

The Differential Validity of Teacher Ratings of Inattention/Overactivity and Aggression

Marc S. Atkins,^{1,4} William E. Pelham,² and Mark H. Licht³

The ability of school-based measures of child behavior to predict differentially the IOWA Conners teacher rating scale Inattention/Overactivity (IO) and Aggression (A) factors was evaluated in a sample of 71 school-aged boys. Multiple measures and multiple methods of assessment were utilized to provide a comprehensive assessment of social and academic behavior. These included direct observations of children in the classroom and on the playground, examination of the organization of children's desks, measures of academic performance, peer nominations of popularity and rejection, and sociometric ratings using the Pupil Evaluation Inventory. Despite moderately high correlations between the teacher rating factors ($r = .60$), considerable evidence was provided for differential validity on measures of academic performance, peer rating measures, and measures of disruptive or inappropriate classroom and playground behavior. These differences between IO and A factors indicated that the factors were logically independent in important ways, supporting prior work validating separate hyperactivity and aggression dimensions.

Manuscript received in final form November 30, 1988.

The authors are grateful to Richard Milich, Joseph Torgeson, and Richard Tate for critical readings of this manuscript; to Charlotte Johnston for invaluable administrative assistance; to Susan Friedan, Stephanie Martinez, Nancy Hudgens, Denise Gammonsly, Kate Taylor, Roz Sweet, and Michelle Willis for assistance in data collection and computer input; and to Edward Vertuno, Marvel Lou Sanden, Sally Butzin, Eileen McDaniel, Sadette Lathrop, Judy Arnette, Gloria Poole, Anne Black, Dianne Ferrell, Patrica Tolbert, and Margie Rush of the Florida State University Developmental Research School for their outstanding cooperation and support.

¹Department of Pediatric Psychology, Children's Seashore House, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104.

²Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania 15213.

³Department of Psychology, Florida State University, Tallahassee, Florida 32306.

⁴Address all correspondence to Marc S. Atkins, Department of Pediatric Psychology, Room 2307, Children's Seashore House, Children's Hospital of Philadelphia, 34th and Civic Blvd., Philadelphia, Pennsylvania 19104.

Teacher ratings are useful for the identification of hyperactive and conduct-disordered children in both clinical and research settings because of the ease and economy of data collection and the unique perspective offered by a teacher regarding a child's performance on a variety of work and play tasks. To be maximally useful for diagnosis, however, teacher ratings should make global distinctions between deviance and normality as well as discriminate among different types of childhood disorders. The differentiation of hyperactivity and conduct disorder is of particular importance because groups of children identified by standardized teacher ratings as hyperactive have been shown to be heterogeneous regarding the presence of aggressive behavior (Hinshaw, 1987). Groups of hyperactive children with aggressive symptoms have a poorer prognosis and different associated variables at intake and follow-up than nonaggressive hyperactive children (Johnston & Pelham, 1986; McGee, Williams, & Silva, 1984; Milich, Oney, & Landau, 1982).

The most commonly used teacher rating scale of hyperactivity, the Conners Teacher Rating Scale (CTRS; Conners, 1969), provides largely overlapping Hyperactivity and Conduct Problems factors (Lahey, Green, & Forehand, 1980; Prinz, Connor, & Wilson, 1981). In recognition of this problem, Loney and Milich (1982) identified items on the CTRS correlated to external measures of hyperactivity but not to measures of aggression, and CTRS items related to measures of aggression but not to measures of hyperactivity. The resulting 10-item scale, the IOWA (Inattention Overactivity With Aggression) Conners Rating Scale, was evaluated further and shown to provide relatively independent scores on Inattention/Overactivity and Aggression as compared with similar factors on the CTRS (Loney & Milich, 1982). Further evaluations with this measure in clinically referred samples demonstrated considerable evidence for its validity as a measure of hyperactivity and aggression (e.g., Milich & Fitzgerald, 1985; Milich & Landau, 1988).

In contrast to evidence for the independence of the IOWA Inattention/Overactivity and Aggression factors in clinic samples, evidence for the independence of these factors in normative samples is less clear. For example, teacher-rated Hyperactivity and Aggression factors based on criteria similar to those used by Loney and Milich (1982) have been shown to be more highly correlated in nonproblem samples as compared with samples of clinically diagnosed hyperactive and aggressive children (O'Leary & Steen, 1982). Because rating scales are commonly used in nonclinical samples for screening and identification, it is important to evaluate the independence of these factors in nonreferred samples as well. To date, however, no study utilizing a nonreferred sample has investigated the relationship of the IOWA factors to measures of social and academic behavior, which are important characteristics of childhood hyperactivity and aggression and are thus the criteria against which rating scales should be validated.

The purpose of this investigation was to evaluate the utility of separate hyperactivity and aggression dimensions in a sample of school-age boys using teacher ratings derived from the IOWA Conners. Given the lack of widely agreed-upon specific criterion measures of hyperactivity and aggression (Pelham, 1982), multiple measures and multiple methods of assessment were utilized to provide a comprehensive assessment of potential differences between factors and a broader test of the differential validity of these ratings than has yet been reported.

METHOD

Subjects and Setting

Subjects were 71 boys in grades 1 through 5 (mean age = 106.4 months; range = 76 to 134 months) selected according to the following procedures. Ten boys were identified as extreme on teacher ratings using the DSM-III criteria for ADD during a prior investigation in this school (Atkins, Pelham, & Licht, 1985), and 12 boys were identified as verbally or physically disruptive (e.g., off task, interrupting, out of seat) on classroom observations prior to data collection. This procedure identified 22 ADD/disruptive boys.⁵ The remaining boys in the sample ($N = 58$) were selected randomly from the remainder of the school population with the exception of 14 boys who were excluded owing to low rates of verbal and physical behaviors during classroom observations to maximize the opportunity for obtaining a range of scores on teacher ratings. When combined with the ADD/disruptive boys, a total of 8 boys from each of 10 classrooms were included in the sample. Because these procedures identified all of the ADD/disruptive boys in this school but only 60% (58/97) of the remaining boys, 60% of the ADD/disruptive boys (13/22) were randomly selected so that the proportion of ADD/disruptive boys equaled the proportion of randomly selected boys. These subject selection procedures were necessary to ensure that a representative sample of ADD/disruptive boys would be included owing to the restricted range of ratings of hyperactivity and aggression found in samples of grade school children (Lambert, Sandoval, & Sassone, 1981; Schachar, Sandberg, & Rutter, 1986).

⁵It was not possible to identify aggressive children by observations owing to the low rate of aggressive behaviors in classroom settings. For example, in this sample, Physical Intrusion, Name Calling, and Threatening each occurred in less than 1% of observations during an entire school year (Atkins et al., 1988).

Assessment Measures⁶

Teacher Ratings. The CTRS was completed on each subject by his teacher, and the IOWA Conners Inattention/Overactivity (IO) and Aggression (A) scales were computed. Reliability and validity evidence for both scales indicated their appropriateness for the purposes of this study (see Atkins & Milich, 1988).

Classroom Observations of Conduct and Attention Deficit Disorder. This system (COCADD) was adapted from the Time Sample Behavioral Checklist (Paul, Power, Engle, & Licht, 1987) and consisted of 32 codes in five domains (Position, Physical-Social Orientation, Vocal Activities, Non-vocal Activities, Play Activities). Coding involved brief (2-sec) but frequent observations of each child. As shown in Table I, 16 scores were derived from these 32 codes. Twelve of these scores were composites derived by grouping conceptually similar codes. Scores consisted of the percentage of observation intervals in which each code or combination of codes was observed.

Desk Observations. Desk observations involved the unobtrusive examination of children's desks when they were out of the room. Items were derived from responses on a teacher questionnaire and interviews with each teacher concerning the behavioral components of children's classroom organization. The desk observations were divided rationally into two categories, neatness and preparedness, with five resulting dependent measures as shown in Table II. Neatness items referred to the arrangement of appropriate and absence of inappropriate materials in, on, and around the children's desks. Preparedness items referred to the presence of required books and supplies. Dependent measures were the percentage of observations each item was scored as present.

Academic Performance. Four measures of academic performance were obtained. All available daily written assignments were scored for percentage of assignments completed (ratio of completed assignments to assigned assignments), and percentage of problems correct (ratio of problems completed correctly to problems attempted). The other two measures were derived from standardized testing. One score was an achievement score expressed as a national percentile rank averaged over all academic categories on the Comprehensive Test of Basic Skills (McGraw-Hill, 1974), and the other score was an academic aptitude score from the Test of Cognitive Skills (McGraw-Hill, 1983).

Peer Ratings. Two types of peer ratings were obtained. A positive-negative peer nomination procedure was employed to assess popularity and

⁶Complete descriptions of codes, definitions, and assessment procedures are available from the first author.

Table I. Summary Definitions for COCADD Variables

Classroom observations

Attending: Attending to persons, situations, or materials necessary for the completion of the assigned task

Overactive: Running or climbing/hanging while not attending

Distracted: Listening while not attending

Verbal disruptive: Screaming, verbal intrusion, crying

Verbal off-task: Talking to others, talking to self, commanding or nonvocal communication while not attending

Verbal aggression: Cursing, name calling/teasing, or threatening

Physical aggression: Destroying property or physical intrusion

Stealing/cheating: Stealing or cheating

Playground observations

Verbal disruptive: Screaming, verbal intrusion, or crying

Verbal aggression: Cursing, name calling/teasing, or threatening

Physical aggression: Destroying property, or physical intrusion

Stealing cheating: Stealing, cheating, or rule breaking

High active play: Running or climbing/hanging

Solitary play: No shared materials or cooperative interaction

Parallel play: Shared materials without cooperative interaction

Group play: Cooperative interaction with or without shared materials

rejection (Roff, Sells, & Golden, 1972). Children chose three classmates whom they liked and three classmates whom they disliked. A class picture was used as a prompt for first-graders. Dependent measures were the number of same-sex positive and negative nominations received. A second peer rating measure, the Pupil Evaluation Inventory (PEI; Pekarik, Prinz, Liebert, Weintraub, & Neale, 1976), was also administered to obtain peer evaluations of target children's classroom behavior. On the PEI, every child in each class was asked to nominate those classmates who fit the description for each of 34 items describing a variety of positive and negative behaviors. Aggression, Withdrawal, and Likability factor scores (Pekarik et al., 1976), which were the percentage of same-sex peers who nominated a child on items of that factor, served as dependent measures. A 17-item version of the PEI was used for first-graders (Pekarik et al., 1976).

Table II. Summary Definitions for Desk Observation Variables

Neatness items

Neat desktop: Appropriate arrangement of books and papers and absence of trash or other inappropriate materials on desktop

Neat inside desk: Absence of trash or other inappropriate materials inside desk

Neat desk area: Absence of trash or other inappropriate materials in area surrounding desk

Preparedness items

Has books: Presence of required books in or on desk

Has materials: Presence of pencil and/or eraser in or on desk

Assessment Procedures

COCADD observations were collected by two paid professional observers who had worked with the code for a year prior to this study. Desk observations and scoring of academic work were collected by trained undergraduates. One week of COCADD and desk observations preceded data collection in each classroom. Positive and negative peer nominations were administered individually. PEI ratings were administered individually for first-graders and those second-graders designated as poor readers by their teachers. All other children completed the PEI independently in 1-hour class periods. For each peer rating instrument, standard instructions were used stressing the need for confidentiality and for careful responding (see Johnston, Pelham, Crawford, & Atkins, 1988). All measures were collected blind to subject selection criteria.

Each classroom was scheduled for 5 consecutive days of concurrent COCADD observations, desk observations, and scoring of daily academic work. The CTRS was completed during the observation week or the following week. Peer ratings were obtained the 1st or 2nd week following the collection of the other dependent measures. In each classroom, COCADD classroom observations were scheduled for 4 hours daily throughout the day (excluding special activities), and playground observations were scheduled for at least 30 minutes of outdoor free-play activity daily. Each child was scheduled for 30 daily COCADD observations in the classroom and 10 daily COCADD observations during playground free-play time, yielding an average of 150 classroom observations and 50 playground observations for each child during the week of data collection.

RESULTS

Interrater Reliability for Observational Measures

Interrater reliability assessments occurred on 10% of COCADD observations and 18% of desk observations. Agreement between raters was assessed by computing interrater Kappa coefficients and yielded a median Kappa of .86 (range of .67 to 1.00) for COCADD classroom measures and a median Kappa of .90 (range of .79 to 1.00) for COCADD playground measures. Three COCADD measures did not occur during reliability observations (Distracted, Verbal Disruptive on the playground, Stealing/Cheating on the playground). The median Kappa coefficient for the five desk observation measures was .67 (range of .63 to .79).

Partial Correlations and Differential Validity

The influence of the subjects' ages, which had a large range in this sample, was controlled by partialing it from all correlations presented below. Separate correlations were calculated between each IOWA factor and each COCADD classroom, COCADD playground, desk observation, academic performance, and peer rating variable to determine the degree to which these variables were associated with each of the factors. When a variable was significantly correlated with one or both IOWA factors, the significance of the difference between its correlation with each factor was tested using the r to Z transformation for same-sample correlations (McNemar, 1969). In addition, when a variable was significantly correlated with both IOWA factors, differential validity was further assessed by computing correlations between the variable and each factor partialing both age and the other factor from the relationship (Milich et al., 1982). These second-order partial correlations indicated the relationship of the IOWA factors to each of the variables, controlling for age and for the alternate IOWA factor. One COCADD variable was not observed (Stealing/Cheating on the playground) and was omitted from all analyses.

The IOWA IO and A factors were significantly intercorrelated ($r = .60, p < .01$), consistent with the correlation previously reported between these two factors in normative samples (Loney & Milich, 1982; Murphy, Pelham, & Milich, 1985). The correlation between variables and the IOWA factors, controlling for age, are presented in Table III.

Ten variables were correlated significantly with the IO factor and not with the A factor. Tests for differences between the correlation of these variables with the IO factor as compared with the A factor revealed significant differences for three academic performance variables (Percentage Correct, Achievement, Aptitude), two playground variables (High Active, Group Play), one desk variable (Neat Desktop), and two peer rating variables (Positive Nominations, PEI Withdrawal). Two variables were correlated significantly with the A factor and not with the IO factor (Verbal Aggression on the playground, Assignments Completed), but only one correlation (Assignments Completed) was significantly different from the correlation with the IO factor. Four variables were significantly correlated with both teacher rating factors. Of these, two variables correlated significantly more highly with the IO factor than with the A factor (Negative Nominations, PEI Likability), and one variable correlated significantly more highly with the A factor than with the IO factor (Verbal Disruptive in the classroom). All significant correlations were in the expected direction.

Table IV presents second-order partial correlations for the four variables that were significantly correlated with both teacher rating factors, con-

Table III. Correlations Between Variables and the IOWA Conners Inattention/Overactivity (IO) and Aggression (A) Factors^a

	IO	A	<i>t</i> ^b
COCADD Classroom			
Attending	.02	.13	—
Overactive	.23 ^c	.11	n.s.
Distracted	-.11	-.16	—
Verbal disruptive	.21 ^c	.41 ^d	2.00 ^d
Verbal off-task	-.07	-.06	—
Verbal aggressive	-.01	-.14	—
Physical aggressive	.01	.05	—
Stealing/cheating	.04	-.07	—
COCADD playground			
Verbal disruptive	-.05	-.16	—
Verbal aggressive	.09	.20 ^c	n.s.
Physical aggressive	.01	.11	—
High active	.36 ^d	.18	1.76 ^c
Solitary play	.05	-.15	—
Parallel play	.18	-.03	—
Group play	-.27 ^d	-.01	2.51 ^d
Desk observations			
Neat desktop (<i>n</i> = 63)	-.22 ^c	-.02	1.89 ^c
Neat inside desk (<i>n</i> = 63)	.08	.02	—
Neat desk area (<i>n</i> = 46)	.19	.00	—
Has books (<i>n</i> = 71)	-.09	.13	—
Has materials (<i>n</i> = 46)	-.29 ^c	-.17	n.s.
Academic performance			
Percentage correct	-.24 ^c	.13	3.70 ^d
Assignments completed	-.02	.27 ^d	2.82 ^d
Achievement	-.44 ^d	-.15	2.98 ^d
Attitude	-.34 ^d	.00	3.41 ^d
Peer ratings (<i>n</i> = 69)			
Positive nominations	-.44 ^d	-.08	3.77 ^d
Negative nominations	.57 ^d	.22 ^c	3.94 ^d
PEI aggression	.55 ^d	.56 ^d	n.s.
PEI withdrawal	.43 ^d	.19	2.43 ^d
PEI likability	-.64 ^d	-.26 ^c	4.59 ^d

^aAll analyses controlling for age. *N* = 71 unless otherwise noted.

^b*t* test for difference between correlations if one or both IOWA factors significantly related to variable.

^c*p* < .05.

^d*p* < .01.

trolling for age and for the alternate IOWA factor. Three variables remained significantly correlated with an IOWA factor after controlling for the alternate IOWA factor (Verbal Disruptive with A; Negative Nominations and PEI Likability with IO). One variable, the PEI Aggression factor, was correlated significantly with both IOWA factors after partialing out the alternate IOWA factor. All of these correlations were in the expected direction.

Table IV. Second-Order Partial Correlations Coefficients for Variables Significantly Correlated with Both the IOWA Conners Inattention/Overactivity (IO) and Aggression (A) Factors^a

	IO	A	<i>t</i> ^b
COCADD Classroom			
Verbal disruptive	-.05	.36**	4.27**
Peer ratings			
Negative nominations	.56**	-.17	12.09**
PEI aggression	.33**	.34**	n.s.
PEI likability	-.62**	.19	19.60**

^aAll correlations controlling for age and for the alternate IOWA factor.

^b*T* test for difference between correlations.

^c*p* < .05.

^d*p* < .01.

DISCUSSION

The results provided considerable evidence for the validity of the IOWA IO and A factors, especially considering the moderately high correlation between these factors. Sixteen of 29 variables were correlated significantly with one or both IOWA factors and 12 variables provided evidence for the differential validity of the IOWA factors as indicated by a significant difference between correlations. For three variables, differential validity was further indicated by a significant partial correlation for one factor controlling for the other factor when correlations with both factors were significant. Evidence for the importance of differences between teacher rating scores was indicated by examination of the relationship of the IOWA factors to specific variables. This is especially evident for the academic and peer rating variables.

Three of the four academic variables significantly differentiated the IO factor from the A factor, consistent with prior work demonstrating the relationship of measures of academic performance to psychiatric chart ratings of hyperactivity (Milich et al., 1982), and to teacher ratings of attention deficit disorder (Atkins et al., 1985; Atkins, Pelham, & Licht, 1988). The unique relationship of academic variables to the IO factor is an important validation of this factor because it is estimated that greater than 50% of children diagnosed as hyperactive have learning or achievement deficits (McGee & Share, 1988), and subgrouping hyperactive children on the dimension of concurrent learning problems has been recommended (Pelham, 1982). Our results highlight the importance of this issue by demonstrating that hyperactivity is uniquely related to a child's academic performance relative to aggression. Further, because the academic performance variables provided at least partially operationalized definitions for rating scale items related to inattention

(e.g., “fails to finish things he starts”), these results support the use of objective measures of academic performance to assess attentional deficits in the classroom.

Similar to the academic performance variables, considerable support for the differential validity of the IOW factors was provided by the peer rating variables. Evidence for unique variance was found for four of five peer rating variables with the IO factor. In addition, COCADD playground observations provided further evidence for the relationship between IOWA ratings and peer-related measures. Two playground measures (High Active, Group Play) were significantly more highly correlated with the IO factor than with the A factor, and one playground measure (Verbal Aggressive) was correlated with the A factor and not with the IO factor, although the difference between the correlations was not significant.

The relatively strong relationship between measures of peer relations and the IO factor as compared with the A factor suggests that the excessive or bothersome behaviors associated with peer rejection and low popularity are more highly related to hyperactivity than to aggression (Pelham & Milich, 1984). The one exception to this—the PEI Aggression factor was related to both IOWA factors—appears to be a function of items reflecting both hyperactivity (e.g., “Can’t sit still,” “Don’t pay attention”) and aggression (e.g., “Start fights,” “Say they can beat everybody up”) on the 21-item PEI Aggression factor. In a subsequent evaluation of these data, those items from the PEI Aggression factor conceptually similar to the IOWA IO factor were significantly more highly correlated with the IO factor than with the A factor (Atkins, Johnston, Pelham, & Licht, 1987).

The differential validity of the IOWA factors was less substantial for COCADD classroom or desk observation variables, as compared with academic and peer variables. However, the COCADD and desk variables did provide modest support for the differential validity of the IOWA factors and conformed to prior research. For example, two of five desk observation variables (Neat Desktop, Has Materials) were correlated significantly with the IO factor and not with the A factor. Further, in a prior study, when children identified as ADD by teacher ratings were compared with a control group of classroom peers, the desk variables contributed discriminant information independent of classroom observations and academic performance measures (Atkins et al., 1985).

For the COCADD classroom variables, one measure (Overactive) was correlated with the IO factor and not with the A factor, although the magnitude of difference was not significant, and one variable was related uniquely to an IOWA factor (Verbal Disruptive to the A factor). The unique relationship of Verbal Disruptive to the A factor in this study is interesting for several reasons. First, this result replicates similar findings by Abikoff, Gittelman,

and Klein (1980) comparing hyperactives with controls, as well as a previous study of ours (Atkins et al., 1985) comparing teacher-identified ADD children with controls. Our results suggest that it may have been the presence of aggressive children in each of these samples and not hyperactive/ADD children that accounted for the discrimination from controls. Second, this finding is especially noteworthy given the inclusion of a similar item ("Often interrupts or intrudes on others") in the DSM-III-R diagnostic criteria for Attention-Deficit Hyperactivity Disorder (American Psychiatric Association, 1987). The results suggest that verbal interruption might be more appropriately listed as a criterion for Oppositional Defiant Disorder in DSM-III-R. Third, this result supports the IOWA A factor as a measure of defiant and oppositional behavior rather than overt aggression, especially considering the omission of items related to overt aggression on the A factor.

The remaining COCADD classroom variables were not correlated significantly with either teacher rating factor, despite close conceptual relationships between many COCADD variables and items contained on the IOWA factors. For example, COCADD classroom variables assessing attentional problems (Distracted, Verbal Off-Task, Attending) were not correlated significantly with the IO factor, and direct measures of aggression (e.g., Verbal and Physical Aggressive) failed to correlate with the A factor. This highlights a common criticism of teacher ratings — that is, that they are highly influenced by subjective impressions to the exclusion of objective behavior — and reaffirms the importance of including alternative measures for diagnosis and assessment (Lahey et al., 1987; Loney, 1987).

However, it should be noted that two methodological issues may have also contributed to these low correlations. First, observations in the classroom occurred primarily during structured seat work activities, which may have obscured differences in less-structured task formats (Milich & Fitzgerald, 1985; Milich & Landau, 1988). Second, several of the COCADD codes measured low-frequency events that may not have been assessed adequately by the time-sampling strategy. The use of alternative sampling strategies for these low-frequency events, such as event sampling (see Mariotto & Licht, 1986), may have provided greater correspondence with IOWA IO and A ratings, especially given the potential influence of low-rate but highly salient events (e.g., physical assaults, destroying property) on these ratings (Schachar et al., 1986). In addition, this may explain the relatively few dependent measures related to the IOWA A factor as compared with the IO factor in this sample.

In summary, considerable evidence was provided for the differential validity of the IOWA IO and A factors. The IO factor was uniquely related to high activity in the classroom and on the playground, messy and disorganized desks without necessary materials, incorrect academic work, low aca-

demic achievement and aptitude, and peer relationships marked by low acceptance, high rejection, and withdrawal from group activities on the playground. The A factor was uniquely related to disruptive behavior in the classroom, completed academic assignments, and aggressive peer ratings. The many measures related uniquely to one or the other factor provided impressive evidence for the independence of the IOWA factors, especially given the overlap between factors as indicated by their moderately high intercorrelation. This reaffirms a point made repeatedly by Loney and Milich (e.g., Loney, 1987; Loney & Milich, 1982; Milich et al., 1982) that a moderately high correlation between factors does not preclude differential validity, and supports a now considerable body of work validating the utility of separate constructs for hyperactivity and aggression (Hinshaw, 1987).

REFERENCES

- Abikoff, H., Gittelman, R., & Klein, D. (1980). Classroom observation code for hyperactive children: A replication of validity. *Journal of Consulting and Clinical Psychology, 48*, 555-565.
- American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, DC: Author.
- Atkins, M. S., Johnston, C., Pelham, W. E., & Licht, M. H. (1987). *Differentiating hyperactivity from aggression on the Pupil Evaluation Inventory*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Boston.
- Atkins, M. S., & Milich, R. (1988). The IOWA Conners Rating Scale. In A. S. Bellack & M. Hersen (Eds.), *Dictionary of behavioral assessment techniques* (pp. 273-275). New York: Pergamon Press.
- Atkins, M. S., Pelham, W. E., & Licht, M. H. (1985). A comparison of objective classroom measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology, 13*, 155-167.
- Atkins, M. S., Pelham, W. E., & Licht, M. H. (1988). The development and validation of objective classroom measures for conduct and attention deficit disorders. In R. J. Prinz (Ed.), *Advances in behavioral assessment of children and families* (Vol. 4, pp. 3-31). Greenwich, CT: JAI Press.
- Conners, C. K. (1969). A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry, 126*, 884-888.
- Hinshaw, S. P. (1987). On the distinction between attentional deficits/hyperactivity and conduct problems/aggression in child psychopathology. *Psychological Bulletin, 101*, 443-463.
- Johnston, C., & Pelham, W. E. (1986). Teacher ratings predict peer ratings aggression at 3-year follow-up in boys with attention deficit disorder with hyperactivity. *Journal of Consulting and Clinical Psychology, 54*, 571-572.
- Johnston, C., Pelham, W. E., Crawford, J. J., & Atkins, M. S. (1988). A psychometric study of positive and negative nominations and the Pupil Evaluation Inventory. *Journal of Abnormal Child Psychology, 16*, 617-626.
- Lahey, B. B., Green, K. D., & Forehand, R. (1980). On the independence of ratings of hyperactivity, conduct problems, and attention deficits in children: A multiple regression analysis. *Journal of Consulting and Clinical Psychology, 48*, 566-574.
- Lahey, B. B., McBurnett, K., Piacentini, J., Hartdagen, S., Walker, J., Frick, P., & Hynd, G. (1987). Agreement of parent and teacher rating scales with comprehensive clinical assessments of attention deficit disorder with hyperactivity. *Journal of Psychopathology and Behavioral Assessment, 9*, 429-439.

- Lambert, N. M., Sandoval, J., & Sassone, D. M. (1981). Prevalence of hyperactivity and related treatments among elementary school children. In K. D. Gadow & J. Loney (Eds.), *Psychosocial aspects of drug treatment for hyperactivity* (pp. 249-291). Boulder, CO: Westview Press.
- Loney, J. (1987). Hyperactivity and aggression in the diagnosis of attention deficit disorder. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in clinical child psychology* (Vol. 10, pp. 99-135). New York: Plenum.
- Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. In M. Wolraich & D. K. Routh (Eds.), *Advances in behavioral pediatrics* (Vol. 2, pp. 113-147). Greenwich, CT: JAI Press.
- Mariotto, M. J., & Licht, M. H. (1986). Ongoing assessment of functioning with DOC systems: Practical and technical issues. In G. Paul (Ed.), *Assessment in residential treatment settings: Principles and methods to support cost-effective quality operations*. Champaign, IL: Research Press.
- McGree, R., & Share, D. L. (1988). Attention deficit disorder-hyperactivity and academic failure: Which comes first and what should be treated? *Journal of the American Academy of Child and Adolescent Psychiatry*, 27, 318-325.
- McGee, R., Williams, S., & Silva, P. A. (1984). Behavioral and developmental characteristics of aggressive, hyperactive, and aggressive-hyperactive boys. *Journal of the American Academy of Child Psychiatry*, 23, 270-279.
- McGraw-Hill. (1974). *Comprehensive tests of basic skills* (Form S). Monterey, CA: Author.
- McGraw-Hill. (1983). *Test of cognitive skills*. Monterey, CA: Author.
- McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York: Wiley.
- Milich, R., & Fitzgerald, G. (1985). Validation of inattention/overactivity and aggression ratings with classroom observations. *Journal of Consulting and Clinical Psychology*, 53, 139-140.
- Milich, R., & Landau, S. (1988). Teacher ratings of inattention/overactivity and aggression: Cross-validation with classroom observations. *Journal of Clinical Child Psychology*, 17, 92-97.
- Milich, R., Loney, J., & Landau, S. (1982). Independent dimensions of hyperactivity and aggression: A validation with playroom observation data. *Journal of Abnormal Psychology*, 91, 183-198.
- Murphy, D., Pelham, W. E., & Milich, R. (1985). *Reliability and validity data for the IOWA Conners Teacher Rating Scale*. Paper presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Houston.
- O'Leary, S. G., & Steen, P. L. (1982). Subcategorizing hyperactivity: The Stony Brook Scale. *Journal of Consulting and Clinical Psychology*, 50, 426-432.
- Paul, G. L., Power, C. T., Engel, K. L., & Licht, M. H. (1987). The Time-Sample Behavioral Checklist observer manual. In G. L. Paul (Ed.), *Observational assessment instrumentation for service and research—The Time Sample Behavioral Checklist: Assessment in residential treatment settings, Part 2* (pp. 11-50). Champaign, IL: Research Press.
- Pekarik, E. G., Prinz, R. J., Liebert, D. E., Weintraub, S., & Neale, J. M. (1976). The Pupil Evaluation Inventory: A sociometric technique for assessing children's social behavior. *Journal of Abnormal Child Psychology*, 4, 83-97.
- Pelham, W. E. (1982). Childhood hyperactivity: Diagnosis, etiology, nature, and treatment. In R. Gatchel, A. Baum, & J. Singer (Eds.), *Behavioral medicine and clinical psychology: Overlapping disciplines* (pp. 261-327). Hillsdale, NJ: Erlbaum.
- Pelham, W. E., & Milich, R. (1984). Peer relations in children with hyperactivity/attention deficit disorder. *Journal of Learning Disabilities*, 17, 560-567.
- Prinz, R. J., Connor, P. A., & Wilson, C. C. (1981). Hyperactive and aggressive behaviors in childhood: Intertwined dimensions. *Journal of Abnormal Child Psychology*, 9, 191-202.
- Roff, M., Sells, S. B., & Golden, M. M. (1972). *Social adjustment and personality development in children*. Minneapolis: University of Minnesota Press.
- Schachar, R., Sandberg, S., & Rutter, M. (1986). Agreement between teachers' ratings and observations of hyperactivity, inattentiveness, and defiance. *Journal of Abnormal Child Psychology*, 14, 331-345.