

Assessment of Parenting Practices in Families of Elementary School-Age Children

Karen K. Shelton, Paul J. Frick, and Jane Wootton

Department of Psychology, University of Alabama

Tested multimethod and multi-informant assessment of parenting practices in families of clinic-referred children between the ages of 6 and 13 ($n = 124$) and in families of community volunteer children ($n = 36$) who were comparable to the clinic group on age and sex of child, family ethnicity, and parental marital status. In general, children's report was not useful for assessing the parenting constructs using either a global report format or multiple telephone interviews. This was especially true for younger children (below age 9) and for child report on the telephone interviews, whereby children tended to respond using a consistent response set. In contrast, both assessment formats for obtaining parental report showed good utility. Reports from parents (in most cases the child's mother) generally were not strongly associated with measures of socially desirable responding, and parental report showed expected age trends and expected associations with socioeconomic status. Most important, both parental report formats were useful for differentiating families of children with disruptive behavior disorders (defined by teacher report alone) from families of normal volunteer children screened for disruptive behavior disorders.

There is a substantial literature linking several specific dimensions of parenting to disruptive child behaviors (for review, see Frick, 1994; Loeber & Stouthamer-Loeber, 1986; Steinmetz, 1979). Loeber and Stouthamer-Loeber (1986) conducted a meta-analysis of over 300 studies and found that the strongest and most consistent effects were for measures of parental monitoring and supervision of their child and measures of parental involvement with their child. The importance of these two dimensions of parenting has been supported by several more recent studies (Cernkovich & Giordano, 1987; Dishion, Patterson, Stoolmiller, & Skinner, 1991; Frick et al., 1992; Laub & Sampson, 1988; Van Voorhis, Cullen, Mathers, & Garner, 1988; Wilson, 1987). Parental discipline practices have also been linked to disruptive child behaviors. Specifically, inconsistent use of discipline, failure to use positive change strategies (e.g., positive reinforcement for appropriate behavior), and excessive use of corporal punishment have been linked to disruptive child behavior (Bierman & Smoot, 1991; Frick et al., 1992; Laub & Sampson, 1988; Patterson, Dishion, & Bank, 1984; Wells & Rankin, 1988). These aspects of discipline take on added importance because the most successful interventions for childhood behavior problems focus on improving these parenting practices (see Kazdin, 1995).

Although the association between parenting practices and child behavior problems is indisputable, how parenting fits into causal theories of child behavior problems is hotly debated (Dodge, 1990; Frick & Jackson, 1993; Lytton, 1990). As a result, more sophisticated research is needed to test alternative causal models. However, this research has been hampered by inadequate and inconsistent methods for assessing parenting in school-age samples. The most common method of assessing parenting is by asking a single informant, either parent or child, one or two questions related to each parenting construct. For example, the modal number of questions used to assess the constructs of parental involvement, supervision, harsh discipline, and inconsistent discipline across the studies reviewed by Loeber and Stouthamer-Loeber (1986) was one.

There are several standardized questionnaires used to assess family functioning in the research literature (e.g., Epstein, Baldwin, & Bishop, 1983; Moos & Moos, 1981; Roberts, Block, & Block, 1984). However, these questionnaires measure "parenting style," or the emotional climate in the home, rather than parenting practices (for a discussion of the importance of this distinction, see Darling & Steinberg, 1993). One standardized questionnaire that was designed to measure parenting practices is the Child's Report of Parental Behavior Inventory (CRPBI; Schaefer, 1965). It includes items that assess parental involvement, positive parenting, and consistency in discipline. However, the CRPBI lacks items that assess harsh discipline and

Requests for reprints should be sent to Paul J. Frick, Department of Psychology University of Alabama, P.O. Box 870348, Tuscaloosa, AL 35487

parental supervision. Also, its reliance on child self-report may make it inappropriate for younger school-age children.

Even if a questionnaire with adequate content existed, some researchers have questioned the reliance on this methodology to assess parenting practices (e.g., Zahn-Waxler, 1993). The difficulty, however, is determining viable alternatives or adjunctive methods of assessment. Observing parent-child interactions either in the home or in the clinic has been popular in the study of parenting behaviors in preschool and young school-aged children (e.g., Forehand & McMahon, 1981; Patterson, 1982; Robinson & Eyberg, 1981). However, behavioral observations may not be as useful for studying parenting in families of older school-age children. First, the reactivity to observation seems to increase with the age of the child, making the observation of parent-child interactions of older children less ecologically valid (see Keller, 1986). Second, it is difficult in the laboratory or in the natural setting to set up situations that elicit the parenting behaviors most important in older children (e.g., being left without adult supervision). Third, even if there were a valid method of conducting behavioral observations, the costly nature of such systems make them prohibitive for many research and clinical applications.

Recognizing the limitations of observational measures for this age group, Patterson and colleagues at the Oregon Social Learning Center (OSLC) developed telephone interviews to assess the frequency with which parents engaged in different parenting practices over discrete time periods, such as in the last 24 hours or over the past 3 days (Capaldi & Patterson, 1989; Patterson, Reid, & Dishion, 1992). The phone interviews were repeated several times to gather an adequate sample of parenting behaviors. Unfortunately, the OSLC system had a number of limitations. Most important, the assessment of constructs was not uniform either across constructs or across methods. For example, different time frames were used to assess the various parenting constructs within the interview format. Different answer formats were used even within the same parenting construct. Further, different questions were used to assess the same construct (e.g., supervision) across method (i.e., interview and questionnaire) and across informant (i.e., parent and child). This methodology makes it difficult to evaluate and compare the psychometric properties of the various assessment modalities.

To overcome these limitations, Frick (1991) developed an assessment system modeled after the OSLC system called the Alabama Parenting Questionnaire (APQ). Like the OSLC system, it was designed to tap the most important aspects of parenting practices related to disruptive behavior problems in children: parental involvement, monitoring/supervision, use of positive parenting techniques, inconsistency in disci-

pline, and harsh discipline. Also like the OSLC system, it was designed to assess these parenting constructs across multiple sources (parent and child) and using multiple formats (global report on questionnaires and behavior frequencies via telephone interviews). However, in developing the APQ, questions were designed to be analogous across all assessment sources and answer formats. As a result, the APQ allows for a comparison across multiple informants and across multiple assessment modalities.

In this article, we present the initial development of the APQ in a sample of elementary school-aged children and their primary custodial caretaker (in most cases, the child's mother). The goal of the study was to compare the assessment of parenting practices across informants and across methods using several indices of reliability and validity. In terms of reliability, the internal consistency of rationally derived scales that tap the relevant dimensions of parenting was tested to determine (a) whether the items seem to be measuring a homogenous construct, and (b) whether the construct can be measured with similar item content across informants and methods of assessment. Also, the generalizability of both parental and child report across repeated administrations of the telephone interview was tested to determine if a stable estimate of parenting practices can be obtained using this assessment format.

Also in this study, we began the process of establishing the validity of the APQ scores. First, we tested the convergent validity of APQ scores across informants and assessment methods and the divergent validity of the scores within an assessment method but across parenting domains. Second, we tested the degree to which a response set for providing socially desirable responses may have influenced APQ scores. Third, we tested whether or not parenting, as measured by the APQ, is correlated with age of the child and socioeconomic status in ways predicted by past research (Wauchope & Straus, 1990). Fourth, we tested whether or not the APQ scores differentiated families with children with behavior disorders from normal control families. This last test was considered the most important test of the validity of APQ scores because the APQ was designed for use in research on the relation between parenting practices and disruptive behavior disorders in children. Therefore, our sampling strategy, selection of measures, and procedures were designed primarily to provide an adequate test of this aspect of the validity of the APQ scores.

Method

Participants

Participants were 160 children ranging in age from 6 to 13 and their primary custodial caretakers recruited

from two sources. The first group were 124 caretakers, 95% of whom were mothers or mother figures, and children drawn from 134 consecutive referrals to the Alabama School-Aged Assessment Service (ASAS) between January, 1991 and May, 1994. The ASAS is a university-based outpatient diagnostic and referral service for children with behavioral, emotional, or learning disorders. Children who scored in the mentally retarded range ($n = 8$) or who had not lived with at least one primary caretaker within the last month ($n = 2$) were excluded from the study. Demographic characteristics of this clinic-referred sample are summarized in Table 1.

The second sample was a volunteer sample that was recruited to serve as a comparison group for the clinical sample. To enhance the comparability of the volunteer sample to the clinical sample, recruitment was conducted in two phases. The first phase was done through newspaper announcements and presentations at Parent-Teacher Association meetings at local schools. Parents of 6- to 13-year-old children were asked to participate in a study of "typical parenting practices." All volunteers were accepted at this stage. The next phase of recruitment involved targeting specific schools to recruit families of lower socioeconomic status and minority families to have the volunteer sample more closely approximate the clinic sample. Also, only families with male children were recruited at this stage, again to more closely approximate the predominantly male clinic sample. In both phases, all children who participated received a gift certificate from a local fast-food restaurant, and parents were eligible to win one of two \$100 gift certificates at a local department store.

Table 1. *Demographic Characteristics of the Sample*

Characteristic	Clinic	Volunteer
<i>n</i>	124	36
Sex (% Male)	81	73
Age		
<i>M</i>	8.7	9.1
<i>SD</i>	2.0	2.4
Ethnicity (% Caucasian)	75	76
Duncan's Socioeconomic Index		
<i>M</i>	37.5	47.8*
<i>SD</i>	24.7	24.2
Parental Marital Status (% Married)	46	59
Number of Children in the Home		
<i>M</i>	2.7	2.16*
<i>SD</i>	1.7	.83
Full Scale IQ		
<i>M</i>	94.5	—
<i>SD</i>	13.6	—

Note: Full Scale IQ was measured with the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974) for the first 102 consecutive referrals and the Wechsler Intelligence Scale for Children-Third Edition (Wechsler, 1991) for the last 25 clinic-referred children.

* $p < .05$.

The demographic characteristics of the volunteer sample are also summarized in Table 1. The recruitment procedures led to two samples that did not differ significantly on age (independent sample t test) or on sex, ethnicity, and parental marital status (χ^2). However, the volunteer sample consisted of families from higher socioeconomic statuses, measured by Duncan's Socioeconomic Index (Hauser & Featherman, 1977), and the volunteer sample had fewer children in the home. Also, all caretakers in the volunteer sample were mothers or mother figures.

Measures

APQ. An initial item pool for the APQ (Frick, 1991) consisted of items that had been used in past research to assess parental involvement with their child, use of positive reinforcement, monitoring and supervision of the child, consistency in applying discipline, and use of corporal punishment. The main sources of these items were studies reviewed by Loeber and Stouthamer-Loeber (1986), measures reported by Patterson and colleagues (Capaldi & Patterson, 1989), unpublished interviews developed by Loeber and his colleagues (Loeber, Stouthamer-Loeber, Van Kammen, & Farrington, 1987), and the CRPBI (Schaefer, 1965). After deleting redundant items, 35 items assessing the five constructs of interest remained. The items assessing each construct are presented in Table 2. Also included in the APQ items were 7 additional items measuring specific discipline practices other than corporal punishment. These items were included so that corporal punishment items were not asked in isolation of other forms of discipline, which could place an implicit negative bias toward these items.

Items were placed into four assessment formats: parent and child global report forms, and parent and child telephone interviews. Child-report items were all worded to refer to parenting in general within the family (e.g., "How often are you out with friends your parents do not know?"). The only exceptions were the items measuring involvement, which were repeated once with the child answering for his or her mother and once answering for his or her father. All items were analogous across each format.

After initial item construction, items were read by several parents and read to several children between the ages of 6 and 13 years, and changes were made to improve the clarity of items. Items on the global report forms were designed to be rated on a 5-point frequency scale ranging from 1 (*never*) to 5 (*always*) to represent the "typical" frequency in the home. Each item on the telephone interviews was designed to be answered with the respondent's best estimate of the number of occurrences of that behavior over the previous 3 days. The average frequency of each item across the four inter-

Table 2. Alabama Parenting Questionnaire: Scale Composition

Involvement
1. You have a friendly talk with your child.
4. You volunteer to help with special activities that your child is involved in (e.g., sports, Boy/Girl Scouts, church youth groups).
7. You play games or do other fun things with your child.
9. You ask your child about his/her day in school.
11. You help your child with his/her homework.
14. You ask your child what his/her plans are for the coming day.
15. You drive your child to a special activity.
20. You talk to your child about his/her friends.
23. Your child helps plan family activities.
26. You attend PTA meetings, parent/teacher conferences, or other meetings at your child's school.

Positive Parenting
2. You let your child know when he/she is doing a good job with something.
5. You reward or give something extra to your child for obeying you or behaving well.
13. You compliment your child when he/she does something well.
16. You praise your child if he/she behaves well.
18. You hug or kiss your child when he/she has done something well.
27. You tell your child that you like it when he/she helps around the house.

Poor Monitoring/Supervision
6. Your child fails to leave a note or to let you know where he/she is going.
10. Your child stays out in the evening past the time he/she is supposed to be home.
17. Your child is out with friends you do not know.
19. Your child goes out without a set time to be home.
21. Your child is out after dark without an adult with him/her.
24. You get so busy that you forget where your child is and what he/she is doing.
28. You don't check that your child comes home from school when he/she is supposed to.
29. You don't tell your child where you are going.
30. Your child comes home from school more than an hour past the time you expect him/her.
32. Your child is at home without adult supervision.

Inconsistent Discipline
3. You threatened to punish your child and then do not actually punish him/her.
8. Your child talks you out of being punished after he/she has done something wrong.
12. You feel that getting your child to obey you is more trouble than it's worth.
22. You let your child out of a punishment early (e.g., lift restrictions earlier than you originally said).
25. Your child is not punished when he/she has done something wrong.
31. The punishment you give your child depends on your mood.

Corporal Punishment
33. You spank your child with your hand when he/she has done something wrong.
35. You slap your child when he/she has done something wrong.
38. You hit your child with a belt, switch, or other object when he/she has done something wrong.

Other Discipline Practices
34. You ignore your child when he/she is misbehaving.
36. You take away privileges or money from your child as a punishment.
37. You send your child to his/her room as a punishment.
39. You yell or scream at your child when he/she has done something wrong.
40. You calmly explain to your child why his/her behavior was wrong when he/she misbehaves.
41. You use time out (make him/her sit or stand in corner) as a punishment.
42. You give your child extra chores as a punishment.

Note: The items in the table are worded according to the Parent Global report form.

views was used to compute scale scores. Therefore, scales scores were the average frequencies of behaviors within a 3-day interval across the four telephone interviews.

APQ data on some participants were not used in analyses. First, data on phone interviews were not used unless the participant completed at least three of the four scheduled interviews. This led to the elimination of 9 (7%) parent phone interviews and 14 (11%) child phone interviews in the clinic sample. Second, an additional 12 (10%) child phone interviews in the clinic sample and 3 (8%) child phone interviews in the normal sample were judged to be unusable because the child gave obviously deviant responses. This was operationally defined as a child reporting that a behavior had occurred more than 100 times on at least one item or at least 80 times on two or more items.¹ Third, 8 cases (5%) in the clinic sample involved the child's father, who was the primary caretaker, completing the APQ. These cases were eliminated in between-group analyses because all cases in the volunteer sample involved the child's mother or mother figure completing the APQ.

Global questionnaires were read to all children in both groups to ensure that reading level did not affect children's responses. The questionnaires were read by research assistants trained in standardized administration procedures who were instructed to read the printed instructions and APQ questions verbatim and not to help the child in interpreting questions. Phone interviews were administered by the same research assistants with similar instructions for standardized administration procedures. They were also trained to request specific number of occurrences from the parent and child for each question. Interviews were conducted at least 3 days apart over a 2- to 4-week period. Parents and children were always interviewed on the same day. Research assistants were not blind to whether the child was a member of the clinic or volunteer sample. However, within the clinic group, assistants were blind to the diagnostic status of the child and his or her reason for referral.

Disruptive behavior disorder diagnoses. One of the primary ways in which the validity of the APQ was assessed was by testing the association of the APQ subscales with a diagnosis of a disruptive behavior disorder (DBD) in the child. All children in the clinic sample were assessed for the presence of attention-deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), and conduct disorder (CD), ac-

cording to the criteria of the third and revised edition of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R;* American Psychiatric Association, 1987) using the NIMH Diagnostic Interview Schedule for Children-Version 2.3 (DISC-2.3; Shaffer, Fisher, Piacentini, Schwab-Stone, & Wicks, 1991). The DISC-2.3 is a structured psychiatric interview that was administered to each child's parent (DISC-P) and teacher (DISC-T).

Interviews were administered by advanced graduate students in clinical psychology or a licensed psychologist trained in assessing childhood psychopathology and in standardized DISC administration procedures. The same interviewer conducted the DISC-P and DISC-T for each child. Forty-seven (38%) of the DISC-P interviews were observed through one-way mirrors. Observers independently coded parents' responses, and the kappa statistic was calculated to determine interviewer and observer agreement on both the symptom and diagnostic levels for all cases in which the symptom or diagnosis was coded as present at least twice by the primary interviewer. Kappas for a diagnosis of ADHD and CD were both 1.0 and the kappa for a diagnosis of ODD was .93. Kappas for the ADHD symptoms ranged from .76 to 1.0 (median = 1.0). Kappas for the CD symptoms ranged from .79 to 1.0 (median κ = 1.0). Kappas for the ODD symptoms ranged from .91 to 1.0 (median κ = 1.0).

Because the same questions were asked of the child's teacher by the same interviewer, it is likely that the high reliability found for the parent interview also applies to the teacher interview. However, all teachers in the clinic sample also completed the Comprehensive Behavior Rating Scale for Children (CBRSC; Neeper, Lahey, & Frick, 1990). Children with a teacher diagnosis of ADHD, CD, or ODD were compared with all other clinic-referred children on the relevant CBRSC scales, to determine if the teacher's report of disruptive behaviors was consistent across assessment modalities. Children with a DBD diagnosis according to teacher report on the DISC had significantly higher scores on the Inattention-Disorganization, $t(124) = 4.24, p < .001$, Motor Hyperactivity, $t(124) = 11.03, p < .001$, and Oppositional/Conduct Disorders, $t(124) = 8.20, p < .001$, scales of the CBRSC.

Conducting DISC interviews in the volunteer sample was not feasible. However, the sample was screened for DBD diagnoses for between-group comparisons using the DBD Rating Scale (DBDRS; Pelham, Gnagy, Greenslade, & Milich, 1992) completed by each child's parent and teacher. The DBDRS asks the respondent to rate on a 4-point scale ranging from 0 (*not at all*) to 3 (*very much*) the degree to which a child displays each symptom of ADHD, CD, and ODD. The DBDRS, and its *DSM-III* predecessor, the SNAP Checklist (Atkins, Pelham, & Licht, 1985), have been widely used in

¹Children who gave these obviously deviant responses were significantly younger than the rest of the sample, $t(161) = 2.90, p < .001$. Six were age 6, five were age 7, and one child each was age 8, 9, 10, or 11.

research to screen for DBDs in school-age samples. It has been shown to be both reliable and valid for this purpose (Pelham et al., 1992).

Pelham et al. (1992) found that considering only symptoms endorsed very much by an informant as being significant seemed to be optimal for representing DBD diagnoses in a school-age sample. However, because in this study we wanted to ensure a normal control sample for between group comparisons, a more lenient criterion was used. Symptoms were considered present if they were rated as either pretty much (2) or very much (3). Any child who had sufficient symptoms rated at this level of severity to meet *DSM-III-R* criteria for any DBD diagnosis by either parent or teacher ($n = 7$; 19%) was eliminated from the volunteer sample for between-group analyses.

Procedure

The measures used in this study were included as part of a comprehensive psychological evaluation for children in the clinic-referred group. Upon their initial visit to the clinic, all referred children and their primary custodial parents were asked to give consent for the use of assessment data in research. They were told that their willingness to participate in research would in no way affect the clinical services that they received. None of the parents or children refused to participate in the study. Following informed consent, the parents were administered a semistructured interview to obtain demographic information and were administered the DISC-P. Following the DISC-P, the parents completed the APQ global report form. While the parent data were being collected, children were administered an intelligence test to screen for mental retardation and then were administered the APQ global report form. The child's teacher was contacted and administered the DISC-T by telephone within the week following the ASAS evaluation. Parent and child phone interviews were also initiated within the week following the evaluation. The interviews were conducted at least 3 days apart over a 2- to 4-week period.

In the volunteer sample, parents and children who agreed to participate were mailed a consent form, a sheet requesting demographic information, the DBDRS, and the parent APQ global report form. Following return of the consent forms, the child's teacher was mailed the DBDRS with a self-addressed return envelope. The children were contacted and completed the APQ global report form by phone. The initial phone interviews with the parent and child were then scheduled. The time frame for completing the interviews (at least 3 days apart over a 2- to 4- week period) was the same as in the clinic sample.

Results

Reliability of APQ Scores

The first analyses tested the internal consistency of the rationally derived scales of APQ to see if (a) the APQ items were measuring relatively homogeneous constructs, and (b) whether a similar scale structure could be used across informants to measure the same construct. Internal consistency estimates in the full sample and in the clinic and volunteer samples are presented in Table 3. Several trends emerged from these data. First, the three-item Corporal Punishment (CP) scale consistently had poor internal consistency across the samples and assessment formats. Second, the child interview format showed the highest internal consistency across scales. Third, the Poor Monitoring/Supervision (MS) scale showed low internal consistency in the parent interview format. Fourth, inspections of item-total correlations suggested that eliminating any of the item(s) would not have appreciably increased the internal consistency of scales in any assessment format.

The temporal stability of the interview format was tested by determining the consistency of the scale scores across the four interview times. The interviews were conducted at least 3 days apart across a 2- to 4-week time period. The consistency estimate was the coefficient alpha for the scores across the four interview times. These estimates are also reported in Table 3. Generally, scores were highly consistent across interview times, suggesting that the interviews provided a stable estimate of the construct. The lowest consistency was found for parent report on the MS and the CP scales.

Validity of APQ Scores

In Table 4, the convergent validity of APQ scores across informant and assessment methods for each parenting construct are presented. The Involvement (IN) and Positive Parenting (PP) constructs generally showed the highest correlations across informant and methods of assessment. Across the five constructs, the highest correlations, all of which were statistically significant, tended to be within each informant but across methods. For parents, the correlations for a construct across global report and interview formats ranged from .30 to .55 ($M = .37$). For children, the correlations across method of assessment ranged from .26 to .46 ($M = .35$). The next highest correlations tended to be across informants but within the same mode of assessment. Parent-child correlations within the global report format ranged from .08 to .28 ($M = .19$). Parent-child correlations within the interview format ranged from .08 to .32 ($M = .21$). The lowest

correlations, showing statistical significance in only 3 of 10 correlations, were the correlations that crossed both informant and method of assessment, ranging from $-.05$ to $.30$ ($M = .10$).

The correlations among APQ scales within each assessment format were also calculated as a test of the scales divergent validity. The MS, Inconsistent Discipline (ID), and CP scales generally showed good divergent validity. Specifically, the average intercorrelations for these three scales (ignoring direction of correlations) within the parent global report format was $.19$ (range = $.08$ – $.27$), within the parental interview format was $.16$ (range = $.01$ – $.56$), and within the child global report format was $.13$ (range = $.01$ – $.48$).² In contrast, the PP and IN scales were highly correlated across informant and assessment format ($.41$ – $.85$; $M = .67$), suggesting that these may be measuring a single dimension of parenting. All scales from the child interview format were highly intercorrelated ($.33$ – $.86$; $M = .64$). Not only were the correlations high within this assessment method, but the two positive practices scales were positively correlated with the three negative practices scales. This pattern suggests that children tended to use a response set in responding to interview questions, answering either high or low on all items.

This response set likely accounts for the high internal consistency estimates found for the child interview scales. We repeated these correlations within the child interview format dividing the sample into younger (below age 9; $n = 86$) and older children (age 9 or older; $n = 74$) and found similar correlations in both age groups.

²The most appropriate test of the multi trait multi method correlation matrix would have been through structural equation modeling, in which the variance attributable to the parenting constructs could have been estimated, and the variance could have been accounted for by method (both informant and assessment technique). However, our sample was not large enough to conduct such an analysis. As a preliminary test of the construct validity of the APQ scales, we tested the difference between the average correlations within each method (divergent validity) with the average correlations within each construct across methods (convergent validity). Initially there were no significant differences found. However, when the high correlations between the two positive parenting scales were excluded from these analyses, the average correlations within the parent global report (.19) differed significantly from the average correlation within each construct across the two parent report methods (.37; $z = 1.83$, $p < .05$). Similarly, the average correlation within the parent interview method (.16) was significantly lower than this estimate of convergent validity ($z = 1.98$, $p < .05$). Also, the average correlation within the child global report format (.13) was significantly lower than the average correlation within each construct across the two child report methods (.35; $z = 1.95$, $p < .05$). As would be expected given the high correlations within the child interview format, even after eliminating the correlation between the two positive parenting scales, the average correlation within this assessment format was higher than the estimate of convergent validity. Therefore, these results indicate good divergent validity for the APQ scales with the exception of the two positive parenting scales and all scales using the child interview format.

This response set seriously compromises the validity of child report using the interview format. We tested whether or not interviewers could detect this response set in children's responses. At the end of each child interview, the interviewer rated his or her perceptions of the accuracy of the child's responses on a 1 (*very inaccurate*) to 5 (*very accurate*) scale. Thirty-five children (27%), 28 in the clinic sample and 7 in the volunteer sample, were judged to be very inaccurate on at least one interview. The children judged to be inaccurate were much younger ($M = 7.4$ years, $SD = 1.6$) than the remainder of the sample, $t(161) = 6.05$, $p < .001$. The correlations among the scales after eliminating these young children reduced the response set somewhat, although the scales using the child interview format still tended to be highly intercorrelated.

The correlations between APQ scales and the child's age are presented in Table 5. All parenting constructs showed a significant association with age in at least two assessment formats. The most consistent age trends were the negative correlations between age and the IN, PP, and CP scales. The MS scale was positively correlated with age on both parent report measures, as would be predicted. However, children's report on the MS scale using the interview format was negatively correlated with age, again calling into question the validity of this method in obtaining child report. Correlations between APQ scales and socioeconomic status (SES) were less consistent (see Table 5). For both assessment formats, parental report of IN was positively associated with SES and parental report of CP was negatively correlated with family SES as would be predicted.

Potential differences on the APQ scales between the two ethnic groups represented in the sample (Caucasian $n = 123$, African American $n = 37$) were also tested. However, because the African American children came from homes with lower SES scores, $t(155) = 3.59$, $p < .001$, scores on the parenting scales were compared across ethnic groups only after controlling for SES scores. There were several differences on the APQ scales across ethnic groups. African American parents reported lower scores on the IN scale using either the global report, $F(2, 151) = 3.78$, $p < .05$) or phone interview, $F(2, 143) = 3.63$, $p < .05$, formats. Furthermore, African American parents reported significantly higher scores on the MS scale using the global report format, $F(2, 151) = 20.57$, $p < .001$. African American parents also reported higher scores on the ID scale using the interview format, $F(2, 143) = 4.66$, $p < .05$ and African-American children reported higher scores on the ID scale using the global report format, $F(2, 147) = 8.05$, $p < .01$.

APQ scales were correlated with an index designed to detect a tendency to give socially desirable re-

Table 3. Internal Consistency of Items on the APQ Scales and Temporal Stability of Scales Across the Four Interviews

APQ Scale	Coefficient Alpha (Unstandardized)			
	Internal Consistency			Temporal Stability (Full Sample)
	Full Sample	Clinic	Volunteer	
Parent Global				
<i>n</i>	160	124	36	
Involvement	.80	.80	.80	
Positive Parenting	.80	.79	.85	
Poor Monitoring/Supervision	.67	.63	.75	
Inconsistent Discipline	.67	.64	.74	
Corporal Punishment	.46	.45	.49	
Parent Interview				
<i>n</i>	152	116	36	152
Involvement	.76	.76	.77	.89
Positive Parenting	.89	.89	.91	.88
Poor Monitoring/Supervision	.21	.37	—	.66
Inconsistent Discipline	.76	.76	.64	.85
Corporal Punishment	.09	.27	.22	.69
Child Global				
<i>n</i>	155	122	33	
Involvement (Mother)	.72	.72	.71	
Involvement (Father)	.83 (<i>n</i> = 120)	.83 (<i>n</i> = 91)	.83 (<i>n</i> = 29)	
Positive Parenting	.74	.75	.72	
Poor Monitoring/Supervision	.69	.68	.66	
Inconsistent Discipline	.56	.53	.66	
Corporal Punishment	.44	.41	.58	
Child Interview				
<i>n</i>	129	96	33	129
Involvement (Mother)	.91	.92	.90	.87
Involvement (Father)	.88 (<i>n</i> = 104)	.90 (<i>n</i> = 76)	.75 (<i>n</i> = 28)	.70 (<i>n</i> = 104)
Positive Parenting	.90	.88	.93	.81
Poor Monitoring/Supervision	.93	.94	.87	.80
Inconsistent Discipline	.90	.88	.95	.83
Corporal Punishment	.81	.81	.81	.81

Note: Internal consistency estimates were based on Cronbach's alpha for unstandardized variables. Coefficient alpha was also used to assess the temporal stability of the scales across the four telephone interviews. Differences in sample sizes for the Involvement (Father) scale were a result of eliminating children without a male caretaker in the home. Alphas less than zero are not reported.

sponses and these correlation coefficients are also reported in Table 5. The K-scale of the Minnesota Multiphasic Personality Inventory—Second Edition (MMPI-2; Hathaway & McKinley, 1989) was used to assess this response set in mothers, and the Lie scale of the Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985) was used to assess this response set in children. These measures were completed by participants in the clinic sample only. The correlations suggest that a socially desirable response set was not substantially influencing APQ scores for most scales. Two notable exceptions were the moderate and significant correlations with the social desirability index for the ID (−.41) and CP (−.40) scales using the parent interview format. The ID scale by parental global report and Maternal Involvement (M-IN) by child report were also significantly correlated with social desirability indices (−.24 and .23, respectively).

The final analyses were conducted to determine the association between the APQ scales and diagnoses of DBDs. Only information obtained from female caretakers (95% of the clinic sample and 100% of the volunteer sample) were used in this analysis because there were no fathers reporting in the volunteer sample. Because the APQ scales tended to have skewed distributions and because the between-group analyses involved unequal group sizes, nonparametric procedures were chosen for these analyses. First, an *elevation* on each APQ scale was defined as a score more than 1 SD below the mean of the full volunteer sample for the positive parenting scales (IN and PP) or greater than one standard deviation above the mean for the negative parenting scales (MS, ID, or CP). Second, the seven children in the volunteer sample who met our lenient screening criteria for DBDs were eliminated from the normal control group for all analyses. Third, different criteria for forming the DBD group within the clinic sample were used

for analyses of child and parent report on the APQ. Specifically, for child report, both parent and teacher reports of DBD diagnoses were used to form the DBD group, as recommended by Piacentini, Cohen, and Cohen (1992). However, to avoid inflated associations due to shared method variance (i.e., parent reporting on diagnoses and parenting) only the teacher reports of DBD diagnoses were used to form groups when parent report on the APQ was tested.

Based on this procedure, maternal report of parenting was compared in families of clinic-referred children with a DBD diagnosis defined by teacher report ($n = 64$) and in a normal control group of children screened for DBD behaviors ($n = 29$). The DBD group was lower on Duncan's SEI, $t(87) = 2.89, p < .01$, and had fewer girls, $\chi^2(1, N = 93) = 3.8, p < .05$, than the normal control group. Therefore, SES and sex were controlled in logistic regression analyses. Analyses of elevations on parent global report scales indicated that children in the DBD group were more likely to have one or more elevations across the five APQ scales than the normal

control group, Wald $\chi^2(1, N = 93) = 5.07, p < .05$, controlling for both SES and sex. The results are illustrated in Figure 1. This finding was largely due to differences in the number of elevations on the three negative parenting scales across groups. When elevations on either of the positive parenting scales were analyzed separately, logistic regression analyses did not reveal an effect of group, Wald $\chi^2(1, N = 93) = .37, ns$. However, an effect for group was found for elevations on the three negative parenting scales, Wald $\chi^2(1, N = 93) = 9.44, p < .01$.

Analyses using scales based on maternal report using the interview format revealed almost identical results. In these analyses, elevations on the MS scale were not included due to its poor reliability. Again, logistic regression analyses revealed an effect for group on the likelihood of having an elevation on any of the four APQ scales included in analyses, Wald $\chi^2(1, N = 93) = 3.84, p < .05$. These results are illustrated in Figure 2. However, as was the case for parental global report, the group differences were primarily due to elevations on the negative parenting scales, Wald $\chi^2(1, N = 93) = 5.3, p < .05$.

To test the validity of children's report of parenting, children with a DBD diagnosis by either parent or teacher report ($n = 95$) were compared with the normal control group ($n = 29$). The DBD group had significantly lower Duncan's SEI ratings, $t(118) = 2.48, p < .01$, and therefore, SES was controlled in logistic regression analyses. Although there was a tendency for families in the DBD group to have more elevations on the APQ scales than the normal control group, this did not reach statistical significance in logistic regression analyses of either the global report format, $\chi^2(1, N = 124) = .74, ns$ or the interview format, $\chi^2(1, N = 124) = 2.88, ns$. The results for the interview format were repeated eliminating those children whose responses were judged to be invalid by the interviewer and the logistic regression analysis again revealed no significant group differences on the likelihood of having elevations on the APQ scales, $\chi^2(1, N = 124) = 2.36, ns$.

Table 4. Cross-Method and Cross-Informant Correlations of APQ Scales

	Parent Global	Parent Interview	Child Global
Involvement (Mother)			
Parent Global ^a	1.00		
Parent Interview ^b	.55***	1.00	
Child Global ^c	.23**	.09	1.00
Child Interview ^d	.22**	.22**	.35***
Involvement (Father)			
Child Global ^e			1.00
Child Interview ^f			.46***
Positive Parenting			
Parent Global	1.00		
Parent Interview	.32***	1.00	
Child Global	.25**	.18*	1.00
Child Interview	.30***	.32***	.30***
Poor Monitoring/Supervision			
Parent Global	1.00		
Parent Interview	.33***	1.00	
Child Global	.08	.06	1.00
Child Interview	-.05	.08	.26**
Inconsistent Discipline			
Parent Global	1.00		
Parent Interview	.30***	1.00	
Child Global	.10	-.05	1.00
Child Interview	.02	.12	.33***
Corporal Punishment			
Parent Global	1.00		
Parent Interview	.37***	1.00	
Child Global	.28***	.11	1.00
Child Interview	.15	.30***	.40***

Note: Differences in sample sizes for the Involvement (Father) scale were a result of eliminating children without a male caretaker in the home.

^a $n = 160$. ^b $n = 152$. ^c $n = 155$. ^d $n = 129$. ^e $n = 120$. ^f $n = 104$.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

In this study, we tested a multi-informant and multi-method system for assessing parenting practices in the families of school-age children. This system, the APQ, was specifically designed to assess parenting constructs using analogous items across different informants and assessment formats. As a result, this system allows for a direct comparison of the reliability and validity of various methods of assessing parenting practices in the families of elementary school-age children. Therefore, these data have important implications not only for the

Table 5. Correlations Between APQ Scales and Age of Child, Socioeconomic Status, and Indexes of a Socially Desirable Response Set

Scale	Age		Duncan's SEI		MMPI-2 K-Scale		RCMAS Lie Scale	
	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>	<i>r</i>	<i>n</i>
Involvement (Mother)								
Parent Global	-.16*	160	.16*	160	.19	81		
Parent Interview	-.28***	152	.16*	152	.02	80		
Child Global	-.12	155	-.01	155			.23*	121
Child Interview	-.39***	129	-.15	129			.06	95
Involvement (Father)								
Child Global	-.04	120	.14	120			.10	91
Child Interview	-.33***	104	-.06	104			.04	76
Positive Parenting								
Parent Global	-.26***	160	-.05	160	.15	81		
Parent Interview	-.24**	152	.12	152	.07	80		
Child Global	-.14	155	.04	155			.17	121
Child Interview	-.33***	129	-.06	129			.02	95
Poor Monitoring/Supervision								
Parent Global	.25**	160	-.08	160	.07	81		
Parent Interview	.30***	152	.14	152	-.12	80		
Child Global	.12	155	-.06	155			.01	121
Child Interview	-.27**	129	-.06	129			-.08	95
Inconsistent Discipline								
Parent Global	.02	160	-.06	160	-.24*	81		
Parent Interview	-.06	152	-.15	152	-.41***	80		
Child Global	-.18*	155	-.02	155			-.01	121
Child Interview	-.36***	129	-.12	129			-.07	95
Corporal Punishment								
Parent Global	-.11	160	-.18*	160	-.14	81		
Parent Interview	-.22**	152	-.21**	152	-.40***	80		
Child Global	-.26**	155	-.07	155			.13	121
Child Interview	-.32***	129	-.09	129			-.03	95

Note: MMPI-2 = Minnesota Multiphasic Personality Inventory—Second Edition (Hathaway & McKinley, 1989). RCMAS = Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1985). Both measures of a socially desirable response set were given only in the clinic sample. Duncan's SEI = Duncan's Socioeconomic Index (Hauser & Featherman, 1977). Differences in sample sizes for the Involvement (Father) scale were a result of eliminating children without a male caretaker in the home.

* $p < .05$. ** $p < .01$. *** $p < .001$.

use of the APQ in future projects but for determining the optimal method for assessing parenting practices in families of school-age children.

Despite the focus on child report of parenting practices in adolescent samples (see Loeber & Stouthamer-Loeber, 1986), these data call into question the utility of child report in younger samples. First, a number of the very young children (especially ages 6–7) gave such deviant responses on the telephone interview that their responses could not be used in analyses. Second, even after eliminating children with these clearly deviant responses, children tended to respond to interview questions regarding the frequency of parental behavior using a consistent response set, giving either very high frequencies or very low frequencies of behavior. Interviewers were only moderately successful in detecting this response set, which was generally found in the youngest children (ages 6–8). Third, although the child global report format seemed less susceptible to this response set, possibly

because of the limited choices in responses, parenting constructs based on this format still failed to differentiate families of children with DBD diagnoses and volunteer families.

In contrast, both the global and interview formats seemed to be useful for obtaining parental reports of parenting practices. Most importantly, both formats distinguished families of children with DBD diagnoses and normal control families, consistent with a great deal of past research (Cernkovich & Giordano, 1987; Dishion et al., 1991; Frick et al., 1992; Laub & Sampson, 1988; Loeber & Stouthamer-Loeber, 1986; Van Voorhis et al., 1988; Wilson, 1987). Importantly, this association between parenting and DBD diagnoses could not be attributed to shared method variance, because the DBD diagnoses were formed by teacher report alone for these analyses. We feel that this evidence for the validity of the parent report formats of the APQ is the most crucial piece of information, given that the APQ was specifically designed for use in research

to further study the association between parenting and DBDs in families of school-age children.

One exception to these generally favorable findings for the parental report formats was the assessment of parental monitoring and supervision using the interview format. The MS scale had quite low internal consistency among items, despite an adequate number

of items to measure the construct, and there was low temporal stability for this construct over the four interviews. It seems that this construct, which is primarily composed of low base-rate behaviors (see Table 2), is not adequately captured within the 3-day time window used by the APQ telephone format even across four time periods. Instead, the larger time window afforded by the

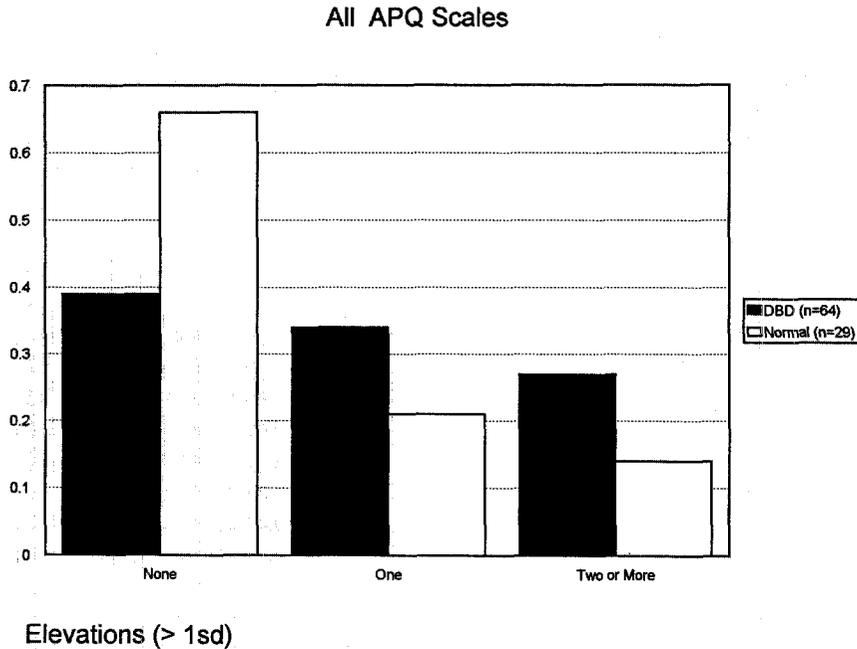


Figure 1. Group differences in the proportion of children with elevations across the five APQ scales using the APQ maternal global report format.

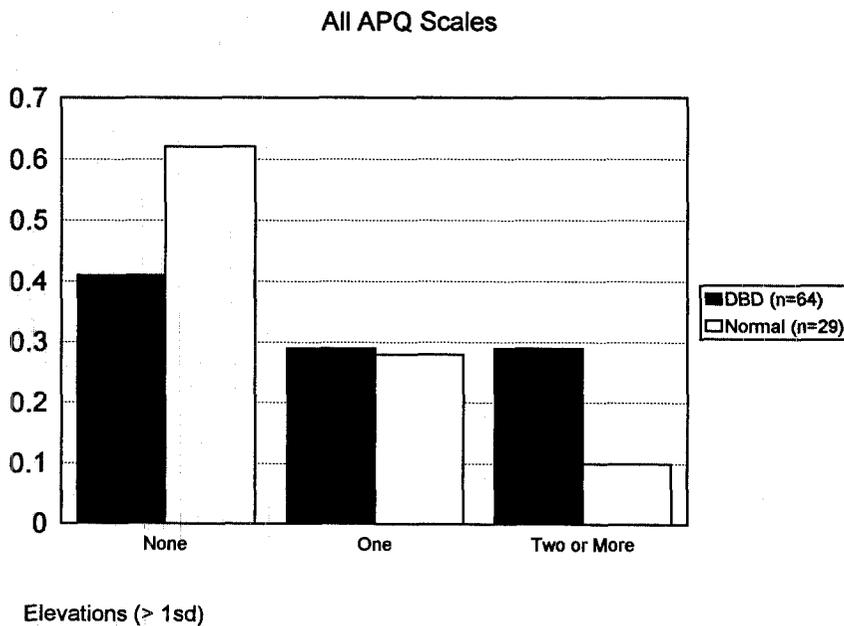


Figure 2. Group differences in the proportion of children with elevations across the five APQ scales using the APQ maternal telephone interview format.

global format, which asks for ratings of "typical" practices, seems to be necessary to adequately assess this construct.

The CP scale also had poor internal consistency for both interview and global formats for parents. However, its low internal consistency is likely to be largely due to the scale having only three items. Consistent with this possibility, across both parent-report formats the CP scale showed the expected negative correlations with age and SES (Wauchope & Strauss, 1991). Also, the CP scale contributed to the ability of the APQ to distinguish between families with DBD children and normal control families. Therefore, these pieces of evidence for the validity of the CP scale suggest that the internal consistency estimates may have underestimated its reliability. It may be that parents tend to use one preferred method of corporal punishment and as a result, there is not a high degree of intercorrelation among the corporal punishment items.

One puzzling finding was that, in contrast to past research (Loeber & Stouthamer-Loeber, 1986), the two positive parenting scales did not contribute substantially to the validity of the APQ to differentiate families with DBD children from non-DBD volunteer families. Our findings indicate that this could not be attributable to socially desirable responding. Instead, the discrepancy with past research is more likely to be a function of the age of the sample and the use of parent report. Specifically, past studies that found an association between parental involvement and child conduct problems generally used child report measures in adolescent samples (see Table 1 in Loeber & Stouthamer-Loeber, 1986). Therefore, the differences may reflect developmental differences in the association of positive parenting practices with childhood DBD and/or suggest that the association is unique to adolescents' perceptions of parental involvement. Importantly, the two positive parenting scales did detect developmental changes in parenting practices. Both the PP and IN scales were negatively correlated with age across both parental report modalities. This pattern is consistent with past research showing decreases in positive parenting practices and increases in parent/child distance as the child approaches adolescence (see Paikoff & Brooks-Gunn, 1991).

All of these results must be interpreted within the context of several methodological issues. These data were based largely on a clinic-referred sample. The volunteer sample was small and recruited specifically to be a comparison group for the clinic-referred sample. As a result, these data are based largely on families of Caucasian clinic-referred boys from lower to lower-middle SES. The generalizability to other more diverse samples requires further testing, especially in light of several differences in parenting practices across ethnic groups. Also, both parental report formats were ob-

tained primarily from mothers or other female caretakers who made up 95% of the clinic sample and 100% of the volunteer sample. To make the clinic and volunteer sample more comparable, only those cases with maternal report in the clinic sample were used in between group validation analyses. Therefore, the generalizability of the findings to fathers' perceptions of parenting is questionable.

To be useful in many research and clinical contexts, the APQ must be tested in larger more representative community samples that would provide better normative information on parenting practices as assessed by the APQ. Also, the sample size in this study was too small to determine if there was factor analytical support for the rationally derived scales on the APQ. Based on the internal consistency estimates and the scale intercorrelations, it is likely that the construct validity of the scales would be confirmed by factor analysis with two exceptions. The two positive parenting scales were highly intercorrelated (see Footnote 2) and therefore, seem to assess a single dimension of positive parenting. Also, the poor differentiation among items on the child interview format, due the problematic response set, would likely prevent differentiation among constructs within this report format.

With these cautions in mind, we feel that these findings have important implications for assessing parenting practices in families of school-age children. They indicate that, with a few exceptions previously noted, parent report using either the global or interview formats is more useful for assessing parenting practices than child report in this age group. These findings hopefully will set the stage for further refinement of this and other standardized methods of assessing parenting practices that are desperately needed for advancing research on the association between parenting practices and behavior problems in children (Frick & Jackson, 1993). As research expands in this area, it hopefully will refine the assessment process for both research and clinical uses. For example, although both the global and interview format from parental report seemed to be valid in documenting differences in families of children with DBDs and families of normal control children, the interview format, with its focus on a more discrete time period, may prove to be better in detecting changes over time, such as after a family-based treatment program or in longitudinal studies.

References

- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed, rev.). Washington, DC: Author.
- Atkins, M. S., Pelham, W. E., & Licht, M. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology*, *13*, 155-167.

ASSESSMENT OF PARENTING

- Bierman, K. L., & Smoot, D. L. (1991). Linking family characteristics with poor peer relations: The mediating role of conduct problems. *Journal of Abnormal Child Psychology, 19*, 341-356.
- Capaldi, D. M., & Patterson, G. R. (1989). *Psychometric properties of fourteen latent constructs from the Oregon Youth Study*. New York: Springer-Verlag.
- Cernkovich, S. A., & Giordano, P. C. (1987). Family relationships and delinquency. *Criminology, 25*, 295-321.
- Darling, N., & Steinberg, L. (1993). Parenting style as context: An integrative model. *Psychological Bulletin, 113*, 487-496.
- Dishion, T. J., Patterson, G. R., Stoolmiller, M., & Skinner, M. L. (1991). Family, school and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology, 27*, 172-180.
- Dodge, K. A. (1990). Nature versus nurture in childhood conduct disorder: It is time to ask a different question. *Developmental Psychology, 26*, 698-701.
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy, 9*, 171-180.
- Forehand, R., & McMahon, R. J. (1981). *Helping the non-compliant child: A clinician's guide to parent training*. New York: Guilford.
- Frick, P. J. (1991). *The Alabama Parenting Questionnaire*. Unpublished instrument, University of Alabama.
- Frick, P. J. (1994). Family dysfunction and the disruptive behavior disorders: A review of recent empirical findings. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in clinical child psychology* (Vol. 17, pp. 203-226). New York: Plenum.
- Frick, P. J., & Jackson, Y. K. (1993). Family functioning and childhood antisocial behavior: Yet another reinterpretation. *Journal of Clinical Child Psychology, 22*, 410-419.
- Frick, P. J., Lahey, B. B., Loeber, R., Stouthamer-Loeber, M., Christ, M. A. G., & Hanson, K. (1992). Familial risk factors to conduct disorder and oppositional defiant disorder: Parental psychopathology and maternal parenting. *Journal of Consulting and Clinical Psychology, 60*, 49-55.
- Greenbaum, P. E., Dedrick, R. F., Prange, M. E., & Friedman, R. M. (1994). Parent, teacher, and child ratings of problem behaviors of youngsters with serious emotional disturbances. *Psychological Assessment, 6*, 141-148.
- Hathaway, S. R., & McKinley, J. C. (1989). *Minnesota Multiphasic Personality Inventory—2nd edition*. Minneapolis: University of Minnesota.
- Hauser, R. M., & Featherman, D. L. (1977). *The process of stratification*. New York: Academic.
- Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence*. New York: Sage.
- Keller, H. R. (1986). Behavioral observation approaches to personality assessment. In H. M. Knoff (Ed.), *The assessment of child and adolescent personality* (pp. 353-390). New York: Guilford.
- Laub, J. H., & Sampson, R. J. (1988). Unraveling families and delinquency: A reanalysis of the Gluecks' data. *Criminology, 26*, 355-379.
- Loeber, R., & Stouthamer-Loeber, M. (1986). Family factors as correlates and predictors of juvenile conduct problems and delinquency. In M. Tonry & N. Morris (Eds.), *Crime and justice* (Vol. 7, pp. 29-149). Chicago: University of Chicago Press.
- Loeber, R., Stouthamer-Loeber, M., Van Kammen, W. B., & Farrington, D. P. (1987). *Parent and child interviews*. Unpublished instrument for the Pittsburgh Youth Study, Pittsburgh, PA.
- Lytton, H. (1990). Child and parent effects in boys' conduct disorder: A reinterpretation. *Developmental Psychology, 26*, 683-697.
- Moos, R. H., & Moos, B. (1981). *Family Environment Scale manual*. Palo Alto, CA: Consulting Psychologists Press.
- Neeper, R., Lahey, B. B., & Frick, P. J. (1990). *The Comprehensive Behavior Rating Scale for Children manual*. San Antonio, TX: Psychological Corporation.
- Paikoff, R. L., & Brooks-Gunn, J. (1991). Do parent-child relationships change during puberty? *Psychological Bulletin, 110*, 47-66.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.
- Patterson, G. R., Dishion, T. J., & Bank, L. (1984). Family interaction: A process model of deviancy training. *Aggressive Behavior, 10*, 253-267.
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.
- Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 210-218.
- Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones. *Journal of Abnormal Child Psychology, 20*, 51-63.
- Reynolds, C. R., & Richmond, B. O. (1985). *Revised Children's Manifest Anxiety Scale—RCMAS*. Los Angeles: Western Psychological Services.
- Roberts, G. C., Block, J. H., & Block, J. (1984). Continuity and change in parents' child-rearing practices. *Child Development, 55*, 586-597.
- Robinson, E. A., & Eyberg, S. M. (1981). The Dyadic Parent-Child Coding System: Standardization and validation. *Journal of Consulting and Clinical Psychology, 49*, 245-250.
- Schaefer, E. S. (1965). Children's reports of parental behavior: An inventory. *Child Development, 36*, 417-424.
- Shaffer, D., Fisher, P., Piacentini, J., Schwab-Stone, M., & Wicks, J. (1992). *Diagnostic Interview Schedule for Children, Version 2.3*. New York: Columbia University.
- Steinmetz, S. K. (1979). Disciplinary techniques and their relationship to aggressiveness, dependency, and conscience. In W. R. Burr, R. Hill, F. I. Nye, & I. L. Reiss (Eds.), *Contemporary theories about the family* (Vol. 1, pp. 405-438). New York: Free Press.
- Van Voorhis, P., Cullen, F. T., Mathers, R. A., & Garner, C. C. (1988). The impact of family structure and quality on delinquency: A comparative assessment of structural and functional factors. *Criminology, 26*, 235-259.
- Wauchope, B., & Straus, M. A. (1990). Physical punishment and physical abuse of American children: Incidence rates by age, sex, and occupational class. In M. A. Straus & R. J. Gelles (Eds.), *Physical violence in American families* (pp. 133-148). New Brunswick, NJ: Transaction Publishers.
- Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children—Revised (WISC-R)*. New York: Psychological Corporation.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children—3rd edition (WISC-III)*. New York: Psychological Corporation.
- Wells, L. E., & Rankin, J. H. (1988). Direct parental controls and delinquency. *Criminology, 26*, 263-285.
- Wilson, H. (1987). Parental supervision re-examined. *British Journal of Criminology, 27*, 275-301.
- Zahn-Waxler, C. (1993). Warriors and worriers: Sex and psychopathology. *Developmental Psychopathology, 1/2*, 79-90.

Received April 25, 1995

Final revision received October 10, 1995