|---|---|---|---|---|---|
| | | Dependent variable: Self-concept ($n$ = 321) | | | |
| 1 | HI - T | .126 | .126 | 45.90 | <.001 |
| 2 | HI - P | .150 | .024 | 8.96 | <.003 |
| 3 | BMSA | .161 | .012 | 4.50 | <.035 |
| | | Dependent variable: Problem behaviors ($n$ = 336) | | | |
| 1 | HI - P | .562 | .562 | 428.80 | <.001 |
| 2 | BMSA | .602 | .039 | 32.90 | <.001 |
| 3 | HI - T | .625 | .024 | 21.06 | <.001 |
| | | Dependent variable: Social skills ($n$ = 338) | | | |
| 1 | HI - P | .412 | .412 | 235.40 | <.001 |
| 2 | BMSA | .504 | .093 | 62.60 | <.001 |
| 3 | HI - T | .512 | .007 | 4.94 | <.030 |
| | | Dependent variable: School adjustment ($n$ = 324) | | | |
| 1 | HI - T | .442 | .442 | 255.30 | <.001 |
| 2 | HI - P | .476 | .034 | 20.80 | <.001 |
| 3 | — | | | | |
| | | Dependent variable: Academic achievement ($n$ = 341) | | | |
| 1 | HI - P | .104 | .104 | 39.20 | <.001 |
| 2 | — | | | | |
| 3 | — | | | | |

[a] HI - T = Hyperactivity Index rated by teachers; HI - P = Hyperactivity Index rated by parents; BMSA = Behavioral Management Self Assessment Scale.
[b] Indicates the significance level for the increment in $R^2$ at each step.

second gate (HI-P) contributed significantly to the prediction of ADHD over and above the first gate, but did not make a further contribution to the prediction of ODD or the DBD-C. The third gate (BMSA) did not contribute significantly to the prediction of ADHD over and above the first two gates, but did additionally contribute significantly to the prediction of ODD and the DBD-C.

**Table V.** Hierarchical Regression of Screening Measures on Disruptive Behavior Disorders Symptoms

| Gate | Variable[a] | $R^2$ | $R^2$ change | $F^b$ | $p^b$ |
|---|---|---|---|---|---|
| | Dependent variable: Attention deficit hyperactivity disorder symptoms ($n$ = 289) | | | | |
| 1 | HI - T | .401 | .401 | 195.48 | <.001 |
| 2 | HI - P | .411 | .010 | 4.82 | <.03 |
| 3 | BMSA | .413 | .002 | 2.25 | .14 |
| | Dependent variable: Oppositional defiant disorder symptoms ($n$ = 289) | | | | |
| 1 | HI - T | .319 | .319 | 136.18 | <.001 |
| 2 | HI - P | .319 | .000 | .01 | .91 |
| 3 | BMSA | .329 | .010 | 6.13 | <.01 |
| | Dependent variable: Disruptive behavior disorders — composite ($n$ = 289) | | | | |
| 1 | HI - T | .425 | .425 | 214.94 | <.001 |
| 2 | HI - P | .426 | .001 | 1.59 | .21 |
| 3 | BMSA | .435 | .008 | 5.17 | <.02 |

[a] HI - T = Hyperactivity Index rated by teachers; HI - P = Hyperactivity Index rated by parents; BMSA = Behavioral Management Self Assessment Scale.

[b] Indicates the significance level for the increment in $R^2$ at each step.

### *Group Comparisons on Child and Family Characteristics as a Function of Risk Status.*

Table VI summarizes the results of analyses of variance (ANOVAs) and nonparametric tests performed among risk groups along with corresponding means and standard deviations for screening measures, child status and disruptive behavior dimensions data, and selected family characteristics. Children who screened positive at all three gates were relabeled as the high-risk group. The moderate-risk group included children who screened positive at Gates 1 and 2, but who failed to qualify at Gate 3 as a consequence of a subthreshold score on the BMSA. Gate 2 screens who failed to return BMSAs were not included in the moderate-risk group to guard against contamination. The possibility existed that families who failed to return BMSAs may have screened positive on this measure and consequently classified as high risk. The low-risk group included children who screened negative at Gate 1 and were se-

**Table VI.** Child and Family Characteristics by Risk Group (Standard Deviations in parentheses)[a]

| | Low risk (LR) $n$ = 144 | | Moderate risk (MR) $n$ = 146 | | High risk (HR) $n$ = 42 | | |
|---|---|---|---|---|---|---|---|
| | *M* | (*SD*) | *M* | (*SD*) | *M* | (*SD*) | Statistical test |
| **Screening measures** | | | | | | | |
| HI - T (Gate 1) | 0.18 | (0.24) | 2.18 | (0.40) | 2.17 | (0.40) | $F = 1416$, $df = 2, 329$, $p < .0001$; LR<HR, MR |
| HI - P (Gate 2) | 0.59 | (0.33) | 1.89 | (0.41) | 2.20 | (0.29) | $F = 591.3$, $df = 2, 329$, $p < .0001$; LR<MR<HR |
| BMSA (Gate 3) | 30.6 | (6.2) | 32.8 | (5.7) | 48.2 | (4.6) | $F = 59.4$, $df = 2, 319$, $p < .0001$; LR<MR<HR |
| **Child characteristics** | | | | | | | |
| Status variables | | | | | | | |
| Age | 9.0 | (1.1) | 9.0 | (1.2) | 9.1 | (1.2) | ns |
| Intelligence | 109.4 | (11.8) | 105.9 | (12.4) | 104.2 | (1.3) | $F = 4.5$, $df = 2, 328$, $p < .05$; LR<MR, HR |
| WRAT-R Spelling | 100.0 | (16.0) | 91.2 | (16.8) | 89.0 | (16.7) | $F = 13.3$, $df = 2, 328$, $p < .0001$; LR<MR, HR |
| WRAT-R Arithmetic | 100.2 | (14.8) | 92.8 | (17.2) | 88.9 | (13.1) | $F = 12.2$, $df = 2, 328$, $p < .0002$; LR<MR, HR |
| W-J Broad Reading | 111.9 | (13.4) | 104.2 | (14.6) | 100.5 | (14.6) | $F = 16.0$, $df = 2, 328$, $p < .0001$; LR<MR, HR |
| Sex (% male) | 74% | | 80% | | 86% | | ns |
| Disruptive Behavior Dimensions (BASC-TRS) | | | | | | | |
| Aggression | 46.3 | (7.1) | 57.3 | (12.2) | 63.3 | (13.9) | $F = 59.4$, $df = 2, 319$, $p < .0001$; LR<MR<HR |
| Conduct Problems | 46.5 | (6.2) | 54.3 | (11.7) | 59.6 | (15.5) | $F = 33.9$, $df = 2, 320$, $p < .0001$; LR<MR<HR |
| Hyperactivity | 45.1 | (8.2) | 59.0 | (10.7) | 61.1 | (11.1) | $F = 85.0$, $df = 2, 320$, $p < .0001$; LR<MR, HR |
| Attention Problems | 46.9 | (8.8) | 62.6 | (8.9) | 62.4 | (8.7) | $F = 124.3$, $df = 2, 320$, $p < .0001$; LR<HR, MR |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Disruptive Behavior Dimensions (BASC-PRS) | | | | | | | |
| Aggression | 47.8 | (6.4) | 61.5 | (12.2) | 70.4 | (10.5) | $F = 92.9$, $df = 2, 291$, $p < .0001$; LR<MR<HR |
| Conduct Problems | 45.5 | (6.9) | 59.9 | (13.6) | 71.0 | (16.7) | $F = 80.2$, $df = 2, 291$, $p < .0001$; LR<MR<HR |
| Hyperactivity | 43.7 | (7.2) | 64.5 | (13.9) | 73.4 | (13.4) | $F = 136.5$, $df = 2, 290$, $p < .0001$; LR<MR<HR |
| Attention Problems | 45.8 | (7.9) | 65.1 | (8.2) | 68.0 | (6.7) | $F = 218.7$, $df = 2, 289$, $p < .0001$; LR<HR, MR |
| | | | | Family Characteristics | | | |
| SES | 48.9 | (11.5) | 46.6 | (11.2) | 44.1 | (11.5) | $F = 3.2$, $df = 2, 309$, $p < .05$ |
| Family size (sibship) | 2.6 | (0.81) | 2.4 | (1.2) | 2.4 | (1.3) | ns |
| Mothers's depression | 45.1 | (6.3) | 48.6 | (8.0) | 58.7 | (11) | $F = 42.8$, $df = 2, 290$, $p < .0001$; LR<MR<HR |
| Single parent (%) | 7% | | 22% | | 26% | | Pearson $\chi^2 = 14.8$, $p < .001$ |

[a] Analysis of variance results with Tukey HSD *post hoc* tests are presented for continuous variables; Pearson chi square results are presented for sex and single-parent variables. HI – T = Hyperactivity Index rated by teachers; HI – P = Hyperactivity Index rated by parents; BMSA = Behavioral Management Self Assessment Scale; WRAT-R = Wide Range Achievement Test — Revised; W-J = Woodcock-Johnson Tests of Achievement; BASC-TRS = Behavioral Assessment System for Children — Teacher Report Scale; BASC-PRS = Behavioral Assessment System for Children — Parent Rating Scale; SES = socioeconomic status.

lected via a stratified randomization procedure. Comparisons of the Gate 1 screener showed that high- and moderate-risk groups did not differ in severity of teacher ratings on the HI-T. At Gate 2, however, high-risk children showed more severe behavior ratings than did moderate risk children. All groups differed at Gate 3, with high-risk subjects exhibiting the poorest ratings on parent discipline. The ratio of boys to girls was similar across groups and low- risk children had higher IQs and academic achievement scores than either moderate- or high-risk subjects. More specific and comprehensive measures of disruptive behavior were completed by parents and new classroom teachers 9 months following the preliminary screening. A revealing pattern of differences emerged. As expected, both high- and moderate-risk groups differed from the low risk group on all dimensions. High- and moderate-risk groups did not differ from each other on teacher ratings of attention problems and hyperactivity or on parent ratings of attention problems. However, the high-risk group showed significantly more deviant ratings on aggression and conduct problems from both sources and were rated more hyperactive by parents. An overall significant group effect was noted for SES, although *post hoc* comparisons failed to show individual group differences. Both high- and moderate-risk groups had three times more single-parent families than the low-risk group. Interestingly, maternal depression was significantly different across all three groups, with mothers of high risk children most depressed, followed by moderate- and low-risk mothers, respectively. The Pearson product–moment correlation between maternal depression and the BMSA was highly significant ($r = .46$, $p < .001$), suggesting a moderately strong relationship between depression and deficient parenting practices.

Children who screened positive at Gate 2 but failed to return completed BMSAs may represent a subgroup of high- *or* moderate-risk children. To help clarify this issue, we reanalyzed available group data from above, this time including subjects who did not return BMSAs as a fourth group (i.e., nonreturn group). The nonreturn and high-risk groups did not differ in SES, although both were lower than low- and moderate-risk groups. The nonreturn group included significantly more single parent families (38%) than any of the other groups. Children in the nonreturn group had lower IQs than their high-risk peers (mean IQ = 98; $SD = 11.7$). On the teacher-rated disruptive behavior dimensions, high-risk, moderate risk and nonreturn groups were similar for attention problems and hyperactivity. High-risk and nonreturn groups showed similar levels of conduct problems but high-risk subjects were more aggressive.

*Predictors of High-Risk Group Membership.*

In order to examine child and family characteristics that distinguish high-risk from moderate risk children, a stepwise logistic regression was performed. Predictor variables included measures of family coping skills (BASC), family communication and involvement (BASC), maternal depression (BASC), parental locus of control (PLOC), and teacher ratings of the child's hyperactivity and aggressiveness (BASC). Interactions were also created expressing the relationship between maternal depression and child aggressiveness and family coping and child aggressiveness. The factors that were significantly related to high-risk group membership included poorer ratings of family coping (odds ratio = .93, $p < .006$) and family communication (odds ratio = .95, $p < .018$) and an external parental locus of control (odds ratio = 1.06, $p < .007$). High levels of maternal depression were also associated with high risk status although the relationship did not quite reach significance. None of the interaction terms reached statistical significance.

## DISCUSSION

A 10-item teacher rating scale of child disruptive behavior (Gate 1), a 10-item parent rating scale of child disruptive behavior (Gate 2), and a 15-item parent rating scale of family discipline practices (Gate 3) were completed in sequential fashion to identify children at heightened risk for the development of CD and related problem outcomes. Convergent validity of the gating measures was suggested by significant correlations (in the expected directions) with adjustment constructs obtained from multiple sources (see Table II). In addition, the gating measures successfully discriminated children with higher from those with lower levels of adjustment, and appeared to identify a progressively more impaired sample across successive gates (see Tables III and IV). Screening measures were also effective in predicting diagnostic ratings of psychiatric symptomatology. Teacher screening at Gate 1 was a significant predictor of diagnostic ratings for both ADHD and ODD while parent screening at Gate 2 added to the prediction of ADHD but not ODD. Gate 3 further contributed to the prediction of ODD but not ADHD (see Table V).

Multigate assessment can be a cost-effective strategy for selecting at-risk participants for school-based intervention trials. Within this sampling framework, rating scales have many advantages that make them indispensable for screening purposes (see Verhulst & Koot, 1992, p. 67). Our early gates assessed both teacher and parent ratings of disruptive behavior using

the Conners Hyperactivity Index. Because of its brevity, ease of completion, sound psychometric properties, widespread usage, and ability to identify a mixed sample of aggressive and hyperactive children, the Hyperactivity Index was considered ideal for large-scale screening. Teachers, as first-line screeners, identified approximately 9% of the population. Adding parent ratings of behavior (and parent consent) as a second screen narrowed the risk group significantly (5.5%), and in doing so identified more severely and pervasively maladjusted children. The third and final gate employed a rating scale of perceived family discipline practices, a variable suspected of mediating the progression of early disruptive behavior to more serious CD. This gate resulted in a substantial reduction in the number of high-risk children identified (0.6% of the original population). Despite representing only a small percentage of the total population, the children in this group showed a risk profile that is consistent with CD risk factors documented in the literature (see Loeber & Dishion, 1983; Loeber & Stouthamer-Loeber, 1986, for reviews). This small group showed high levels of inattention, hyperactivity, aggression, and conduct problems as rated by both parents and teachers. These behaviors were associated with a pattern of serious and pervasive maladjustment that clearly distinguished high-risk from low- and moderate-risk counterparts. In addition to being defined by poor discipline practices, a substantial proportion of the variance in high risk group membership was accounted for by other parenting process variables including impaired involvement/communication, ineffective coping, external locus of control, and maternal depression.

The present study differs from much of the CD research in that a suburban population was studied. Although all levels of SES were represented, the majority of families were middle SES. The lower-SES families that did exist were geographically and socially embedded in affluent, upwardly mobile communities. This ecological context provided a unique opportunity to assess the epidemiological characteristics and prognostic significance of known risk indicators in a highly "buffered" setting. Many of the psychosocial risk variables that have been linked to CD, such as neighborhood decay and violence, racial discrimination, severe family stress (see Rutter & Giller, 1983), and serious economic disadvantage, including substandard housing and large family size (Farrington, 1987), were virtually nonexistent in our population. Nevertheless, we did find some evidence that lower SES (based on occupational status and educational attainment) and single parenthood were associated with having a troublesome child. Neither of these variables, however, distinguished high- from moderate-risk families. These findings offer some evidence to reject a simple linear relationship between these psychosocial variables and aggressive behavior.

The causal pathways of CD are undoubtedly complex, with the effect of variables such as SES and single parenthood mediated in large part by the extent to which they set in motion specific micropathological processes that have direct impacts on the child's psychological development. We found some evidence, for example, that the most significant predictors of high risk group membership, and conceivably the most powerful predictors of CD development, were variables related to family process, namely, ineffective discipline, poor family communication and involvement, poor maternal coping skills, and an external parent locus of control. Within this context, it is also interesting to note that there was a relatively strong positive relationship between maternal depression and higher levels of risk. At least one recent study, however, found that the association between maternal depression and childhood disruptive behaviors was the result of psychosocial factors (e.g., social disadvantage, marital discord) common to both outcomes (Fergusson & Lynskey, 1993).

Several theoretical models provide roles for unskilled parenting practices (Patterson, Capaldi, & Bank, 1991) and maternal depression (Snyder, 1991) as important mechanisms involved in the development of childhood disruptive behavior, particularly in how ineffective discipline and maternal distress may interfere with caregiving during the early years. Ineffective parenting behaviors and maternal depression, however, may be spuriously associated with child conduct problems as a consequence of another causal factor. Frick *et al.* (1992), for example, found that the variance in CD was negligible when the effect of parental antisocial personality disorder (APD) was controlled. This might suggest the importance of a genetic predisposition as a partial explanation for CD in some populations. Further, the evidence that ADHD and major depressive disorder are variable phenotypic expressions of the same genotype casts a new perspective on the relationship between maternal depression and child behavior problems (Biederman *et al.*, 1992).

There are several limitations that may have affected the utility of the present multiple-gate screening procedure. A parent self-report measure (BMSA) was developed to function as a brief screener for parent discipline practices. The self-report nature of this scale may have produced distortions in ratings as a result of social desirability rating bias. As a consequence, the present parent practice ratings may have underestimated the true degree of unskilled discipline in the home. Direct home observations, laboratory-based analogues, or structured interviews may have provided more accurate assessments, but at the cost of more intrusive and labor-intensive efforts. Among those eligible for Gate 3 screening, approximately 40% failed to return completed BMSAs. It is possible that parents experiencing the most difficulties in management practices were those who

failed to complete the scales. Our noncompliant group had the lowest SES rating and approximately one-third were single-parent families, a percentage higher than that observed in the high risk group. The targeted children in this group had lower IQs than those in the high-risk group but showed similar levels of teacher-rated attention problems, hyperactivity, and conduct problems. In all likelihood, a significant number of these cases would have screened positive for deficient discipline practices and conferred high risk membership if the parent had completed the assessment. Thus, our 0.6% high risk estimate appears to be an underestimate of the percentage of true high-risk cases in this population. Better methods for obtaining consent and facilitating participation in assessment will increase the precision of high risk identification.

Despite these limitations, the present research offers guidance for the selection of participants for preventive intervention trials. Should an intervention target all children who screen positive for disruptive behavior or just the subsample with deficient parenting practices (see Sandler, Braver, Wolchick, Pillow, & Gersten, 1991)? The solution may be dictated by the expense of the intervention and the intensity required to produce significant change in the targeted mediator. For example, an inexpensive and less intensive intervention can be offered to children who screen positive for early misbehavior but screen negative for parenting difficulties (i.e., moderate risk). A more costly and intensive program may be required for children who screen positive for both variables (i.e., high risk). Our future work will involve updating the utility of our multigating system, discriminating variables with the highest validity for the progression toward CD and assessing the capacity of our mediator-based selection measure to identify participants who respond differentially to interventions of various intensity.

## APPENDIX

### *Behavioral Management Self Assessment (BMSA)*

1. When I ask my child to do something, I am clear and to the point in my request.
2. During the day, I try to take notice when my child is being good and let him/her know I like how he/she is behaving.
3. When my child gives me a hard time ("whining, yelling") after I ask him/her to do something, I give up because it is too much of a hassle to continue.
4. I praise my child for doing something I like or approve of.
5. I am not consistent in disciplining my child.

6. I do a good job of keeping track of my child's misbehavior.
7. To change my child's undesirable behavior, I try to correct little problems first and gradually work up to what I want him/her to do.
8. When I have had a problem with my child, I set aside some time so that we can talk about the problem together.
9. I have to nag and/or scold my child to get him/her to do something I have asked.
10. When my child fails to do what I ask, I end up doing it.
11. When I punish my child I do it quickly, and do not let things get out of hand.
12. I am firm and consistent in disciplining my child.
13. I threaten my child if he/she does not do what I want.
14. I yell or scream at my child when he/she gets on my nerves.
15. When I give my child commands, I do not follow through to see that he/she obeys.

## REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

Barkley, R. A. (1988). Child behavior rating scales and checklists. In M. Rutter, A. H. Tuma, & I. S. Lann (Eds.), *Assessment and diagnosis in child psychopathology* (pp. 113- 155). New York: Guilford Press.

Biederman, J., Faraone, S. V., Keenan, K., Benjamin, J., Krifcher, B., Moore, C., Spich-Buckminster, S., Ugaglia, K., Jellinke, M. S., Steingard, R., Spencer, T., Norman, D., Kolodny, R., Kraus, I., Perrin, J., Keller, M. B., & Tsuang, M. T. (1992). Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder. *Archives of General Psychiatry, 49,* 728-738.

Campis, L. K., Lyman, R. D., & Prentice-Dunn, S. (1986). The parental locus of control scale: Development and validation. *Journal of Clinical Child Psychology, 15,* 260-267.

Costello, E. J. (1989). Developments in child psychiatric epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry, 28,* 836-841.

Dryfoos, J. G. (1990). *Adolescents at risk.* New York: Oxford University Press.

Farrington, D. (1987). Early precursors of frequent offending. In J. Q. Wilson & G. C. Loury (Eds.), *From children to citizens: Families schools and delinquency prevention* (pp. 27-50). New York: Springer-Verlage, pp. 27-50.

Frick, P. J., Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., Christ, M. G., & Hanson, K. (1992). Familial risk factors to oppositional defiant disorder and conduct disorder: parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology, 60,* 49-95.

Fergusson, D. M., & Lynskey, M. T. (1993). The effects of maternal depression on child conduct disorder and attention deficit behaviors. *Social Psychiatry and Psychiatric Epidemiology, 28,* 116-123.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on Revised Conners Parent and Teacher Rating Scales. *Journal of Abnormal Child Psychology, 6,* 221-236.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System: Manual.* Circle Pines, MN: American Guidance Service.

Hollingshead, A. (1975). *Four Factor Index of Social Status.* Unpublished working paper.

Jastak, G., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test — Revised.* Wilmington, DE: Jastak Associates.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test (K-BIT).* Circle Pines, MN: American Guidance Service.

Loeber, R., & Dishion, T. J. (1983). Early predictors of male delinquency: A review. *Psychological Bulletin, 74,* 68-99.

Loeber, R., Dishion, T. J., & Patterson, G. R. (1984). Multiple gating: A multistage assessment procedure for identifying youths at risk for delinquency. *Journal of Research in Crime and Delinquency, 21,* 7-31.

Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. in M. Tonry & N. Norris (Eds.), *Crimes and justice: An annual review of research* (Vol. 7, pp. 29-149). Chicago: University of Chicago Press.

Muehrer, P., & Koretz, D. S. (1992). Issues in preventive intervention research. *Current directions in psychological Science, 1,* 109-112.

Moos, R., & Moos, B. S. (1981). *Family Environment Scale: Manual.* Palo Alto: Consulting Psychologists Press.

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., Ryles, J. A., Crawford, J. W., Blum, H. M., Byrne, C., Thomas, H., & Woodward, C. A. (1987). Ontario child health study — II. Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry, 44,* 832-836.

Patterson, G. R., Capaldi, D., & Bank, L. (1991). An early starter model for predicting delinquency. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of child aggression* (pp. 139-168). Hillsdale, NJ: Erlbaum.

Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher Ratings of DSM-III-R symptoms for the disruptive behavioral disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 210- 218.

Piers, E. V. (1984). *Piers-Harris Children's Self-Concept Scale — revised: Manual.* Los Angeles: Western Psychological Services.

Reynolds, C. R., & Kamphaus, R. W. (1992). *BASC: Behavioral Assessment System for Children.* Circle Pines, MN: American Guidance Service.

Roberts, M. W., Joe, V. C., & Rowe-Hallbert, A. (1992). Oppositional child behavior and parent locus of control. *Journal of Clinical Child Psychology, 21,* 170-177.

Rutter, M., & Giller, H. (1983). *Juvenile delinquency: Trends and perspectives.* New York: Penguin Books.

Sandler, I. N., Braver, S. L., Wolchik, S. A., Pillow, D. R., & Gersten, J. C. (1991). Small theory and the strategic choices of prevention research. *American Journal of Community Psychology, 19,* 873-879.

Snyder, J. (1991). Discipline as a mediator of the impact of maternal stress and mood on child conduct problems. *Development and Psychopathology, 3,* 263-276.

Strayhorn, J. M., & Weidman, C. S. (1988). A parent practices scale and its relation to parent and child mental health. *Journal of the American Academy of Child & Adolescent Psychiatry, 27,* 613- 618.

Verhulst, F. C., & Koot, H. M. (1992). *Child psychiatric epidemiology: concepts, methods and findings.* Newbury Park, CA: Sage.

Walker, H. M., & McConnell, S. R. (1988). *Walker-McConnell Scale of Social Competence and School Adjustment.* Austin, TX: Pro-Ed.

Woodcock, R. W., & Johnson, M. B. (1990). *Woodcock-Johnson Psychoeducational Battery — Revised: Tests of achievement.* Allen, TX: DML Teaching Resources.