

## SCREENING FOR DEPRESSIVE DISORDER IN CHILDREN AND ADOLESCENTS: VALIDATING THE CENTER FOR EPIDEMIOLOGIC STUDIES DEPRESSION SCALE FOR CHILDREN

MICHAEL FENDRICH,<sup>1,2</sup> MYRNA M. WEISSMAN,<sup>1,2</sup> AND VIRGINIA WARNER<sup>2</sup>

Fendrich, M. (Dept. of Psychiatry, College of Physicians and Surgeons of Columbia U., New York, NY 10032), M. M. Weissman, and V. Warner. Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for Children. *Am J Epidemiol* 1990; 131:538-51.

The utility of the Center for Epidemiologic Studies Depression Scale for Children (CES-DC), a modified version of the Center for Epidemiologic Studies Depression Scale, was explored in a sample of children, adolescents, and young adults at high or low risk for depression according to their parents' diagnosis. Proband parents were participants in the Yale Family Study of Major Depression who had children between the ages of 6 and 23 years. Diagnostic and self-report information on offspring was collected over two waves, spaced 2 years apart, from 1982 to 1986. Support was obtained for the reliability and validity of the CES-DC as a measure of depressive symptoms, especially for girls and for children and adolescents aged 12-18 years. Children with major depressive disorder or dysthymia, as defined by the *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition (DSM-III), had elevated scores in comparison with all other respondents. The CES-DC lacked diagnostic specificity; children with a range of current DSM-III diagnoses had elevated scores on the measure. A cutoff point of 15 and above for screening children and adolescents for current major depressive disorder or dysthymia may be optimal. Depressed respondents scoring below this cutoff point (false negatives) showed better social adjustment than true positives; nondepressed respondents scoring above this cutoff point (false positives) showed worse adjustment than true negatives. Factor analysis was used to construct an abbreviated, four-item version of the scale. The abbreviated scale was shown to be useful as a screen.

adolescence; child; depressive disorder; psychiatry; questionnaires

The value of self-report symptom scales for obtaining estimates of symptom prevalence and for screening psychiatric disorders in the community is well established

Received for publication December 28, 1988, and in final form August 24, 1989.

Abbreviations: CES-D, Center for Epidemiologic Studies Depression Scale; CES-DC, Center for Epidemiologic Studies Depression Scale for Children; DSM-III, *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition.

<sup>1</sup> Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York, NY.

<sup>2</sup> New York State Psychiatric Institute, New York, NY.

Reprint requests to Dr. Michael Fendrich, Depart-

ment of Psychiatry, College of Physicians and Surgeons of Columbia University, 722 West 168th Street, Box 14, New York, NY 10032.

This research was supported in part by National Institute of Mental Health Grants MH36197 and MH28274 and by Grant 86-213, "Child and Adult Depressive Disorders: A Test of Continuities Using Family-Genetic Data," from the John D. and Catherine T. MacArthur Foundation Mental Health Research Network on Risk and Protective Factors in the Major Mental Disorders.

for adult populations (1). With recent research suggesting a high prevalence of depressive illness among children and adolescents (2), there is an obvious need for valid screening instruments for use in younger populations. One way to address this need is by adapting instruments used in adult populations. The present study assesses the validity of the Center for Epidemiologic Studies Depression Scale for Children (CES-DC) (3), a measure adapted from a widely used instrument for measuring depression in adults, the Center for Epidemiologic Studies Depression Scale (CES-D) (4). The present study had three main aims: 1) to examine the utility of the CES-DC as a measure of depressive symptoms by evaluating its internal consistency reliability (using Chronbach's alpha), as well as its concurrent and convergent validity; 2) to evaluate the performance of the CES-DC as a screen for depressive disorder and make recommendations about appropriate cutoff points; and 3) to evaluate the performance of a screen based on an abbreviated version of the CES-DC.

#### *Description of the CES-D and CES-DC*

The CES-D is a 20-item questionnaire consisting of items selected from other depression scales covering six major symptom areas, including depressed mood, feelings of guilt/worthlessness, a sense of helplessness/hopelessness, psychomotor retardation, loss of appetite, and sleep disturbance. Each item is rated on a scale of 0 to 3 in terms of frequency of occurrence during the past week, with total scores ranging from 0 to 60; higher scores indicate more symptomatology. In adults, the CES-D has been validated as an indicator of depressive symptoms and as a screen for clinical depression (5, 6). Items for the CES-DC were modified from the CES-D by Orvaschel (see Weissman et al. (3)) to facilitate comprehension and enhance the instrument's relevance to a younger population. The modifications are shown in the Appendix.

#### *Previous research on use of the CES-D for children*

Support for the utility of the CES-D for children is mixed. Schoenbach et al. (7, 8) demonstrated the feasibility of using a modified CES-D (response alternatives were simplified) to assess depressive symptoms in a sample of junior high school students. Two studies, including a small pilot study (3) of 28 children of psychiatrically ill parents and a large community study of over 1,000 rural public school students (9), found that scores on a modified CES-D were correlated with scores on another self-report symptom checklist, the Children's Depression Inventory (10). Faulstich et al. (11) found that CES-DC scores of 39 child psychiatric patients with *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition (DSM-III) (12), diagnoses of depression were not significantly different from scores of 39 randomly selected child psychiatric patients without depression. High scale reliability coefficients reported by both Schoenbach et al. and Faulstich et al. (Chronbach's alpha coefficients greater than 0.80) were comparable to those reported by Radloff (4) in community samples of white adults. Of the four studies discussed, only one (11) attempted to validate the measure for use in children and adolescents by investigating its association with DSM-III clinical diagnosis. The use of a patient sample, however, limits the generalizability of the study's findings.

#### MATERIALS AND METHODS

##### *Sample*

This study was based on a longitudinal sample of 220 children, adolescents, and young adults at high or low risk for major depression by virtue of the presence or absence of major depression in one or more of their parents. A detailed description of the sample of parents and children has been provided elsewhere (1, 13). Briefly, the proband parents were participants in the Yale Family Study of Major Depression (14) who

had children between the ages of 6 and 23 years. The depressed probands had primarily received treatment at the Yale University Depression Research Unit (New Haven, CT). Major depression in parents was defined according to Research Diagnostic Criteria (15) (modified to require 4 weeks' duration of symptoms and impairment in a major social role) and was assessed using the Schedule for Affective Disorders and Schizophrenia, Lifetime Version (16). The parents composing the nondepressed group were from a 1975 community study in New Haven, Connecticut (17), and had reported no history of psychiatric illness in at least five direct interviews conducted over an 8-year period. All of the probands were white and group-matched by age and sex. Diagnostic and self-report information on offspring was collected over two waves, spaced 2 years apart, from 1982 to 1986. The present study focused on responses to the CES-D at the initial assessment (wave 1) and included 166 respondents who completed all 20 items on this measure and who were screened by means of an independent psychiatric assessment. Compared with the other 54 children in the study sample, the 166 completers were significantly less likely to be children of one or more depressed parents ( $\chi^2 = 8.27, 1 \text{ df}; p < 0.01$ ).

Table 1 shows the age and sex distribu-

tion of the sample by parents' diagnostic status. The majority of the children (64 percent) had one or more parents with a lifetime history of major depression. Boys represented 44 percent of the sample and girls represented 56 percent. The respondents ranged in age from 6 to 23 years. Sixteen percent of the children were 6-11 years of age, 45 percent were 12-18, and 40 percent were 19-23. Of the 26 children in the youngest age group (6-11 years), five were under 8 years of age. Age group was not associated with parents' diagnostic status.

#### Assessment of children

The psychiatric assessment of the children was made using the Schedule for Affective Disorders and Schizophrenia for School-Aged Children, Epidemiologic Version (18), a widely used research instrument for obtaining current and lifetime psychiatric symptoms in children. Interviewers who were blind to the parents' diagnoses interviewed a parent (usually the mother) about the child and, at a later time, the child. Two sets of diagnostic assessments were made on each child, including one at wave 1 and one 2 years later at wave 2.

In addition to the CES-DC, two other

TABLE 1

Age and sex distribution of 166 children, adolescents, and young adults assessed by the Center for Epidemiologic Studies Depression Scale for Children, by parents' diagnostic status, New Haven, Connecticut, 1982-1986

Sex and age (years)	≥1 parent depressed		Neither parent depressed		Total
	No.	%	No.	%	
Boys					
6-11	10	9.3	2	3.4	12
12-18	18	16.8	13	22.0	31
19-23	20	18.7	10	16.9	30
Girls					
6-11	9	8.4	5	8.5	14
12-18	27	25.2	16	27.1	43
19-23	23	21.5	13	22.0	36
Total no. of children	107		59		166
Total no. of families	51		25		76

measures based on self-administered reports completed several weeks prior to the diagnostic interview were considered in the analysis, including a 50-item scale to assess a child's self-esteem (the Coopersmith Self-Esteem Inventory (19)) and a 30-item semantic differential scale to assess a child's feelings of undesirability (the Child Trait Checklist (20)). The child's intelligence quotient was assessed using the Peabody Picture Vocabulary Test (21). The analysis also included an index reflecting a psychiatrist's evaluation of the overall social functioning of the child based on all available information (the Children's Global Assessment Scale (22)).

### *Diagnoses*

The diagnoses constituting the validity criteria for this study were made according to a "best estimate" procedure (23) in which a child psychiatrist and psychologist who were not involved in the interviewing reviewed all sources of information (including the Schedule for Affective Disorders and Schizophrenia for School-Aged Children, Epidemiologic Version, interviews with both mother and child) and independently assigned a current and/or lifetime DSM-III diagnosis. Discrepancies in diagnoses by independent evaluators were resolved by a third source, who also independently and blindly reviewed all available information. We were interested in determining the validity of the CES-DC as a measure of depression in children and as a screen for clinically significant DSM-III diagnoses of depression. Because comorbidity between disorders was common, we developed a diagnostic hierarchy, placing individuals in diagnostic categories in the following order: major depressive disorder according to "strict criteria" (impairment in a social role and 4 weeks' duration of illness), major depressive disorder according to DSM-III, dysthymia, minor depression, any anxiety disorder, conduct disorder or attention deficit disorder, any other diagnosis, and no diagnosis. Subjects with diagnoses indicative of current, clinically significant depres-

sion (according to DSM-III) were considered criterion cases in the assessment of validity; this included anyone with major depressive disorder or dysthymia.

Most of the analyses for the validation study were based on diagnoses evaluated at the initial interview (wave 1). Information about current diagnostic status was also obtained from follow-up interviews with children and/or their parents 2 years after the initial interview (wave 2). This information, available for 141 of the 166 CES-DC completers (84.9 percent), was used to assess the predictive validity of the CES-DC. All other variables included in the analysis were assessed at wave 1.

## RESULTS

### *Reliability*

Reliability is a necessary prerequisite for validity (24). Chronbach's alpha coefficient, an overall index of the parallelism of items in a scale (i.e., of internal consistency), provides a convenient measure of reliability. An overall alpha coefficient of 0.89 (lower bound of 95 percent confidence interval, 0.86) (25) indicates excellent internal consistency reliability. Alpha coefficients were similar for boys and girls and for children with and without a depressed parent. An alpha coefficient of 0.78 (lower bound of 95 percent confidence interval, 0.64) for children aged 6-11 years suggested a trend toward less internal consistency reliability in younger children as compared with older children and adolescents.

### *Validity*

Figure 1 and table 2 provide information about the concurrent validity of the measure by illustrating the distribution of CES-DC scores (figure 1) and mean CES-DC values (table 2) within each diagnostic subgroup. To the extent that elevated scores on the CES-DC are associated with current DSM-III diagnoses of major depressive disorder or dysthymia, the concurrent validity of the measure is supported. To the extent that elevated scores on the

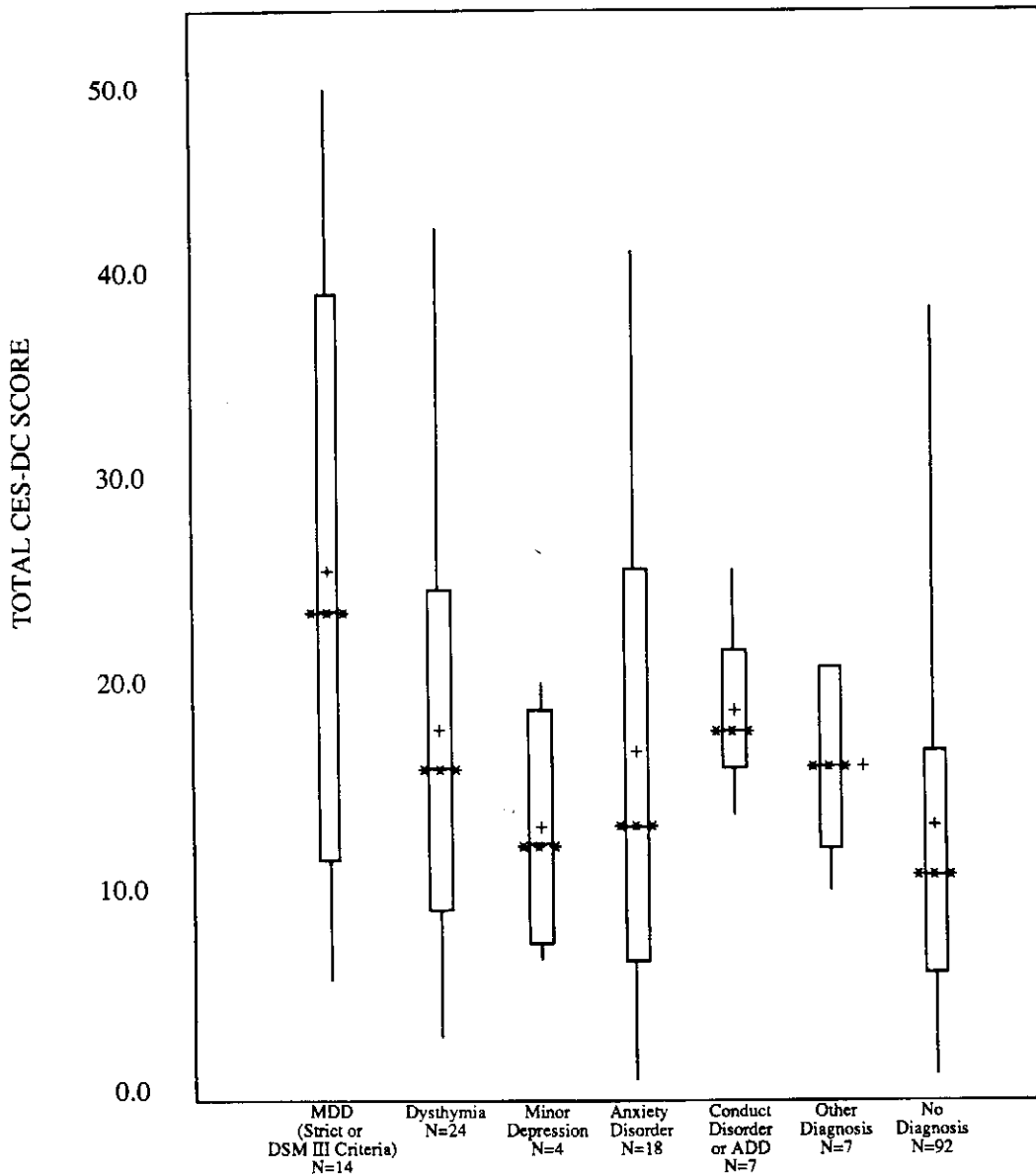


FIGURE 1. Distribution of scores on the Center for Epidemiologic Studies Depression Scale for Children (CES-DC), by current diagnosis, in 166 children, adolescents, and young adults aged 6–23 years, New Haven, Connecticut, 1982–1986. +, mean; \*\*\*, median; vertical lines, range. Boxes extend from the 25th to the 75th percentile of CES-DC scores for each diagnosis. (MDD, major depressive disorder; DSM-III, *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition; ADD, attention deficit disorder).

CES-DC are specific to criterion diagnoses (i.e., they are observed for current major depressive disorder or dysthymia and not for other current diagnoses), the validity of the measure as a screen for depression is supported.

#### Concurrent validity

Using a series of “box and whisker” plots (26), figure 1 shows the distribution CES-DC scores of subjects according to their hierarchically defined current DSM-III di-

TABLE 2

Mean CES-DC\* scores of children, adolescents, and young adults aged 6-23 years, by DSM-III\* diagnosis in the child, New Haven, Connecticut, 1982-1986

Child's current hierarchical† DSM-III diagnosis	Mean CES-DC score (SD*)	No. of children
Major depression (strict criteria)‡	27.6 (16.99)	10
Major depression	21.8 (13.57)	4
Dysthymia	18.5 (11.91)	24
Minor depression	13.0 (6.06)	4
Anxiety disorder	16.6 (11.83)	18
Attention deficit disorder/conduct disorder	19.0 (4.04)	7
Other diagnosis	16.1 (4.22)	7
No diagnosis	12.8 (8.70)	92

\* CES-DC, Center for Epidemiologic Studies Depression Scale for Children; DSM-III, *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition; SD, standard deviation.

† The diagnostic hierarchy placed children with current DSM-III diagnoses into one diagnostic category using the following hierarchical order: major depression (strict criteria), major depression (DSM-III criteria), dysthymia, minor depression, anxiety disorder, attention deficit disorder or conduct disorder, other diagnosis, and no diagnosis.

‡ Children with depression who reported a duration of depressive illness lasting longer than 3 weeks and an impairment in a major social role (e.g., inability to attend school) were classified as fulfilling strict criteria for major depression.

agnosis. The boundaries of each box stretch from the 25th percentile to the 75th percentile of the distribution for each diagnosis. The lines above and below each box indicate the range. The figure shows that the distribution of scores for those with major depressive disorder (by strict or DSM-III criteria) was elevated in comparison to the distribution of scores for those with other diagnoses. The distribution of scores for those with dysthymia was below the distribution for those with major depressive disorder and paralleled the distribution of scores for children diagnosed with anxiety disorders. Children with dysthymia had an elevated median in comparison with those with minor depression, anxiety disorder, and no diagnosis; they also showed more extreme high scores than every non-depressed group. The scores of subjects with no diagnosis were concentrated well below scores for subjects with major depressive disorder or dysthymia.

Table 2 lists mean CES-DC scores by the hierarchically defined current DSM-III diagnosis. One-way analysis of variance indicated a significant main effect for current diagnostic status on total score ( $F = 3.69$ ,

7 and 158 df;  $p < 0.01$ ). The relatively small number of subjects in each of the diagnostic groups limited the power to test differences between means; significance levels of post hoc tests should be viewed with caution. Post hoc comparisons using the Scheffé technique (27) indicated that children with major depressive disorder or dysthymia had significantly higher mean scores than all other children in the sample (21.2 vs. 13.9;  $p < 0.05$ ) and children with no current DSM-III diagnosis (21.2 vs. 12.8;  $p < 0.01$ ). These findings, along with observations from figure 1, lend support to the concurrent validity of the CES-DC.

#### Validity as a screen

The mean CES-DC score for children with current major depressive disorder or dysthymia was higher than that for children with other current DSM-III diagnoses (21.2 vs. 16.6). A post hoc comparison (using the Scheffé criteria) that was not significantly different from zero suggested that the CES-DC may have a limited capacity to discriminate between children with major depressive disorder or dysthy-

mia and children with other DSM-III disorders. Figure 1 supports the contention that CES-DC score distributions for children in all diagnostic categories are elevated in comparison with the scores for children without a current diagnosis.

Two other analyses assessed the scale's ability to discriminate between children diagnosed with current depressive disorder and children diagnosed with lifetime (but not current) depressive disorder. Current depressives showed elevated mean CES-DC scores compared with lifetime depressives (21.2 vs. 17.3;  $t = 1.27$ , 65 df; not significant). When the analysis was restricted to just children with major depressive disorder, however, children with current major depressive disorder had significantly higher scores on the CES-DC than children with lifetime major depressive disorder (25.9 vs. 17.6;  $t = 2.28$ , 54 df;  $p < 0.05$ ).

#### Convergent validity

Since self-devaluation and low self-esteem are characteristic of depression (28), correlations with two independent measures tapping this construct—the Child Trait Checklist (20) and the Coopersmith Self-Esteem Inventory (19)—were explored. Additionally, since elevated symp-

tomatology should be reflected in impaired social functioning, associations with the Children's Global Assessment Scale (22) were explored. The CES-DC correlated 0.27 with the Child Trait Checklist, 0.46 with the Coopersmith Self-Esteem Inventory, and 0.43 with the Children's Global Assessment Scale. All three coefficients were significantly different from zero ( $p < 0.01$ ); the magnitude of the latter two correlations indicates that the Children's Global Assessment Scale and the Coopersmith Self-Esteem Inventory each explain about 20 percent of the total variation in CES-DC scores.

#### Concurrent validity within demographic subgroups

Table 3 compares the concurrent validity of the scale across different subgroups of respondents. The effect size estimate, " $d$ ," provides a standard unit for assessing the impact that membership in the criterion group (i.e., those diagnosed with major depressive disorder or dysthymia) has on CES-DC scores (29). An effect size is measured by the ratio of the difference between the two group means (in this case, the difference between the group mean for those with major depressive disorder or

TABLE 3

Mean CES-DC† scores and effect sizes for children, adolescents, and young adults aged 6–23 years, by age, sex of the child, and the presence or absence of parental depression, New Haven, Connecticut, 1982–1986

	MDD† or dysthymia		No depression		Effect size ( $d$ )
	Mean CES-DC score (SD)†	No. of children	Mean CES-DC score (SD)	No. of children	
Total sample	21.2 (13.76)	38	13.9 (8.89)	128	0.72**
Age (years)					
6–11	8.7 (5.51)	3	12.7 (8.34)	23	–0.49
12–18	26.2 (12.76)	18	13.1 (8.30)	56	1.37**
19–23	18.1 (13.90)	17	15.3 (9.74)	49	0.26
Sex					
Boys	17.6 (13.47)	13	13.2 (8.95)	60	0.45
Girls	23.1 (13.80)	25	14.5 (8.86)	68	0.83**
Parental depression					
≥1 parent depressed	20.8 (14.72)	25	15.5 (9.35)	82	0.49*
Neither parent depressed	22.0 (12.23)	13	11.0 (7.24)	46	1.29**

\*  $p < 0.05$ ; \*\*  $p < 0.01$ .

† CES-DC, Center for Epidemiologic Studies Depression Scale for Children; MDD, major depressive disorder; SD, standard deviation.

dysthymia and the group mean for those without major depressive disorder or dysthymia) over the pooled sample standard deviation. The larger the effect size, the better the scale is able to discriminate between the two groups. Effect sizes not significantly different from zero indicate a lack of scale validity.

For the entire sample, the validity of the measure was supported, since the effect size estimate of 0.72 was significantly different from zero. There were considerable differences in effect sizes when observations were stratified by age and by sex. The effect sizes for the children in the youngest and oldest age groups were not significantly different from zero. (The negative effect size for children in the youngest age group may be an unreliable estimate, since there were only three children aged 6–11 years who had any current major depressive disorder or dysthymia.) The only age group showing an effect size significantly different from zero was the group aged 12–18 years; this group showed the largest effect size of any subgroup: 1.37. The effect size for boys was not significantly different from zero and was half the value of the effect size for girls (which was significantly different from zero). Children from both parent risk groups showed effect sizes that were significantly different from zero. However, the effect size for children with neither parent depressed (1.29) was more than twice the effect size for children with one or more depressed parents (0.49). The relatively small effect sizes within certain subgroups, particularly for the oldest and youngest age groups in the sample, boys, and children with one or more depressed parents, raise questions about the validity of the CES-DC as an index of current depression within these groups.

Indicators of a child's current depression, risk group status (i.e., whether he or she was a child of a depressed parent), age, and sex were considered as covariates in an ordinary least squares regression model predicting total CES-DC score. Significant positive coefficients were obtained for

child's depression and risk group status. There were no significant interactions between child's depression and the three other variables. Thus, despite differences in effect sizes for different subgroups, a diagnosis of major depressive disorder or dysthymia is consistently associated in these data with an elevated CES-DC score.

#### *Screening performance*

Receiver Operating Characteristic curves (plots of sensitivity versus false-positive rates for every possible cutoff point) (30) were inspected for three criterion groups, including major depressive disorder, major depressive disorder or dysthymia, and any diagnosis. The plots suggested that the CES-DC provides good case discrimination for all three criteria; the best discrimination occurs when major depressive disorder is the criterion. The plots suggested potential optimal cutoff points ranging from 14 to 16 and above.

Table 4 shows evaluations of the performance of the scale against the three criterion groups at each of these potential cutoff points. Within any particular criterion group, all three cutoff points yielded similar levels of sensitivity and specificity. Moving up from a cutoff point of 14 and above to a cutoff point of 16 and above resulted in a decrease of sensitivity of 8–10 percent and a gain in specificity of 8–9 percent, depending on the criterion diagnosis. At any cutoff point, the scale appeared to be most sensitive to major depressive disorder. Positive predictive value was consistently lowest for major depressive disorder and highest for "any diagnosis." Negative predictive value was consistently highest for major depressive disorder and lowest for "any diagnosis." The 15-and-above cutoff point appeared to yield the best compromise between sensitivity and specificity, especially when major depressive disorder or dysthymia and "any diagnosis" were the criteria. The 16-and-above cutoff point yielded the best compromise between these indices when major depressive disorder was the criterion.



TABLE 4  
 Validity statistics for the Center for Epidemiologic Studies Depression Scale for Children in children, adolescents, and young adults aged 6-23 years, New Haven, Connecticut, 1982-1986

Cutoff score and criterion diagnosis*	Sensitivity	Specificity	Positive predictive value	Negative predictive value
	% (nos.)	% (nos.)	% (nos.)	% (nos.)
14 and above				
MDD†	79 (11/14)	53 (80/152)	13 (11/83)	96 (80/83)
MDD or dysthymia	68 (26/38)	55 (71/128)	31 (26/83)	86 (71/83)
Any diagnosis	66 (49/74)	63 (58/92)	59 (49/83)	70 (58/83)
15 and above				
MDD	71 (10/14)	57 (86/152)	13 (10/76)	96 (86/90)
MDD or dysthymia	63 (24/38)	59 (76/128)	32 (24/76)	84 (76/90)
Any diagnosis	61 (45/74)	66 (61/92)	59 (45/76)	68 (61/90)
16 and above				
MDD	71 (10/14)	62 (94/152)	15 (10/68)	96 (94/98)
MDD or dysthymia	58 (22/38)	64 (82/128)	32 (22/68)	84 (82/98)
Any diagnosis	55 (41/74)	71 (65/92)	60 (41/68)	66 (65/98)

\* Current best-estimate diagnosis in child.

† MDD, major depressive disorder.

Schoenbach et al. (7) generated diagnoses by adding together scale items according to an algorithm based on Research Diagnostic Criteria (15). In the present sample, the Research Diagnostic Criteria consistently yielded unacceptably low levels of sensitivity (ranging from 7 percent to 14 percent); specificity was very high for all three criterion diagnoses (at least 97 percent).

Unfortunately, all three cutoff points yielded relatively high rates of misclassified cases for all criterion diagnoses. For example, at a cutoff point of 15 and above with major depressive disorder or dysthymia as the criterion, the false-negative rate was 37 percent and the false-positive rate was 41 percent. Although this suggests a limitation of the CES-DC, the possibility exists that there are important differences between individuals who are misclassified.

The 14 subjects with current major depressive disorder or dysthymia scoring below the 15-and-above cutoff point (false negatives) were compared with the 24 subjects with major depressive disorder or dysthymia scoring above the cutoff point (true positives) on age, sex, parental diagnosis, intelligence quotient, self-esteem, and overall social functioning. The false-negative subjects reported significantly higher self-

esteem than the true-positive subjects (the mean Coopersmith Self-Esteem Inventory score for false negatives was 68.9; for true positives, it was 56.7;  $t = 2.60$ , 36 df;  $p < 0.05$ ). The false negatives also showed significantly better social functioning than the true positives (the mean Children's Global Assessment Scale score for false negatives was 59.7; for true positives, it was 51.6;  $t = 2.6$ , 36 df;  $p < 0.05$ ).

Again using the 15-and-above cutoff point, false positives were compared with true negatives. For this analysis, children were stratified into two groups: those with other current diagnoses ( $n = 36$ ) and those without current diagnoses ( $n = 92$ ). Within the group of 36 children with other current diagnoses, the 21 false positives demonstrated significantly lower self-esteem on the Coopersmith Inventory (65.1 vs. 74.5;  $t = 2.23$ , 34 df;  $p < 0.05$ ) and more impaired social functioning on the Children's Global Assessment scores (57 vs. 64.9;  $t = 2.31$ , 34 df;  $p < 0.05$ ) than the 15 true negatives. Similarly, within the group of 92 children without a current diagnosis, the 31 false positives demonstrated lower self-esteem on the Coopersmith Inventory (67.6 vs. 80.3;  $t = 5.19$ , 90 df;  $p < 0.0001$ ) and more impaired social functioning on the Children's Global Assessment Scale (70.4 vs.

78.8;  $t = 3.36$ , 90 df;  $p < 0.01$ ) than the 61 true negatives. The 31 false positives were also more likely than the 61 true negatives to be children of one or more depressed parents ( $\chi^2 = 12.03$ , 1 df;  $p < 0.01$ ) and to have a lifetime (but not current) history of a psychiatric diagnosis ( $\chi^2 = 4.54$ , 1 df;  $p < 0.05$ ).

#### *Predictive validity of the screen*

Table 5 shows evaluations of the utility of the CES-DC as a screen for predicting major depressive disorder or dysthymia at the follow-up interview (wave 2). Children scoring 15 and above on the CES-DC were counted as positive screens. Children with wave 2 diagnoses of major depressive disorder or dysthymia were counted as cases. CES-DC classification status at wave 1 was compared with wave 2 diagnosis for the subgroup of 141 participants who received diagnostic assessments at follow-up. Sensitivity assessed the percentage of wave 2 cases among the positive screens at wave 1. Specificity assessed the percentage of negative screens at wave 1 among those who were not cases at wave 2. These statistics were calculated for the entire subsample and for subgroups broken down according to wave 1 diagnostic status.

Consistently high levels of sensitivity suggested that most children who were cases at wave 2 were screened positive at wave 1. The screen was particularly sensitive for children with wave 1 diagnoses of

depression. Among children with current major depressive disorder or dysthymia at wave 1, the screen had perfect sensitivity. Specificity of the screen was consistently worse than sensitivity. The screen showed particularly poor specificity for children with wave 1 current major depressive disorder or dysthymia (42 percent) and wave 1 lifetime major depressive disorder or dysthymia (43 percent).

#### *Screening from a shortened scale*

Researchers or clinicians interested in screening children for depression may find it less time-consuming to use shortened versions of scales such as the CES-DC. Recently, it has been suggested that the use of factor analysis may facilitate the selection of a reliable and diverse subset of items for a valid screening instrument (31). An attempt was made to assess the screening performance of a shortened scale consisting of the items loading highest on the factors derived from a factor analysis on the total scale.

The VARIMAX-rotated factor solution in the present study yielded four factors paralleling those found in other studies using the CES-D (4), including depressed affect, positive affect, somatic problems, and interpersonal problems. The items loading highest on each of these factors (items 18, 12, 20, and 15) were subsequently added together to create a revised, shortened scale. As a consequence of scale ab-

TABLE 5

*Validity of a CES-DC\* cutoff score of 15 and above for predicting wave 2 (follow-up interview) current MDD\* or dysthymia in children, adolescents, and young adults, New Haven, Connecticut, 1982-1986†*

Wave 1 diagnostic status	No. of children	Sensitivity	Specificity	Positive predictive value	Negative predictive value
		% (nos.)	% (nos.)	% (nos.)	% (nos.)
All children combined	141	80 (8/10)	57 (75/131)	13 (8/64)	97 (75/77)
Current MDD or dysthymia	30	100 (4/4)	42 (11/26)	21 (4/19)	100 (11/11)
Lifetime MDD or dysthymia	56‡	86 (6/7)	43 (21/49)	18 (6/34)	95 (21/22)
No lifetime/current MDD or dysthymia	85	67 (2/3)	66 (54/82)	7 (2/30)	98 (54/55)

\* CES-DC, Center for Epidemiologic Studies Depression Scale for Children; MDD, major depressive disorder.

† All children scoring 15 or higher were screened positive; those scoring below 15 were screened negative. Criterion cases had wave 2 current best-estimate MDD or dysthymia. Analyses were based on 141 children with direct interviews at wave 2.

‡ This includes 30 cases with current MDD or dysthymia at wave 1.

breviation (24), reliability of the shortened version as measured by coefficient alpha dropped to 0.64.

A cutoff point of 3 and above for the shortened scale yielded values for sensitivity and specificity that were most comparable to values obtained for the entire scale at the 15-and-above cutoff point. The 3-and-above cutoff point corresponded to sensitivity of 63 percent and specificity of 55 percent when the criterion diagnostic group was major depressive disorder or dysthymia, sensitivity of 64 percent and specificity of 52 percent when the criterion was DSM-III major depressive disorder, and sensitivity of 62 percent and specificity of 61 percent when the criterion was any diagnosis.

## DISCUSSION

### *Reliability and validity*

The results provided support for the reliability and validity of the CES-DC as a measure of depressive symptomatology in children, adolescents, and young adults. The CES-DC appears to be most valid as a measure of depression for girls and for children aged 12-18 years. Scale reliability and validity may be particularly limited for younger children aged 6-11 years. The analyses revealed the CES-DC to be a general measure of current childhood psychopathology and not just a measure of current depressive disorder.

### *Screening*

The results suggested that the CES-DC performs adequately as a screen, especially when the criterion is restricted to major depressive disorder. A cutoff point of 15 and above for classifying children as cases might maximize both sensitivity and specificity when major depressive disorder or dysthymia is the criterion. The conventional adult cutoff point of 16 and above might be optimal when the criterion is limited to major depressive disorder. The 15-and-above cutoff point was just as valid for classifying respondents with any diagnosis

as it was for classifying children with major depressive disorder or dysthymia. Consistently high sensitivity to depressive diagnoses at follow-up suggests that this cutoff point has prognostic value.

Since a screen for depression needs to be sensitive in order to detect clinical cases, and since the Research Diagnostic Criteria algorithm yielded very low levels of sensitivity, continued use of the conventional scoring method (i.e., symptom summation) is supported. At the 15-and-above cutoff point, the CES-DC yielded sensitivity that was comparable to sensitivity found in adult community samples using a 16-and-above cutoff point (2, 5); at these same cutoff points, the specificity of the CES-DC was considerably lower than that of the CES-D. It was demonstrated that a reduced four-item CES-DC scale proved almost as accurate a screen (at a cutoff point of 3 and above) as the entire set of 20 items. Analyses of misclassified cases yielded encouraging findings about the validity of the CES-DC as a screen. False-negative subjects were functioning significantly better on a number of other measures than the true-positive subjects. The false positives were functioning significantly worse on a similar set of measures and were more likely to have a lifetime diagnosis of major depressive disorder or dysthymia than true negatives. The CES-DC may detect subtle distinctions between children and adolescents that current DSM-III diagnostic classification does not.

### *Limitations*

The sample used in this analysis was derived from a retrospective cohort study and therefore was not representative of the community. Nevertheless, the design increased the probability that criterion cases of depression would be found. Conclusions about the validity of the measure are limited to the extent that certain groups were unrepresented (only white families were included) or underrepresented (e.g., children under age 12). Validity statistics presented in this paper were based on only one sample

of relatively small size ( $n = 166$ ). Suggestions about the performance of the CES-DC or a subset of items from the CES-DC as a screen for depression in children and adolescents need to be verified with larger independent community samples before firm conclusions about the adequacy of the measure as a screen can be drawn.

#### Recommendations

The CES-DC is most appropriate for use as a diagnostic screen for children and adolescents aged 12–18 years. It should not be used as a substitute for clinical diagnosis. Since the CES-DC shows only modest sensitivity and specificity at optimal cutoff points, it cannot be recommended for use as an instrument for estimating the prevalence of depressive disorder in children and adolescents in the community. The scale should, however, be considered for use as a screen for depression in epidemiologic studies using a two-stage sampling design or as an aid in case detection in clinical settings. If the goal of the screen is to detect cases of DSM-III major depressive disorder, then the conventional adult cutoff point of 16 and above is recommended. If the screen is designed to detect a broader group of depressive persons (i.e., including major depressive disorder or dysthymia), then a 15-and-above cutoff point is recommended. Clinicians seeking a quick screen as a preliminary step to a more detailed diagnostic assessment may consider using the abbreviated four-item version of the measure. Clinicians and researchers might consider using the CES-DC to identify potential cases of depression among children and adolescents who do not meet the criteria for clinical diagnosis. As a supplement to clinical diagnosis, the CES-DC might also facilitate the identification of children most in need of immediate intervention.

#### REFERENCES

1. Boyd JH, Weissman MM, Thompson WD, et al. Screening for depression in a community sample: understanding the discrepancies between depression symptom and diagnostic scales. *Arch Gen Psychiatry* 1982;39:1195–1200.
2. Weissman MM, Gammon GD, John KA, et al. Children of depressed parents: increased psychopathology and early onset of major depression. *Arch Gen Psychiatry* 1987;44:847–53.
3. Weissman MM, Orvaschel H, Padian N. Children's symptom and social functioning self-report scales: comparison of mothers' and children's reports. *J Nerv Ment Dis* 1980;168:736–40.
4. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977;1:385–401.
5. Weissman MM, Sholomskas D, Pottenger M, et al. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol* 1977;106:203–14.
6. Myers JK, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry* 1980;137:1081–4.
7. Schoenbach VJ, Kaplan BH, Grimson RC, et al. Use of a symptom scale to study the prevalence of a depressive syndrome in young adolescents. *Am J Epidemiol* 1982;116:791–800.
8. Schoenbach VJ, Kaplan BH, Wagner EH, et al. Prevalence of self-reported depressive symptoms in young adolescents. *Am J Public Health* 1983;73:1281–7.
9. Doerfler L, Felner RD, Rowlison RT, et al. Depression in children and adolescents: a comparative analysis of the utility and construct validity of two assessment measures. *J Consult Clin Psychol* 1988;56:769–72.
10. Saylor CF, Finch AJ, Spirito A, et al. The Children's Depression Inventory: a systematic evaluation of psychometric properties. *J Consult Clin Psychol* 1984;52:955–67.
11. Faulstich ME, Carey MP, Ruggiero L, et al. Assessment of depression in childhood and adolescence: an evaluation of the Center for Epidemiologic Studies Depression Scale for Children (CES-DC). *Am J Psychiatry* 1986;143:1024–7.
12. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3rd ed. Washington, DC: American Psychiatric Association, 1980.
13. Weissman MM, John K, Merikangas KR, et al. Depressed parents and their children: general health, social and psychiatric problems. *Am J Dis Child* 1986;140:801–5.
14. Weissman MM, Kidd KK, Prusoff BA. Variability in rates of affective disorders in relatives of depressed and normal probands. *Arch Gen Psychiatry* 1982;39:1397–1403.
15. Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: rationale and reliability. *Arch Gen Psychiatry* 1978;35:773–82.
16. Endicott J, Spitzer RL. A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. *Arch Gen Psychiatry* 1978;35:837–44.
17. Weissman MM, Myers JK, Harding PS. Psychiatric disorders in a U.S. urban community: 1975–1976. *Am J Psychiatry* 1978;135:459–62.
18. Orvaschel H, Puig-Antich J, Chambers W, et al. Retrospective assessment of prepubertal major depression with Kiddie SADS-E. *J Am Acad Child Adolesc Psychiatry* 1982;21:392–7.
19. Coopersmith S. The antecedents of self-esteem.

- San Francisco, CA: WH Freeman & Company, 1967.
20. John K. The Child Trait Checklist. Unpublished monograph, 1982.
  21. Dunn LM, Dunn LM. Peabody Picture Vocabulary Test—Revised. Circle Pines, MN: American Guidance Service, 1981.
  22. Shaffer D, Gould MS, Brasic J, et al. A children's global assessment scale (C-GAS). *Arch Gen Psychiatry* 1983;40:1228-31.
  23. Leckman JF, Scholomskas D, Thompson WD, et al. Best estimate of lifetime psychiatric diagnosis: a methodologic study. *Arch Gen Psychiatry* 1982; 39:839-83.
  24. Lord FM, Novick MR. Statistical theories of mental test scores. Reading, MA: Addison-Wesley Publishing Company, Inc, 1968.
  25. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 16:420-8.
  26. Tukey JW. Exploratory data analysis. Reading MA: Addison-Wesley Publishing Company, Inc, 1977.
  27. Fleiss JL. The design and analysis of clinical experiments. New York, NY: John Wiley & Sons, Inc, Publishers, 1986.
  28. Seligman MEP. Helplessness: on depression, development, and death. San Francisco, CA: WH Freeman & Company, 1975.
  29. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
  30. Erdreich LS, Lee ET. Use of relative operating characteristic analysis in epidemiology: a method for dealing with subjective judgement. *Am J Epidemiol* 1981;114:649-62.
  31. Shrout PE, Yager TJ. Reliability and validity of screening scales: effect of reducing scale length. *J Clin Epidemiol* 1989;42:69-78.

APPENDIX

Comparison of Center for Epidemiologic Studies Depression Scale (CES-D) items for adults and for children

Adult's item (CES-D)	Children's item (CES-DC*)
I was bothered by things that usually don't bother me.	Same.
I did not feel like eating; my appetite was poor.	I did not feel like eating; I wasn't very hungry.
I felt that I could not shake off the blues even with help from my family or friends.	I wasn't able to feel happy, even when my family or friends tried to help me feel better.
I felt that I was just as good as other people.	I felt like I was just as good as other kids.
I had trouble keeping my mind on what I was doing.	I felt like I couldn't pay attention to what I was doing this week.
I felt depressed.	I felt down and unhappy this week.
I felt that everything I did was an effort.	I felt like I was too tired to do things this past week.
I felt hopeful about the future.	I felt like something good was going to happen.
I thought my life had been a failure.	I felt like things I did before didn't work out right.
I felt fearful.	I felt scared this week.
My sleep was restless.	I didn't sleep as well as I usually sleep this week.
I was happy.	† I was happy this week.
I talked less than usual.	I was more quiet than usual this week.
I felt lonely.	† I felt lonely, like I didn't have any friends.
People were unfriendly.	† I felt like kids I knew were not friendly or that they didn't want to be with me.
I enjoyed life.	I had a good time this week.
I felt sad.	† Same.
I felt that people disliked me.	† I felt people didn't like me this week.
I could not get "going."	† It was hard to get started doing things this week.
Code (adult's response category)	Code (children's response category)
0 Rarely (<1 day)	0 Not at all
1 Some or a little of the time (1-2 days)	1 A little
2 Occasionally or a moderate amount of the time (3-4 days)	2 Some
3 Most or all of the time (5-7 days)	3 A lot

\* CES-DC, Center for Epidemiologic Studies Depression Scale for Children.  
 † Items used for the shortened CES-DC scale.