

Differential Performance of the MacArthur HBQ and DISC-IV in Identifying *DSM-IV* Internalizing Psychopathology in Young Children

JOAN L. LUBY, M.D., AMY HEFFELFINGER, Ph.D., JEFFREY R. MEASELLE, Ph.D.,
JENNIFER C. ABLOW, Ph.D., MARILYN J. ESSEX, Ph.D., LISA DIERKER, Ph.D.,
RICHARD HARRINGTON, M.D., HELENA C. KRAEMER, Ph.D., AND DAVID J. KUPFER, M.D.

ABSTRACT

Objective: Three sites collaborated to evaluate the differential performance of the MacArthur Health and Behavior Questionnaire (HBQ) and the Diagnostic Interview Schedule for Children Version IV (DISC-IV) in identifying *DSM-IV* psychopathology in young children. **Method:** A sample of 120 four- to eight-year-old nonreferred (community) ($n = 67$) and referred (clinical) ($n = 53$) children was examined. Mothers reported on their child's mental health symptoms using the HBQ (a dimensional measure with a clinical cutoff score) and the DISC-IV. Teachers independently reported on the child's symptoms and impairment in school using the teacher HBQ. Children self-reported on their symptoms using the companion Berkeley Puppet Interview. **Results:** On the basis of its derived clinical cutoff scores, the HBQ identified significantly more children with clinical symptoms of *DSM-IV* internalizing psychopathology than the DISC-IV in both referred and nonreferred groups. There was a high level of agreement between the two measures in the identification of externalizing psychopathology. Children identified as having internalizing psychopathology by the HBQ were also rated as highly symptomatic and impaired by teachers. **Conclusion:** The findings support the general validity of the parent HBQ for the assessment of young child psychopathology and the hypothesis that it captures more internalizing psychopathology than the DISC-IV in this population. *J. Am. Acad. Child Adolesc. Psychiatry*, 2002, 41(4):458–466. **Key Words:** internalizing psychopathology, young children, assessment.

The need for identification of childhood mental disorders at the earliest developmental point has become increasingly evident to the field of child mental health. The

Accepted November 13, 2001.

Dr. Luby is Assistant Professor of Psychiatry (Child), Washington University School of Medicine, St. Louis; Dr. Heffelfinger is with the Medical College of Wisconsin, Milwaukee; Drs. Measelle and Ablow are with the University of Oregon, Eugene; Dr. Kraemer is with the Stanford University Medical Center, Stanford, CA; Dr. Harrington is with the Royal Manchester Children's Hospital, Manchester, England; Dr. Dierker is with Wesleyan University, Middletown, CT; Dr. Essex is with the University of Wisconsin School of Medicine, Madison; and Dr. Kupfer is with the University of Pittsburgh School of Medicine.

Oversight for this project was provided by Dr. Essex, core member of the John D. and Catherine T. MacArthur Foundation Research Network on Psychopathology and Development. The MacArthur Research Network on Psychopathology and Development, of which Dr. Kupfer is Chair, supported this study.

Copies of the HBQ may be obtained from the Administrator of the MacArthur Research Network on Psychopathology and Development, Ms. Hermi Woodward, by sending a request to machermi@mindspring.com.

Additional materials for this article appear on the Journal's Web site: www.jaacap.com.

Correspondence to Dr. Luby, Washington University School of Medicine, Box 8134, 660 S. Euclid, St. Louis, MO 63110.

0890-8567/02/4104-0458©2002 by the American Academy of Child and Adolescent Psychiatry.

known chronicity, and in some cases treatment resistance, of early childhood mental disorders when identified later in childhood suggests that early identification could provide a window of opportunity for more effective treatment. At present, attempts to move the field forward in this area are thwarted by the lack of valid and reliable tools to assess psychopathology in young children. To date there have been a dearth of measures specifically designed for the identification of clinical psychopathology in this age group. Although a few validated dimensional measures of psychopathology have been widely used for parent and teacher informants, these measures have lacked sufficient developmental sensitivity and specificity due to the broad age range of children encompassed by a single version (e.g., 4–18 for the Child Behavior Checklist). Along these lines, the use of measures designed for older children with younger populations may fail to capture age-appropriate symptoms, contain questions in which the context may not be pertinent to the life experiences of the child (e.g., school), and fail to take into

account the known important effects of relationship context on symptom manifestations (Zeanah et al., 1997).

The MacArthur Battery for middle childhood was developed to address the current gaps in assessment tools for young child psychopathology. Although clinicians and developmentalists have traditionally viewed "latency," or school age (6–12), as "middle childhood," the MacArthur network focused on the age period between 4 and 8, on the basis of the rich socioemotional development occurring between the psychosocial transition to primary school after age 4 and the biological transition to puberty that begins at ages 7 to 8 (for detailed discussion, see Boyce et al., in press). The cognitive and emotional coherence during this period also allowed the application of a singular assessment methodology to children in this age group (e.g., a puppet interview). The MacArthur battery was designed to capture a multidimensional model of early developmental psychopathology, incorporating the biological, psychosocial, and contextual components of the developmental process (Boyce et al., in press). This multi-informant assessment strategy is compatible with the *DSM* system but extends beyond the categorical definitions of disorders in its scope. The current study was a part of an investigation of the utility, reliability, and preliminary validity of this multi-informant battery for young children. The MacArthur Health and Behavior Questionnaire (HBQ) parent and teacher report components of the battery are dimensional measures designed to assess mental health and impairment (and other aspects of functioning) and generate categorical diagnoses by the use of a clinical cutoff score. As such, the measure offers the benefits of a dimensional continuum while mapping onto the *DSM* categorical criteria. Internal consistency and test-retest reliability of the HBQ and the child-report component of this battery (Berkeley Puppet Interview [BPI]) have been established (Ablow et al., 1999; Essex et al., in press).

In the current study, we investigated the differential performance of the parent HBQ and the Diagnostic Interview Schedule for Children Version IV (DISC-IV) in the identification of *DSM-IV* psychopathology in a clinical and community population of young children. In the absence of a clear gold standard for defining clinical disorders in this age group, referral status, corroborating reports from teachers and the children themselves, and independent measures of impairment were used to investigate the content validity of the HBQ. We hypothesized that the HBQ would capture a higher proportion of children with clinically significant internalizing psy-

chopathology than the DISC-IV. Although the DISC-IV is a well-validated measure for use in older children, its application to children younger than 6 has not been established. For this reason, we hypothesized that it would not contain enough age-appropriate questions for some *DSM* internalizing disorders but would perform comparably to the HBQ in its ability to capture externalizing disorders. Age-adjusted symptom manifestations were thought to be potentially more important in internalizing disorders, in which symptoms were expected to be developmentally unique and more subtle in presentation.

METHOD

The MacArthur Three-City Outcome Study was a multisite collaboration conducted in two U.S. cities (Palo Alto, California and St. Louis, Missouri) and one British city (Manchester, England). The three sites were selected to achieve a socioeconomic, ethnically, and culturally diverse sample. Combined across the three sites, 120 children (Manchester $n = 36$, Palo Alto $n = 43$, St. Louis $n = 41$) aged 4.6 to 7.6, their primary caregivers (95.1% mothers), and their primary teacher or day-care providers participated. Sixty-seven community children were recruited from preschools and elementary schools through letters distributed to parents eliciting participation and from primary care patient registers or advertisements posted in waiting rooms. Fifty-three clinical children were recruited from local mental health clinics on the basis of review of admissions charts that met inclusion criteria. The case-control design was used to test how well the measures differentiated between young children who were referred for mental health services and those who were not. Parents and children were interviewed in the home simultaneously but separately (i.e., different rooms) with the measures described below. Interviewers and coders were blind to the group status of the subjects.

Subjects had to have a legal primary caretaker (defined here as parent, close relative functioning in loco parentis, legal guardian, or foster parent), with whom they had resided for at least 6 months continuously prior to study entry. Children were also required to have been in a school or day-care setting not less than a total of 40 hours. To factor out potential treatment effects, clinic-referred subjects were excluded if they had received more than 1 month of treatment. Subjects were also excluded if they had psychiatric hospitalizations, severe developmental delays, and/or if parents or children could not speak English fluently. In total, 158 families expressed interest in the study: 82.7% of the community families and 74.8% of the clinic-referred families agreed to participate; 38 families were not enrolled because of disinterest or meeting exclusion criteria.

Measures

Diagnostic Interview Schedule for Children Version IV. The National Institute of Mental Health (NIMH) DISC-IV (Shaffer et al., 1998) parent report version was administered to the primary caregiver of the child. This is a structured psychiatric interview with known reliability and validity in children aged 6 and older (for the parent report version). The computerized version of the interview was administered to the parent by trained research assistants using laptop computers in the subject's home. Research assistants at all sites underwent several days of formal training consisting of didactics and practice administrations until proficiency at DISC-IV administration was achieved.

Since this measure was designed and validated for use in school-aged children and adolescents, it was modified slightly for use in this younger sample. Modifications consisted only of the deletion of items pertinent to school and school-based behaviors for children younger than 6. In addition, due to the possibility that parents might become concerned when endorsing symptom states that could be normative in the young child (but clinically important in an older child), a disclaimer statement was used at the onset of the interview to assure parents that endorsed symptoms did not necessarily signify a mental health problem. The following modules were selected for use on the basis of their appropriateness to our young sample: Separation Anxiety Disorder (SAD), Specific Phobia, Generalized Anxiety Disorder (GAD), Selective Mutism, Elimination Disorders, Major Depression (MDD), Attention-Deficit/Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder (ODD), and Conduct Disorder (CD). The criterion validity, acceptability, interrater reliability, and test-retest reliability of earlier versions of the DISC for use in children 9 and older have been well established (Piacentini et al., 1993; Schwab-Stone et al., 1993; Shaffer et al., 1993, 1996). Despite its acceptability (but insufficient established age-specific validity) in children older than 6, the use of the interview in this young sample was considered experimental.

MacArthur Health and Behavior Questionnaire. The HBQ is a questionnaire with both parent and teacher versions, designed to assess the physical and mental health of young children (available upon request from the MacArthur Research Network). An HBQ item list is available via Article Plus at the Journal's Web site, www.jaacap.com. The measure yields dimensional ratings of functioning in the domains of (1) emotional and behavioral symptomatology, (2) physical health, (3) social adaptation, and (e) school adaptation. These four health domains comprised 18 preliminary subscales covering (1) externalizing symptoms, ADHD symptoms, internalizing symptoms, functional impairment, and mental health care utilization (mental health); (2) global physical health symptoms, chronic medical conditions, and physical health care utilization (physical health); (3) peer acceptance/rejection, bullied by peers, prosocial behavior, asocial with peers, social inhibition, and adult-led recreational activities (social adaptation); and (4) school engagement, academic competence, teacher-child closeness, and teacher-child conflict (school adaptation).

The questionnaire is a compilation of child health and behavior items derived from existing measures, including (1) the Ontario Child Health Study Scales (Boyle et al., 1987, 1993), (2) the Prosocial Behavior Scale (Weir and Duveen, 1981), (3) the Medical History Questionnaire from the Rand Health Insurance Study of children (Eisen et al., 1980; Lewis et al., 1989), and (4) additional items written or adapted to more fully assess the four domains. Symptoms of difficulties in social relations, for example, were assessed with items derived from the Child Behavior Scale (Ladd and Proffitt, 1996) and the Preschool Social Behavior Scale (Crick et al., 1997), and teacher-child relationship items were derived from the Student-Teacher Relationship Scale (Pianta, 1996; Pianta et al., 1995). All HBQ scales and items were selected and designed to parallel child-report dimensions in the BPI.

Parent and teacher responses on the HBQ are scored dichotomously (e.g., yes versus no) or on 3- or 4-point Likert scales consisting, for example, of 0 ("never or not true"), 1 ("sometimes true"), and 2 ("often or very true"). The HBQ parent version contains 140 items; the teacher version contains 115 items and omits those related to separation anxiety or the school context. Parents evaluate internalizing symptoms on three scales (Overanxious, Depression, and Separation Anxiety), whereas the teacher version contains only the Overanxious and Depression scales. Both parent and teacher versions include four externalizing scales (Oppositional Defiant, Conduct Problems, Overt Hostility, and Relational Aggression) and two attention scales (Inattention and

Impulsivity), and both versions take 15 to 20 minutes to complete. The internal consistency and test-retest reliability and discriminant validity of this measure have been established (Essex et al., in press).

Berkeley Puppet Interview Symptomatology Scales. The BPI-S is a puppet interview designed to elicit young children's self-reports of psychiatric symptoms. The internal consistency, test-retest reliability, and preliminary validity of this measure have been established (Ablow et al., 1999). The BPI-S has 65 items on nine separate scales, with three internalizing scales (Overanxious, Separation Anxiety, and Depression), four externalizing scales (Oppositional Defiant, Conduct Problems, Overt Hostility, and Relational Aggression), and two attention scales (Inattention and Impulsivity). The interview is videotaped and the child's responses are coded based on a 7 point Likert scale with higher scores representing greater psychopathology.

Statistical Analyses

To establish clinical cutoff points for each diagnosis on the maternal HBQ, receiver operating characteristic (ROC) curve analyses were completed to determine the most effective cutoff point to predict positive DISC-IV diagnosis, high scores on the teacher HBQ, and high scores on the BPI. These cutoff values allowed the maternal HBQ to be used as a categorical variables (0 = low symptoms for disorder, 1 = high symptoms for the disorder) at the point of maximum sensitivity and specificity. This procedure was used only for SAD, CD, ODD, ADHD, because these disorders had sufficient sample sizes of children receiving a DISC-IV diagnosis.

To determine cutoff values for clinical symptoms of MDD and GAD, respective prevalence rates of 4.5% and 6.0% were applied to the frequency distribution of mothers' reports on the HBQ of MDD and GAD symptoms among nonreferred children. These rates were chosen on the basis of the break in the distribution of HBQ scores between most scores and outlying scores in the nonreferred sample, and consideration for prevalence estimates of these disorders in the general population of older children (Bernstein et al., 1996; Birmaher et al., 1996). The HBQ value at the break in the distribution was selected to be the cutoff value for each diagnosis, as this appeared consistent with known prevalence rates for older children.

These cutoff values for each diagnostic category were then applied to both the referred and nonreferred groups to determine clinical group status. Prevalence rates of *DSM-IV* disorders according to the DISC-IV and clinically significant psychopathology according to the HBQ cutoffs within the corresponding *DSM-IV* category were compared in the study population by using the χ^2 statistic. Clinical group status according to the two measures was then compared with referral status. Clinical group status was also compared with teacher HBQ scores, child BPI scores, and measures of impairment in school as rated by the teacher. Post hoc comparisons were conducted to generate effect size for each group comparison in the area of school impairment. Group effect sizes, according to Cohen's (1988) distinction between small (0.20 through 0.49), moderate (0.40 through 0.79), and large (≥ 0.80) effect sizes were used in the interpretation of data. Group comparisons were conducted with *t* tests on normally distributed variables, and Mann-Whitney *U* tests were used for group comparisons of ordinal scales. To test the general validity of the HBQ as a measure of young child psychopathology, analyses of variance (ANOVA) were conducted with the HBQ as a dimensional measure to assess whether the HBQ accurately differentiated between those children who were referred and those with and without DISC-IV diagnoses. Referral status and presence or absence of a DISC-IV diagnoses were independent variables, and outcome values on the HBQ were dependent variables. Post hoc comparisons were conducted to compare each group difference and to generate the effect size for each group comparison.

RESULTS

The demographic characteristics of the three-city sample have been described previously in detail (Ablow et al., 1999). Key characteristics pertinent to this investigation are as follows. The sample consisted of $n = 120$ children (69 boys and 51 girls) across all three sites. There were no differences between clinical or community groups at any of the sites with respect to child gender or mean age (mean age = 5.9 years, $SD = 0.98$). The three sites did differ significantly in terms of household income ($F = 16.36$, $p < .001$). The Palo Alto sample reported a greater household income than the other two sites. With all three sites combined, the community sample reported greater annual income (mean = \$77.7K, $SD = \$16.3K$) than the clinical sample (mean = \$62K, $SD = \$20.3K$) ($F = 4.23$, $p < .05$). There were significant differences in ethnicity between sites ($[\chi^2 = 12, n = 120] = 131.23$, $p < .001$), with the Palo Alto and St. Louis sites having greater numbers of ethnic minorities. All sites were composed of predominantly (>80%) non-Hispanic white subjects (88.6% in the total sample).

The Maternal HBQ as a Categorical Measure of Disorders

Table 1 presents derived clinical cutoffs for the maternal HBQ based on ROC curves in relation to DISC-IV diagnoses, as described above (CD = 0.78, ODD = 0.83, ADHD = 1.05, SAD = 0.65). HBQ cutoff scores for clin-

ical symptoms of MDD and GAD (estimated on the basis of known prevalence rates in older children and rates of outlying scores in the study sample) were 0.85 and 0.90 respectively. Table 1 also presents the number of children with and without DISC-IV diagnoses in each HBQ category. No significant differences by study site in the prevalence of individual diagnoses according to the DISC-IV or clinically significant psychopathology according to parent HBQ cutoffs were found. The κ coefficients comparing the level of agreement between clinical status defined by the DISC-IV versus the HBQ cutoffs are also presented in Table 1.

Low rates of agreement for internalizing psychopathology were found between the DISC-IV and the maternal HBQ, with the DISC-IV identifying fewer children with clinical psychopathology. For MDD, there was no significant agreement (the DISC-IV identified only one child with this disorder), whereas there was significant but low-level agreement in the SAD and GAD categories. In the externalizing categories, the κ coefficients ranged from 0.57 to 0.77, indicating that the two measures have a high moderate level of agreement in their identification of CD, ODD, and ADHD psychopathology.

DISC-IV Versus HBQ Cutoffs: Diagnostic Frequency and Relationship to Referral Status

The ability of each measure to identify clinical psychopathology in individual diagnostic groups in relation

TABLE 1
Agreement Between DISC-IV Diagnosis and Maternal HBQ Cutoffs

DISC Dx	HBQ SAD Scale			HBQ MDD Scale			HBQ GAD Scale		
	Low Sym	High Sym	κ	Low Sym	High Sym	κ	Low Sym	High Sym	κ
	<0.65	>0.65		<0.85	>0.85		<0.90	>0.90	
No	95	19	.33***	109	12	-.02	107	15	.10**
Yes	2	7		1	0		0	1	
DISC Dx	HBQ CD Scale			HBQ ODD Scale			HBQ ADHD Scale		
	Low Sym	High Sym	κ	Low Sym	High Sym	κ	Low Sym	High Sym	κ
	<0.78	>0.78		<0.83	>0.83		<1.05	>1.05	
No	104	4	.77***	78	13	.57***	75	13	.67***
Yes	2	12		8	23		4	28	

Note: Significance of κ correlation between DISC diagnosis and HBQ classification: ** $p < .01$; *** $p < .001$. DISC = Diagnostic Interview Schedule for Children; HBQ = Health and Behavior Questionnaire; SAD = separation anxiety disorder; MDD = major depressive disorder; GAD = generalized anxiety disorder; CD = conduct disorder; ODD = oppositional defiant disorder; ADHD = attention-deficit/hyperactivity disorder; Sym = symptoms.

to referral status was investigated. The frequencies of clinical psychopathology identified by the DISC-IV and the maternal HBQ (using the clinical cutoff scores) in the referred and nonreferred samples were compared by means of χ^2 statistics (Table 2).

The maternal HBQ identified significantly more children in the referred group than the nonreferred group as having clinical psychopathology for every *DSM-IV* diagnostic category studied (MDD, GAD, SAD, CD, ODD, and ADHD). In contrast, the DISC-IV identified significantly more referred children as having SAD, CD, ODD, and ADHD, but not MDD and GAD. Further comparisons using DISC-IV diagnostic groups were not possible for MDD and GAD due to the low frequencies identified. The maternal HBQ identified higher numbers of children in the referred group in all internalizing categories when compared with the DISC-IV.

Discriminant Validity of Maternal HBQ Clinical Cutoffs

To investigate the validity of the HBQ clinical cutoff values, the number of DISC-IV symptoms (symptom counts), teacher report on the HBQ, and child report on the BPI were compared between the HBQ clinical and nonclinical groups. The number of DISC-IV symptoms rather than the categorical diagnosis was used, so that comparisons could be made using the DISC-IV in a dimensional fashion, allowing a test of the measure unrestricted by the *DSM* algorithm. The mean numbers of DISC-IV symptoms per group were compared by using the Mann-Whitney *U* statistic, because the variable was not normally distributed. Mean group performances for each of the other variables were compared with *t* tests.

In the area of internalizing psychopathology, significant relationships were found between the number of DISC symptoms and the HBQ clinical group status for SAD ($U = 925.0, p < .001$), MDD ($U = 500.5, p < .05$) and GAD ($U = 655.5, p < .01$). Significant relationships were also found between teacher HBQ as a dimensional measure and maternal HBQ clinical group status for MDD ($t_{117,1} = 2.47, p < .05$) and GAD ($t_{107,1} = 2.63, p < .05$). No significant relationships were found between these clinical cutoffs and child self-report on the BPI.

Significant relationships were found between the number of DISC-IV symptoms and HBQ cutoff for CD ($U = 162.5, p < .001$), ODD ($U = 576.0, p < .001$), and ADHD ($U = 317.0, p < .001$). Significant relationships were found between teacher HBQ and maternal HBQ cutoffs for CD ($t_{107,1} = 4.64, p < .001$), ODD ($t_{107,1} = 5.95, p < .001$), and ADHD ($t_{117,1} = 9.04, p < .001$). The relationship between the HBQ cutoffs and child report on the BPI was significant in the CD ($t_{116,1} = 2.24, p < .05$) and ODD ($t_{117,1} = 3.37, p < .001$) categories, but not the ADHD category.

Maternal HBQ Cutoffs: Relationship to Impairment

To better investigate the validity of maternal HBQ cutoffs for clinical status in both internalizing and externalizing diagnostic areas, the child's impairment in the school setting was compared in those children above and below the clinical cutoff (Table 3). Several forms of impairment were rated by the teacher in a separate section of the HBQ. School performance was derived by a composite of two items rated from 1 (far below grade level) to 5 (far above grade level). Peer and teacher-child rela-

TABLE 2
Frequencies of DISC-IV Diagnoses, Maternal HBQ Scores, and Referral Status

Diagnoses	DISC-IV		Sig.	Parent HBQ		Sig.
	Nonreferred (<i>n</i> = 67)	Referred (<i>n</i> = 53)		Nonreferred (<i>n</i> = 67)	Referred (<i>n</i> = 53)	
MDD	1 (1.5)	0 (0.0)		3 (4.5)	9 (17.0)	***
GAD	0 (0.0)	1 (1.9)		4 (6.0)	12 (22.6)	***
SAD	2 (3.0)	7 (13.2)	***	7 (10.5)	19 (35.8)	***
CD	2 (3.0)	12 (22.6)	***	1 (1.5)	15 (28.3)	***
ODD	5 (7.5)	25 (47.2)	***	6 (9.0)	30 (56.6)	***
ADHD	5 (7.5)	27 (50.9)	***	5 (7.5)	35 (66.0)	***

Note: Frequency of diagnosis and percent of group with diagnosis (in parentheses) are presented. *N* = 120. Sig. = level of significance. Difference in frequency of disorders between groups: *** $p < .001$. DISC = Diagnostic Interview Schedule for Children; HBQ = Health and Behavior Questionnaire; MDD = major depressive disorder; GAD = generalized anxiety disorder; SAD = separation anxiety disorder; CD = conduct disorder; ODD = oppositional defiant disorder; ADHD = attention-deficit/hyperactivity disorder.

TABLE 3

Note: HBO = Health and Behavior Questionnaire; SAD = separation anxiety disorder; MDD = major depressive disorder; GAD = generalized anxiety disorder; CD = conduct disorder.

^a Higher scores represent better performance.

^b Higher scores represent poorer physical health/more absenteeism.

In the area of internalizing psychopathology, very large effect sizes were found in the expected direction in all three internalizing diagnostic categories, as defined by HBQ cutoffs, in the two more objective areas of school impairment: physical health and absenteeism. There was also a moderate effect size for impairment in the expected direction in peer relationships in children with symptoms of MDD, defined by the HBQ cutoff, compared with those below the cutoff.

To investigate the validity of the HBQ as a dimensional measure of psychopathology in general, mean scores from the maternal HBQ were compared by using a two-way analysis of variance with group (referred or nonreferred) and DISC-IV diagnosis as the factors. Post hoc comparisons were made with *t* tests. Mothers reported significantly greater frequencies of symptoms on the HBQ in those children who were referred and those who had received DISC-IV diagnoses. There were main effects for group and DISC-IV diagnosis in each diagnostic category. The effect sizes for these differences in HBQ scores were large, ranging from 1.38 to 2.46 for all diagnostic categories in both referred and nonreferred groups.

DISCUSSION

Maternal HBQ Identification of Internalizing Psychopathology

Both the DISC-IV and HBQ cutoffs identified significantly more referred children than nonreferred children with clinically significant symptoms, supporting the content validity of both measures for the identification of psychopathology in young children in general. Although there was significant concordance between the two measures in the area of clinical externalizing symptoms, notable differences emerged between the two measures in the identification of clinical internalizing symptoms. Although both measures identified significantly more referred children as having clinical symptoms of SAD, only the HBQ identified significantly more referred children as having clinically significant symptoms of MDD and GAD.

Given the lack of a gold standard for the identification of clinical disorders in children, the question of whether those children identified by the HBQ but not the DISC-IV had clinically significant psychopathology was addressed by an investigation of independent ratings of symptoms and impairment. The hypothesis that the HBQ identified children who had clinically significant internalizing psychopathology was supported by the finding that these children were also rated as more symptomatic by their teachers. Given the well-known lack of concordance between informants in the identification of childhood psychopathology (particularly internalizing psychopathology), these findings are remarkable (Edelbrock et al., 1986; Gagnon et al., 1992; Verhulst and Akkerhuis, 1989). These children were also rated by teachers as impaired in several areas of functioning at school. Specifically, young children who scored above the HBQ threshold for clinical symptoms of MDD were rated by their teachers as impaired in peer relationships. Children above the HBQ threshold for clinical symptoms of SAD, MDD, or GAD missed more days of school and had poorer physical health at school according to their teachers, providing objective evidence of impairment in all three groups.

These findings support the hypothesis that children identified by scores above the clinical cutoffs on the maternal HBQ have clinically significant psychopathology. The fact that some of these markers of impairment were objective (e.g., physical health and days of missed school) lends further credibility to these ratings and the clinical validity of the cutoffs. The lack of concordance between mater-

nal HBQ scores and child report on the BPI in the area of internalizing psychopathology is not surprising, given the previous findings of lower internal consistency on these subscales and the fact that these BPI items are less behaviorally grounded (Ablow et al., 1999). These findings suggest that the HBQ may be a more sensitive measure than the DISC-IV for the identification of internalizing psychopathology in young children.

As would be expected in a sample of young children with psychopathology, high levels of comorbidity were also found. For example, on the basis of HBQ cutoffs, clinically significant symptoms of ADHD were found in 50% of children with clinical MDD symptoms and 75% of those with GAD symptoms. However, a qualitatively different pattern of impairment in children with internalizing versus externalizing psychopathology suggests that the impairment is attributable to the internalizing problems and is not likely to be due to the externalizing comorbidity.

Consistency Between HBQ and DISC-IV: Externalizing Psychopathology

A high level of concordance was found between the DISC-IV and the HBQ in the identification of children with clinical symptoms of CD, ODD, and ADHD within both referred and nonreferred groups. The identification of symptomatic children within these diagnostic categories in the referred group was also supported by teacher and child reports in most categories. These findings support the validity of both measures in young child populations with externalizing psychopathology.

Clinical Implications

The assessment of internalizing psychopathology in young children is an underdeveloped area urgently in need of further measurement development. Available data indicate that the majority of referrals for mental health services in young child populations are for externalizing problems, suggesting that internalizing problems are less likely to be identified and to come to clinical attention (Luby and Morgan, 1997). It is not clear, however, that the population prevalence of internalizing disorders in young children is as low as referral rates would suggest, raising the possibility that these disorders may frequently go unrecognized. Literature suggesting that internalizing disorders have a chronic and relapsing course and are associated with significant morbidity in childhood and potentially across the lifespan supports the need for the earliest

possible identification and treatment (Biederman et al., 1997; Caspi and Silva, 1995; Schwartz et al., 1999; Wright, 1985). Findings from this study suggest that the HBQ might be a more sensitive tool for the identification of internalizing psychopathology in young children than other standard measures such as the DISC-IV. This maternal report checklist is highly feasible for use as a screening tool in clinical mental health or primary care settings. The compatibility of this dimensional measure to the categorical *DSM* system also provides a tangible advantage for use in the established system of health care delivery.

Limitations

A limitation of the current study was the fact that the DISC-IV, in contrast to the HBQ, is a measure that was not specifically designed for use in young child populations. Although age-inappropriate items were deleted for the current study, the DISC-IV was originally designed for use in children 9 and older (and applicable to children as young as 6) and may have failed to capture some symptom states as they manifest in younger children. Since the time of data collection, an experimental version of the DISC, the Young Child DISC (Lucas et al., 1998) has been developed and would be a more optimal comparison measure for the youngest portion of this sample. Potentially related to this issue were the very low rates of MDD and GAD identified by the DISC-IV, which made it impossible to perform comparable ROC analyses to derive cutoffs for these disorders. In lieu of this, expected rates of disorder based on the distribution of outlying scores in the study sample was considered along with known prevalence rates of these disorders for older children. The assumption that rates of these disorders in young children would be similar to those in older children remains unproven; therefore, use of these prevalence rates could have erroneously influenced the derived cutoffs. However, the fact that children who scored above the derived cutoffs also had other markers of valid psychopathology remains an important and compelling finding suggesting the cutoffs are valid. The predominantly white ethnic composition (although both British and American) of the sample also limits the generalizability of the findings and suggests the measure should be tested in ethnically more diverse samples.

Conclusions

The HBQ identified significantly more young children with internalizing psychopathology than did the

DISC-IV. The validity of the HBQ in the identification of this psychopathology and the suggestion that it may be a more sensitive measure than the DISC-IV in this area is supported by higher ratings of symptoms and impairment on teacher reports, as well as objective evidence of impairment in school performance. Establishing the HBQ as a valid parent report measure of internalizing psychopathology in young children represents an important advance in young child assessment. The sensitivity of this measure can only be firmly established by future investigations that include clinician best-estimate diagnoses to ensure that the measure is accurately identifying, rather than overidentifying, clinical cases. Further investigation of the validity of the HBQ in the area of internalizing disorders, incorporating best-estimate diagnoses as well as robust independent external validators of psychiatric status and impairment is warranted. This should include an investigation of associated biological and genetic factors (e.g., family history of related disorders), observational measures of behavior, and data on the longitudinal outcome of young children identified early with internalizing psychopathology by the HBQ. Study findings do suggest that both the HBQ and DISC-IV are useful and valid tools for the identification of young children referred for mental health services in general and those with externalizing psychopathology in particular.

REFERENCES

- Ablow JC, Measelle JR, Kraemer HC et al. (1999), The MacArthur Three-City Outcome Study: evaluating multi-informant measures of young children's symptomatology. *J Am Acad Child Adolesc Psychiatry* 38:1580-1590
- Bernstein GA, Borchardt CM, Perwien AR (1996), Anxiety disorders in children and adolescents: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 35:1110-1119
- Biederman J, Faraone SV, Marris A et al. (1997), Panic disorder and agoraphobia in consecutively referred children and adolescents. *J Am Acad Child Adolesc Psychiatry* 36:214-223
- Birmaher B, Ryan ND, Williamson DE et al. (1996), Childhood and adolescent depression: a review of the past 10 years, part I. *J Am Acad Child Adolesc Psychiatry* 35:1427-1439
- Boyce T, Essex M, Woodward HR, Measelle J, Ablow J, Kupfer D (in press), The confluence of mental, physical, social, and academic difficulties in middle childhood, I: exploring the headwaters of early life morbidities. *J Am Acad Child Adolesc Psychiatry*
- Boyle MH, Offord DR, Hofmann HG et al. (1987), Ontario Child Health Study, I: methodology. *Arch Gen Psychiatry* 44:826-831
- Boyle MH, Offord DR, Racine YA, Fleming JE (1993), Evaluation of the revised Ontario Child Health Study scales. *J Child Psychol Psychiatry* 34:189-213
- Caspi A, Silva PA (1995), Temperamental qualities at age 3 predict personality traits in young adulthood. *Child Dev* 66:486-498
- Cohen J (1988), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum
- Crick NR, Casas JF, Mosher M (1997), Relational and overt aggression in preschool. *Dev Psychol* 33:579-588

- Edelbrock C, Costello AJ, Dulcan MK, Conover NC, Kala R (1986), Parent-child agreement on child psychiatric symptoms assessed via structured interview. *J Child Psychol Psychiatry* 27:181-190
- Eisen M, Donald CA, Ware JE, Brook RH (1980), *Conceptualization and Measurement of Health for Children in the Health Insurance Study*. Santa Monica, CA: US Department of Health, Education, and Welfare
- Essex M, Boyce T, Goldstein L, Armstrong J, Kraemer H, Kupfer D (in press), The confluence of mental, physical, social, and academic difficulties in middle childhood, II: developing the MacArthur Health and Behavior Questionnaire. *J Am Acad Child Adolesc Psychiatry*
- Gagnon C, Vitaro F, Tremblay RE (1992), Parent-teacher agreement on kindergartners' behavior problems: a research note. *J Child Psychol Psychiatry* 33:1255-1261
- Ladd GW, Profilet SM (1996), The Child Behavior Scale: a teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Dev Psychol* 32:1008-1024
- Lewis CC, Pantell RH, Kieckhefer GM (1989), Assessment of children's health status: field test of new approaches. *Med Care* 27:S54-S65
- Luby J, Morgan K (1997), Characteristics of an infant/preschool psychiatric clinic sample: implications for clinical assessment and nosology. *Infant Ment Health J* 18:209-220
- Lucas C, Fisher P, Luby J (1998), *Young-Child DISC-IV Research Draft: Diagnostic Interview Schedule for Children*. New York: Columbia University, Division of Child Psychiatry (disc@worldnet.att.net)
- Piacentini J, Shaffer D, Fisher P, Schwab-Stone M, Davies M, Giori P (1993), The Diagnostic Interview Schedule for Children-Revised Version (DISC-R), III: concurrent criterion validity. *J Am Acad Child Adolesc Psychiatry* 32:658-665
- Pianta RC (1996), *Manual and Scoring Guide for the Student-Teacher Relationship Scale*. Charlottesville: University of Virginia
- Pianta RC, Steinberg MS, Rollins KB (1995), The first two years of school: teacher-child relationships and deflections in children's classroom adjustment. *Dev Psychopathol* 7:295-312
- Schwab-Stone M, Fisher P, Piacentini J, Shaffer D, Davies M, Briggs M (1993), The Diagnostic Interview Schedule for Children-Revised Version (DISC-R), II: test-re-test reliability. *J Am Acad Child Adolesc Psychiatry* 32:651-657
- Schwartz CE, Snidman N, Kagan J (1999), Adolescent social anxiety as an outcome of inhibited temperament in childhood. *J Am Acad Child Adolesc Psychiatry* 38:1008-1015
- Shaffer D, Fisher P, Dulcan M et al. (1996), The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA study. *J Am Acad Child Adolesc Psychiatry* 35:865-877
- Shaffer D, Fisher P, Lucas C, and the NIMH DISC Editorial Board (1998), *Diagnostic Interview Schedule for Children, Version IV*. New York: Columbia University, Division of Psychiatry (disc@worldnet.att.net)
- Shaffer D, Schwab-Stone M, Fisher P et al. (1993), The Diagnostic Interview Schedule for Children-Revised Version (DISC-R), I: preparation, field testing, interrater reliability and acceptability. *J Am Acad Child Adolesc Psychiatry* 32:643-650
- Verhulst FC, Akkerhuis GW (1989), Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *J Child Psychol Psychiatry* 30:123-136
- Weir K, Duveen G (1981), Further development and validation of the Prosocial Behavior Questionnaire for use by teachers. *J Child Psychol Psychiatry* 22:357-374
- Wright HH (1985), Early identification and intervention with children who refuse to speak. *J Am Acad Child Psychiatry* 24:739-746
- Zeanah CH, Boris NW, Heller SS et al. (1997), Relationship assessment in infant mental health. *Infant Ment Health J* 18:182-197

The Adolescent Health Review: A Brief, Multidimensional Screening Instrument. Patricia A. Harrison, PhD, Timothy J. Beebe, PhD, Eunkyung Park, PhD

Purpose: To develop a brief, multidimensional screening instrument for adolescents that addresses psychosocial domains critical to adolescent preventive health care service. **Methods:** Secondary analyses were conducted on survey data obtained in 1995 from a school sample of 76,159 students in grades 9 and 12, as well as 893 adolescents from juvenile correctional facilities, 500 adolescents from chemical dependency treatment programs, and 575 adolescents from residential behavioral treatment programs. A comprehensive set of 300 survey items was used in a series of discriminant analyses to determine which items best distinguished males and females in each clinical sample from their counterparts in the school sample. **Results:** The item selection for the Adolescent Health Review was guided both by empirical analyses and clinical judgment. The final screen is comprised of 33 demographic and clinical items that address a variety of psychosocial domains. The computerized, self-administered screen can be completed in about 3 minutes. The screen is scored automatically and produces an easy-to-read risk-assessment profile. Because screening items were drawn from a large epidemiologic survey, normative profiles are available for each age and gender subgroup. **Conclusions:** A brief, empirically derived screening instrument, designed to address a range of adolescent risks, offers an opportunity for information gathering that otherwise might not be incorporated into routine clinic visits. *J Adolesc Health* 2001;29:131-139. Reprinted by permission of Elsevier Science, copyright 2001 by The Society for Adolescent Medicine.

A vertical yellow bar with a red diamond at the top, located on the left side of the page.

COPYRIGHT INFORMATION

TITLE: Differential Performance of the MacArthur HBQ and
DISC-IV in Identifying DSM-IV Internalizing
Psychopathology in Young Children

SOURCE: Journal of the American Academy of Child and
Adolescent Psychiatry 41 no4 Ap 2002

WN: 0209102114012

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited..

Copyright 1982-2002 The H.W. Wilson Company. All rights reserved.