

# Comparisons of Problems Reported by Youths From Seven Countries

Frank C. Verhulst, M.D.  
 Thomas M. Achenbach, Ph.D.  
 Jan van der Ende, M.S.  
 Nese Erol, Ph.D.  
 Michael C. Lambert, Ph.D.  
 Patrick W.L. Leung, Ph.D.  
 Maria A. Silva, M.D.  
 Nelly Zilber, Dr. ès Sc.  
 Stephen R. Zubrick, B.Sc. M.Sc.,  
 A.M., Ph.D.

**Objective:** This study compared ratings for self-reported behavioral and emotional problems in adolescents from seven countries.

**Method:** Youth Self-Report scores were analyzed for 7,137 adolescents ages 11–18 years from general population samples from Australia, China, Israel, Jamaica, the Netherlands, Turkey, and the United States.

**Results:** Comparisons of problems scores yielded small to medium effect sizes for cross-cultural variations. Youths from China and Jamaica had the highest and youths from Israel and Turkey had the

lowest mean total problems scores. With cross-cultural consistency, girls scored higher for internalizing and lower for externalizing than boys. Cross-cultural correlations were high among the mean item scores.

**Conclusions:** Empirically based assessment provided a robust method for assessing and comparing adolescents' self-reported problems. Self-reports thus supplemented empirically based assessments of parent-reported problems and offered a cost-effective way of identifying problems for which adolescents from diverse cultural backgrounds may need help.

(*Am J Psychiatry* 2003; 160:1479–1485)

Mental health professionals, educators, and policy makers are increasingly challenged by cultural differences both between and within countries. Wars, famine, disasters, and ethnic and religious conflicts are uprooting people around the world. Among the victims are youths who are placed in unfamiliar cultural environments, sometimes without their families. Cultural differences associated with ethnic, linguistic, religious, and regional variations within countries may also pose challenges for determining which youths need professional help, even when they are not displaced from their homes or families.

When evaluating youths of different cultures, mental health professionals must often determine whether ostensible problems reflect merely cultural differences or whether they reflect need for professional help. To assist in making such determinations, standardized assessment instruments should be applicable to youths from various cultures by diverse professionals under different conditions. To be practical and cost-effective, the instruments should not require much professional time, specialized training, or commitment to particular theories of psychopathology. Instead, they should quickly obtain standardized information that will assist a variety of relevant decision-makers.

Because many cultures lack well-standardized indigenous instruments for assessing the problems of youths, instruments developed in one culture are often translated and adapted for use in other cultures. To apply such instruments to new cultures, they should be tested in various ways to maximize the equivalence of data obtained in the different cultures (1). A key step that both enhances the use

of assessment instruments and contributes important knowledge about the distribution of psychopathology within a country is the assessment of epidemiological samples that are representative of a country's population. Data from representative samples are valuable for determining the prevalence of particular problems in a population, for identifying important differences in the number of members of particular groups, such as women and men, and for constructing norms. The norms can then be used to identify deviance in individuals who are subsequently assessed with the instrument that was used in identifying the epidemiological sample. When an assessment instrument has been applied to epidemiological samples in multiple cultures, the results can then be used to identify similarities and differences in the rates of problems from one culture to another and for particular groups, such as women and men, across cultures.

Most cross-cultural epidemiological studies have compared only two cultures at a time; we (2) have provided a review of bicultural comparisons. Although such comparisons reveal similarities and differences between the rates of problem in two cultures, simultaneous comparisons of more cultures make it possible to test the overall range of variation across multiple cultures, to determine where each culture falls within that range, and to identify effects associated with variables such as gender and age with much more confidence than in bicultural comparisons. Multicultural comparisons that include measures of both broad and narrow bands (3, 4) can reveal whether certain

TABLE 1. Effects of Culture, Age, and Gender on Youth Self-Report Scale Scores for 7,137 Adolescents From Seven Cultures<sup>a</sup>

Item	Culture		Age		Gender		Age by Gender		Age by Culture	
	F (df=6, 7035)	Effect Size (%) <sup>b</sup>	F (df=7, 7035)	Effect Size (%) <sup>b</sup>	F (df=1, 7035)	Effect Size (%) <sup>b</sup>	F (df=7, 7035)	Effect Size (%) <sup>b</sup>	F (df=37, 7035)	Effect Size (%) <sup>b</sup>
Total problems	58.7*	5	7.2 <sup>c*</sup>	<1	15.4 <sup>d*</sup>	<1	1.1	—	1.8*	<1
Internalizing	68.7*	6	6.5 <sup>c*</sup>	<1	130.7 <sup>d*</sup>	2	2.3	—	2.1*	1
Externalizing	46.1*	4	10.1 <sup>c*</sup>	1	22.1 <sup>e*</sup>	<1	1.0	—	1.7*	<1
Withdrawn	71.9*	6	11.4 <sup>c*</sup>	1	31.9 <sup>d*</sup>	<1	2.2	—	1.7*	<1
Somatic complaints	62.4*	5	1.3	—	128.1 <sup>d*</sup>	2	1.8	—	2.4*	1
Anxious/depressed	66.7*	5	5.5 <sup>c*</sup>	<1	113.9 <sup>d*</sup>	2	2.3	—	1.4	—
Social problems	72.2*	6	1.9	—	2.1	—	0.7	—	1.5*	—
Thought problems	105.7*	8	3.1 <sup>c*</sup>	<1	9.0 <sup>d*</sup>	<1	1.0	—	2.2*	1
Attention problems	50.1*	4	11.6 <sup>c*</sup>	1	0.3	—	0.8	—	1.7*	<1
Delinquent behavior	75.5*	6	22.3 <sup>c*</sup>	2	69.3 <sup>e*</sup>	1	1.7	—	2.7*	1
Aggressive behavior	36.5*	3	4.6 <sup>c*</sup>	<1	5.9	—	0.8	—	1.5	—

<sup>a</sup> A total of 3,611 boys, 3,526 girls; for Australia, N=576, ages 12–16; for China, N=1,599, ages 12–18; for Israel, N=614, ages 11–17; for Jamaica, N=400, ages 11–18; for the Netherlands, N=1,098, ages 11–18; for Turkey, N=1,341, ages 11–18; and for the United States, N=1,509, ages 11–18.

<sup>b</sup> Effect size indicates percentage of explained variance; given only for significant effects ( $p < 0.01$ ).

<sup>c</sup> Older subjects scored higher than younger subjects; all significant age effects reflected linear trends, with delinquent behavior also reflecting a quadratic trend.

<sup>d</sup> Girls scored higher than boys.

<sup>e</sup> Boys scored higher than girls.

\* $p < 0.01$ .

cultures are exceptionally high or low in reports of general versus specific kinds of problems.

Most cross-cultural comparisons of the problems of children and youths (2) have been based on reports by parents or teachers. When available, parents and teachers are certainly important sources of data about such problems. However, meta-analyses of many studies using different instruments for assessing children and youths (5) have yielded mean correlations of only 0.27 between parent and teacher reports, 0.25 between parent and self-reports, and 0.20 between teacher and self-reports. Although statistically significant, these modest correlations have indicated that no one type of informant can substitute for the others. Instead, comprehensive assessment requires data from multiple informants whenever possible. For displaced youths, parent and teacher reports may be unavailable. Even when they are available, however, such reports cannot fully substitute for youths' own reports of their problems. To evaluate the degree to which cultural differences may be associated with youths' reports of their problems, it is necessary to compare such reports obtained by means of the same standardized procedures in different cultures.

In this study, we wanted to evaluate the degree to which youths' reports of their own problems differed across diverse cultures, so we compared scores on problems scales from the Youth Self-Report (6) completed by 7,137 11–18-year-olds from Australia, China, Israel, Jamaica, the Netherlands, Turkey, and the United States. To evaluate cross-cultural variations in overall problem levels, we compared total problems scores; to identify possible cross-cultural differences in specific kinds of problems, we compared scores for eight empirically based syndromes and scores for internalizing and externalizing groupings of syndromes. In addition, to evaluate effects associated with

gender and age, we compared scores for girls and boys and for different ages between and across cultures.

## Method

### Assessment

The Youth Self-Report (6) is a questionnaire designed to be completed by adolescents ages 11–18 years and contains 101 problem items. The problem items are scored as follows: 0=not true, 1=somewhat or sometimes true, and 2=very true or often true on the basis of the preceding 6 months. The Youth Self-Report can be scored on the total problems scale, which is the sum of the scores for each problem item, and the following eight syndrome scales: withdrawn, somatic complaints, and anxious/depressed (together comprising the broadband internalizing scale); social problems, thought problems, and attention problems (which are not part of either the internalizing or externalizing scales); and delinquent behavior and aggressive behavior (together comprising the externalizing scale).

For four of the seven cultures (China, Israel, the Netherlands, and Turkey), where languages other than English are spoken, the Youth Self-Report was translated and back-translated to approximate the original version as closely as possible.

The reliability and validity of the Youth Self-Report are documented by Achenbach (6). Confirmatory factor analyses of parent, teacher, and self-reports of referred Dutch children (7) supported the overall syndrome structure. Because the availability of reliability and validity data varied across the seven cultures, we computed Cronbach's alphas for each of the 11 Youth Self-Report scales in each culture. The ranges were for withdrawn (0.52–0.64), somatic complaints (0.65–0.76), anxious/depressed (0.79–0.86), social problems (0.46–0.64), thought problems (0.49–0.69), attention problems (0.64–0.74), delinquent behavior (0.51–0.70), aggressive behavior (0.76–0.83), internalizing (0.83–0.89), externalizing (0.82–0.86), and total problems (0.92–0.95).

### Samples

To qualify for inclusion, samples from each culture were required to include completed Youth Self-Reports for at least 75% of the target informants. All samples involved randomized selection from the general population. Data were obtained for adolescents from the following cultures:

Gender by Culture		Age by Gender by Culture	
F (df=6, 7035)	Effect Size (%) <sup>b</sup>	F (df=37, 7035)	Effect Size (%) <sup>b</sup>
2.0	—	1.0	—
2.1	—	1.1	—
1.4	—	0.8	—
1.3	—	1.1	—
3.4*	<1	1.1	—
1.5	—	1.0	—
2.9*	<1	1.1	—
2.3	—	1.1	—
2.2	—	1.3	—
2.0	—	0.6	—
1.2	—	0.9	—

1. Australia (8): The sampling frame was all households in Western Australia. The procedure was random sampling of all households and all children per household. In families in which there was more than one 12–16-year-old, one child was randomly selected for this study. The rate of response was 91%, or 576, of the 12–16-year-olds.
2. China (unpublished data of P.W.L.L.): The sampling frame was all schools from the city of Hong Kong. The procedure was random sampling of the schools; within each class, two students were randomly selected. The rate of response was 86%, or 1,599, of the 12–18-year-olds.
3. Israel (unpublished manuscript of N.Z.): The sampling frame was all households in Jerusalem. The procedure was random sampling of all households; within each household, one child was selected. Only Israeli-born Jewish children were included. The rate of response was 83%, or 614, of the 11–17-year-olds.
4. Jamaica (9): The sampling frame was all schools from the Kingston and Montego Bay areas and rural areas throughout Jamaica. The procedure was random sampling of schoolchildren; within each school, classes from each grade level were randomly selected, and within each class, one adolescent was randomly selected. The rate of response was 90%, or 400, of the 11–18-year-olds.
5. The Netherlands (10): The sampling frame was all children with Dutch nationality throughout the Netherlands. The procedure was a two-stage sampling of municipalities, followed by random selection from municipal registries. The rate of response was 78%, or 1,098, of the 11–18-year-olds.
6. Turkey (11): The sampling frame was all children of Turkish nationality throughout Turkey. The procedure was random sampling of households stratified by region and type of settlement. One child per household was randomly selected. The rate of response was 79%, or 1,341, of the 11–18-year-olds.
7. United States (6): The sampling frame was all households in the 48 contiguous states. The procedure was initial multi-stage random selection of one child per household between ages 4 and 16, with assessment 3 years later of subjects between ages 11 and 18. The rate of response was 89%, or 1,509, of the 11–18-year-olds.

## Analyses

Analyses of variance (ANOVA) were performed with the general linear modeling procedure of SPSS version 9.0 (Chicago, SPSS) that handles unbalanced data and empty cells. We assumed that there was simple random sampling in each of the countries, although in some countries, complex survey designs were used. However, information about sampling in several countries was too limited to include in our analyses. Therefore, estimates of parameters (e.g., means) and precision may have been biased.

## Results

ANOVAs for total problems, internalizing, externalizing, and scores for the eight syndrome scales were performed in a seven-(culture)-by-eight-(ages 11–18)-by-two (gender) factorial design. Age ranges differed somewhat by country (Australia, 12–16 years; China, 12–18 years; Israel, 11–17 years; and Jamaica, the Netherlands, Turkey, the United States, 11–18 years). To test for age effects, a polynomial contrast was applied. Linear and quadratic trends for age were tested.

In view of the high statistical power afforded by the large sample size ( $N=7,137$ ), we reported only the effects that were significant at  $p<0.01$ . Effect sizes were expressed as the percent of explained variance, and they were interpreted according to Cohen's criteria (12) as small (1.0% to <5.9% of variance), medium (5.9% to <13.8%), and large (13.8% or more).

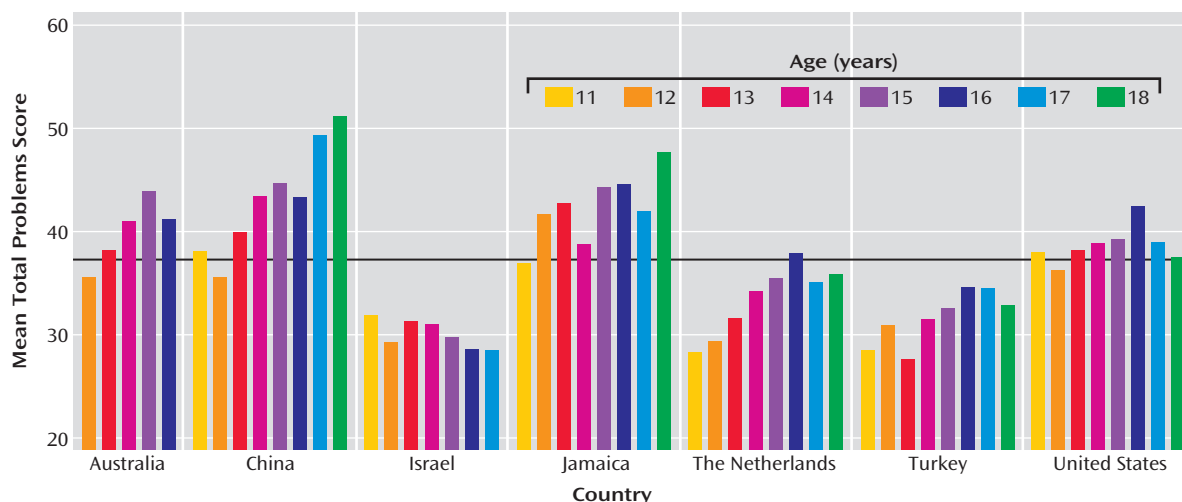
Results of the ANOVAs are displayed in Table 1, which shows F values and effect sizes for culture, age, gender, and their interactions for the Youth Self-Report total problems, internalizing, externalizing, and eight syndrome scales. Table 1 also shows that for all significant age effects, the trend was linear, with delinquent behavior also showing a quadratic age effect. For each scale, Table 2 shows the overall mean score and significant deviations from the overall mean for each culture.

Figure 1 shows the deviation from the overall mean score for Youth Self-Report total problems for each culture, by age.

The mean total problems scores for each culture deviated significantly from the overall mean of 37.6 (Table 2). The largest deviations were for China (mean=6.6) and Israel (mean=-7.2). Culture accounted for 5% of the variance in total problems scores (Table 1), which is a small effect size, according to Cohen's criteria (12). Age accounted for <1% of the variance, with older adolescents scoring higher than younger adolescents. A significant age-by-culture interaction, accounting for <1% of the variance in total problems scores, reflected cross-cultural differences in age effects, with the steepest increase with age for China and an absence of an increase for Israel. Girls earned significantly higher total problems scores than boys. There were no interactions between gender and culture or between gender and age. There was no three-way interaction between culture, gender, and age.

**TABLE 2.** Significant Deviations From Overall Mean Scores on Youth Self-Report Scales for 7,137 Adolescents From Seven Cultures

Scale	Overall Score		Significant <sup>a</sup> Deviation From Overall Mean Score						
	Mean	SD	Australia	China	Israel	Jamaica	The Netherlands	Turkey	United States
Total problems	37.6	21.0	2.6	6.6	-7.2	5.3	-3.6	-5.3	1.7
Internalizing	12.2	8.1		2.4	-3.1	3.7	-2.7	—	—
Externalizing	10.4	6.9	1.3	0.8		—	—	-2.6	1.5
Withdrawn	3.5	2.4		0.3	-0.4	1.4	-1.1	—	—
Somatic complaints	2.7	2.7	0.8	0.4	-1.4	1.1	-0.2	-0.6	—
Anxious/depressed	6.4	4.9		1.8	-1.5	1.3	-1.5	0.6	-0.4
Social problems	2.8	2.3	-0.4	0.9	-0.9	1.1	-0.3	-0.3	-0.2
Thought problems	2.0	2.1	1.0		-0.4	0.6	-0.8	-0.8	0.4
Attention problems	4.6	3.1		1.1	-0.7	-0.3	0.4	-0.7	0.4
Delinquent behavior	2.7	2.4	0.8	-0.2	-0.5	—	0.7	-1.0	0.2
Aggressive behavior	7.7	5.2	0.5	1.0		-0.7	—	-1.2	0.9

<sup>a</sup>  $p < 0.01$ .**FIGURE 1.** Mean Scores for Total Problems on the Youth Self-Report Scale of 7,137 Adolescents From Seven Cultures, by Age

For internalizing, externalizing, and the eight syndrome scales, ANOVAs revealed significant effects of culture on each scale. The effects of culture were small, according to Cohen's criteria (12) for five scales and medium for the other five scales. The deviations from the overall mean of each scale are listed in Table 2. For eight of the 10 scales, significant age effects were found, with higher scores for older youths. The effect sizes were small for four and  $<1\%$  for the other scales. Significant gender effects were found for seven of the 10 scales. The effect sizes were small for four, and  $<1\%$  for the other scales. All effect sizes reflected higher scores for girls than boys, except for the externalizing and delinquent behavior scales, on which boys scored higher than girls.

No two-way interactions between age and gender were found. For seven of the 10 scales, significant age-by-culture interactions were found. However, these effect sizes were small (1%) or very small ( $<1\%$ ). For the internalizing, externalizing, withdrawn, attention problems, and delinquent behavior scales, the interactions reflected differences in the increase in scores with age across cultures, with the steepest increases for China, Australia, and the Netherlands, whereas Israel showed the smallest increase

in scores with age. For somatic complaints, Israel and the United States showed a decrease, whereas China, Australia, and Jamaica showed an increase with age. The Netherlands and Turkey did not show changes with age. For the thought problems scale, all cultures showed increases with age, except Israel and Turkey. Only two interactions between gender and culture were found, both accounting for  $<1\%$  of the variance. For the somatic complaints scale, the difference between boys and girls was much smaller for Israel and Turkey than the gender differences found for the other cultures, although for each culture, the girls scored higher than the boys. On the social problems scale, the boys scored higher than the girls in Israel, Australia, the Netherlands, and the United States, whereas in the other cultures, the girls scored higher than the boys. There were no significant three-way interactions between age, gender, and culture.

To compare the relative magnitude of scores on individual Youth Self-Report problem items across cultures, we computed Pearson correlations between the means of the scores on each problem item in each pair of cultures (Table 3). All bicultural correlations were significant at  $p < 0.001$ . To compute mean correlations, correlations were

**TABLE 3. Bicultural Correlations Between Mean Item Scores on the Youth Self-Report Scale for 7,137 Adolescents From Seven Cultures<sup>a</sup>**

Culture	Country	Mean Correlation With All Other Countries (r)	Correlation With Each Other Country (r)						
			1	2	3	4	5	6	7
1	Australia	0.77	—						
2	China	0.74	0.73	—					
3	Israel	0.76	0.77	0.71	—				
4	Jamaica	0.70	0.69	0.70	0.69	—			
5	The Netherlands	0.73	0.78	0.74	0.75	0.59	—		
6	Turkey	0.69	0.57	0.75	0.73	0.76	0.60	—	
7	United States	0.83	0.92	0.80	0.88	0.77	0.83	0.70	—
Overall		0.75							

<sup>a</sup> All correlations were significant ( $p < 0.001$ ). Mean correlations were computed from Fisher's  $z$  transformations.

converted by means of Fisher's  $z$  transformations, and the mean correlation per country and the overall mean correlation were computed. The mean correlations per country ranged from 0.69 (for Turkey) to 0.83 (for the United States), with an overall mean of 0.75. All mean correlations were large, according to Cohen's criteria (12).

## Discussion

Comparisons of problems reported by adolescents ages 11–18 years from seven cultures ( $N = 7,137$ ) yielded a small effect size of 5% for cross-cultural variations in Youth Self-Report total problems scores. The largest deviation above the overall mean was for China, while the largest deviation below the overall mean was for Israel. The other five cultures had mean total problems scores that were clustered within five points of the overall mean of 37.6 derived from all cultures. Although the Youth Self-Report was used for all cultures, differences in methods, including variations in sampling procedures, variations in sample heterogeneity, and vicissitudes of translation may have contributed to these cross-cultural variations. Therefore, differences could be secondary to variations in translations and sampling. However, despite these variations in methods, most differences were small.

Crijnen et al. (3, 4) compared parent-reported problems with the Child Behavior Checklist (13) in 12 cultures across ages 6–11 and nine cultures across ages 6–17. Five cultures in the nine-culture comparison were represented in the present study (Australia, Israel, Jamaica, the Netherlands, and the United States). Deviations from the overall mean score of parent-reported total problems were in the same direction as the deviations from the overall mean score of self-reported total problems for Israel, the Netherlands, and the United States. For Jamaica and Australia, Youth Self-Report scores were above the overall mean, whereas Child Behavior Checklist scores did not differ significantly from the overall mean. For China, Child Behavior Checklist scores were above the overall mean for ages 6–11, as were Youth Self-Report scores in the present study for ages 12–18, despite the fact that the samples were drawn from different parts of Hong Kong. In general, cross-cultural variations were consistent across parent-reported and self-reported problems.

The effect sizes of culture for the internalizing and externalizing scales were medium (6%) and small (4%), respectively. The cultures that scored the highest on the total problems scale (China and Jamaica) also scored the highest on the internalizing scale. Of the seven cultures in our comparison, Israel and the Netherlands had the lowest internalizing scores. Deviations from the overall mean externalizing scale score were smaller than the deviations from the overall mean internalizing scale score. Turkey scored the lowest on the externalizing scale.

The effect size of culture for the eight Youth Self-Report syndrome scales yielded the largest cross-cultural variation for the thought problems scale and the smallest for the aggressive behavior scale. The deviations from the mean scale scores for each culture were not always consistent across all syndrome scales.

The variations in problems scores among cultural groups may have important implications for evaluating individual youths. For example, higher levels of particular problems in immigrant youths versus indigenous youths may result from stress factors associated with immigration. This may alert clinicians to the need for special help. However, if the level of problems in immigrant youths is comparable to that of youths from their native country, the higher levels of problems in immigrant youths over indigenous youths may reflect cross-cultural differences in the reporting of problems (14). These cross-cultural differences may result from differences in thresholds for reporting particular kinds of problems, from linguistic differences, or from true differences in the prevalence of these problems. This can alert clinicians to the need for special help without automatically implying that the higher level of problems constitutes the presence of disorders.

The largest effect size for culture was found for the thought problems scale (8%). This scale consists of items such as "I hear sounds or voices that other people think aren't there," "I see things that other people think aren't there," "I do things other people think are strange," and "I have thoughts that other people would think are strange." These problems involve an adolescent's interpretations of other people's standards and may be more sensitive to cultural influences than problems that are more straightforward and do not require such interpretations.



Cross-cultural differences in scores for particular syndromes can generate ideas for more detailed investigations of factors that influence problems. To illuminate this, the finding that adolescents in China had much higher scores on the anxious/depressed scale than adolescents in United States stressed the need for other sources of data, such as parent or teacher reports, classroom observations, or interviews. If multiple sources of data agree in showing higher levels of anxiety and depression for Chinese than U.S. adolescents, this may imply that culturally related factors are involved in the development of anxiety and depression.

Gender differences were similar across the seven cultures for total problems, internalizing, and externalizing, with girls scoring higher than boys on total problems and the internalizing scales, and boys scoring higher than girls on the externalizing scale. These gender differences by type of self-reported problems were consistent with the gender differences found for parent-reported internalizing and externalizing problems (3). Except for the somatic complaints scale, there was also cross-cultural consistency in gender differences for the syndrome scales, with girls scoring higher than boys on the withdrawn, anxious/depressed, and thought problems scales and boys scoring higher than girls on the delinquent behavior scale. For the three syndrome scales (somatic complaints, anxious/depressed, and delinquent behavior) with significant gender differences in both the present study and the study of parent-reported problems (4), the differences were in the same direction. Despite the range in cultural, economic, political, and genetic differences, there was consistency in both self- and parent reports of boys having more externalizing and girls having more internalizing problems.

Age differences were less consistent across the eight cultures, although significant effects of the age-by-culture interaction did not exceed 1%. Generally, a majority of cultures scored higher with increasing age on most scales.

### Implications

The relatively small cross-cultural differences in mean problems scores, the cross-cultural similarity in gender differences, and the high bicultural correlations between item scores found in the present study indicated that empirically based standardized self-reports can provide methodologically sound information across diverse cultures. Adolescents from different cultures responded in fairly similar ways to the problem items of the Youth Self-Reports, despite large variations in language, customs, religion, socioeconomic circumstances, and health care systems.

In an earlier study, the use of cross-cultural, standardized, empirically based parental ratings was documented (3). This approach can thus be supplemented with standardized self-reports to form a robust, empirically based assessment method for cross-cultural comparisons of adolescents both within and across individual countries. This is important because adolescents have different per-

spectives on their problems than do parents or teachers. Adolescents typically report more problems than parents or teachers do about them (15, 16). The use of adolescents' standardized self-reports can cost-effectively provide clinicians with appropriate norms against which individual adolescents' problems scores can be evaluated. This holds true both for adolescents with indigenous as well as immigrant backgrounds. For example, the comparison of adolescent immigrants' or refugees' problems scores with those obtained for adolescents in the host country, as well as with adolescents in the native country, both on parent and self-reports, can guide our understanding of the nature of these problems, as well as the development and provision of mental health services. This approach may also help us understand an individual's problems within a cultural context and may guide more detailed assessments and treatment strategies.

The use of the same assessment procedure to obtain standardized parent reports has facilitated international communication, training, and research. However, this is much less so for adolescents' self-reports of their behavioral and emotional problems. Epidemiological studies containing standardized clinical interviews of adolescents in diverse countries are available (10), but these approaches are costly because they are time-consuming and require intensive training of interviewers. The use of standardized self-reports enables us to cost-effectively gather data on large normative samples, which is important for cross-cultural research as well as for evaluating an individual's problems by comparing the individual's scores on each syndrome with those obtained for normative samples of the same age and gender.

Consistent with the gender differences in parent-reported problems, the present study found great cross-cultural consistency in higher externalizing scores for boys and higher internalizing scores for girls.

---

Received March 7, 2002; revisions received Aug. 16 and Dec. 6, 2002, accepted Feb. 12, 2003. From the Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia; the Center for Children, Youth and Families, Department of Psychiatry, University of Vermont, Burlington, Vt.; the Department of Child and Adolescent Psychiatry, Ankara University, Ankara, Turkey; Michigan State University, East Lansing, Mich.; the Department of Psychology, Chinese University of Hong Kong, China; the C.N.R.S. Centre de Recherche Français de Jérusalem, Jerusalem; the Falk Institute for Mental Health and Behavioral Studies, Jerusalem; and the Curtin University Centre for Developmental Health, TVW Telethon Institute for Child Health Research, West Perth, Australia. Address reprint requests to Dr. Verhulst, Department of Child and Adolescent Psychiatry, Erasmus MC-Sophia, dr Molewaterplein 60, 3015 GJ Rotterdam, the Netherlands: m.engel@erasmusmc.nl (e-mail).

Funded by NIMH grant MH-40305, the Sophia Foundation for Medical Research, the Western Australian Health Promotion Foundation (grant number 0253/91), and the Australian Rotary Health Research Fund.

---

### References

1. Canino G, Bravo M: The translation and adaptation of diagnostic instruments for cross-cultural use, in *Diagnostic Assessment*

- in Child and Adolescent Psychopathology. Edited by Shaffer D, Lucas CP, Richters JE. New York, Guilford, 1999, pp 285–298
2. Verhulst FC, Achenbach TM: Empirically based assessment and taxonomy of psychopathology: cross-cultural applications: a review. *Eur Child Adolesc Psychiatry* 1995; 4:61–76
3. Crijnen AAM, Achenbach TM, Verhulst FC: Comparisons of problems reported by parents of children in 12 cultures: total problems, externalizing, and internalizing. *J Am Acad Child Adolesc Psychiatry* 1997; 36:1269–1277
4. Crijnen AAM, Achenbach TM, Verhulst FC: Problems reported by parents of children in multiple cultures: the Child Behavior Checklist syndrome constructs. *Am J Psychiatry* 1999; 156:569–574
5. Achenbach TM, McConaughy SH, Howell CT: Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull* 1987; 101:213–232
6. Achenbach TM: Manual for the Youth Self-Report and 1991 Profiles. Burlington, University of Vermont Department of Psychiatry, 1991
7. De Groot A, Koot HM, Verhulst FC: Cross-cultural generalizability of the Youth Self-Report and Teacher's Report Form cross-informant syndromes. *J Abnorm Child Psychol* 1996; 24:651–664
8. Zubrick SR, Silburn SR, Gurrin L, Teoh H, Shepherd C, Carlton J, Lawrence D: Western Australian Child Health Survey: Education, Health and Competency. Perth, Western Australia, Australian Bureau of Statistics, 1997
9. Lambert MC, Lyubansky M, Achenbach TM: Behavioral and emotional problems among adolescents of Jamaica and the United States: parent, teacher, and self-reports for ages 12 to 18. *J Emotional and Behavioral Disorders* 1998; 6:180–187
10. Verhulst FC, van der Ende J, Ferdinand RF, Kasius MC: The prevalence of DSM-III-R diagnoses in a national sample of Dutch adolescents. *Arch Gen Psychiatry* 1997; 54:329–336
11. Erol N, Simsek Z: Mental health of Turkish children: behavioral and emotional problems reported by parents, teachers and adolescents, in *International Perspectives on Child and Adolescent Mental Health*. Edited by Singh NN, Leung JP, Singh AN. Oxford, UK, Elsevier Science, 2000, pp 223–247
12. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, Lawrence Erlbaum Associates, 1988
13. Achenbach TM: Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, University of Vermont, Department of Psychiatry, 1991
14. Bengi-Arslan L, Verhulst FC, van der Ende J, Erol N: Understanding childhood (problem) behaviors from a cultural perspective: comparison of problem behaviors and competencies in Turkish immigrant, Turkish and Dutch children. *Soc Psychiatry Psychiatr Epidemiol* 1997; 32:477–484
15. Achenbach TM: Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles. Burlington, University of Vermont, Department of Psychiatry, 1991
16. Verhulst FC, Dekker MC, van der Ende J: Parent, teacher and self-reports as predictors of signs of disturbance in adolescents: whose information carries the most weight? *Acta Psychiatr Scand* 1997; 96:75–81