Behavioral and Emotional Problems Reported by Parents of Children Ages 6 to 16 in 31 Societies

Leslie Rescorla, Thomas Achenbach, Masha Y. Ivanova, Levent Dumenci, Fredrik Almqvist, Niels Bilenberg, Hector Bird, Wei Chen, Anca Dobrean, Manfred Döpfner, Nese Erol, Eric Fombonne, Antonio Fonseca, Alessandra Frigerio, Hans Grietens, Helga Hannesdottir, Yasuko Kanbayashi, Michael Lambert, Bo Larsson, Patrick Leung, Xianchen Liu, Asghar Minaei, Mesfin S. Mulatu, Torunn S. Novik, Kyung-Ja Oh, Alexandra Roussos, Michael Sawyer, Zeynep Simsek, Hans-Christoph Steinhausen, Sheila Weintraub, John Weisz, Christa Winkler Metzke, Tomasz Wolanczyk, Hao-Jan Yang, Nelly Zilber, Rita Zukauskiene, and Frank Verhulst

This study compared parents' ratings of behavioral and emotional problems on the *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001) for general population samples of children ages 6 to 16 from 31 societies ($N = 55,508$). Effect sizes for society ranged from .03 to .14. Effect sizes for gender were ≤ .01, with girls generally scoring higher on Internalizing problems and boys generally scoring higher on Externalizing problems. Effect sizes for age were ≤ .01 and varied across types of problems. Total Problems scores for 19 of 31 societies were within 1 *SD* of the overall mean of 22.5. Bisociety correlations for mean item scores averaged .74. The findings indicate that parents' reports of children's problems were similar in many ways across highly diverse societies. Nonetheless, effect sizes for society were larger than those for gender and age, indicating the need to take account of multicultural variations in parents' reports of children's problems.

Children of immigrant parents constitute increasing proportions of populations served by mental health, educational, medical, and welfare systems in many societies. In addition, assessment of needs for child mental health services is a significant public health goal around the world. To meet these challenges for assessment of behavioral and emotional problems in diverse societies, there is a need for instruments that are easily administered, scored, and interpreted by a wide range of practitioners and researchers and that demonstrate multicultural robustness. Multicultural robustness is established through systematic research demonstrating that an instrument performs similarly across many societies in terms of features such as reliability, internal consistency, factor structure, scale scores, and associations of scores with age and gender (Geisinger, 1994).

In the early stages of multicultural research, mental health specialists in a society often evaluate instruments developed in other societies for use in their own. If an instrument is in a foreign language, a translated version is created, and then an independent back-translation into the original language is done to verify that the translation captures the meaning of the original. Ideally, researchers then collect data using the instrument with a large general population sample. When data are available from many societies, they can be analyzed together to compare variations between societies and within societies. Establishing the multicultural robustness of an instrument is thus an incremental process using an *etic* approach to research, whereby the same standardized assessment instrument is used in different societies. This contrasts with an *emic* approach to research, whereby the meanings of the constructs under study are explored within each society.

MULTICULTURAL RESEARCH USING THE CHILD BEHAVIOR CHECKLIST

The present study contributes to an international program of collaborative research investigating the multicultural robustness of the *Child Behavior Checklist* (CBCL; Achenbach, 1991; Achenbach & Rescorla, 2001). The CBCL for ages 6 to 18 obtains parents' reports of children's behavioral and emotional problems and competencies. Scores are obtained for (a) 118 specific problem items plus 2 open-ended items; (b) eight empirically based syndromes, derived using factor analysis (Achenbach, 1991; Achenbach & Rescorla, 2001); (c) two scales derived from second-order factor analyses of the eight syndromes, one labeled Internalizing and comprising the Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints syndromes, and the other labeled Externalizing and comprising the Rule-Breaking Behavior and Aggressive Behavior syndromes; and (d) Total Problems, which consists of the sum of ratings on all 120 problem items. In addition, the 2001 revision to the CBCL features six new scales composed of items judged by clinicians from 16 societies to be consistent with diagnostic categories of the American Psychiatric Association's fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–IV*, 1994; Achenbach, Dumenci, & Rescorla, 2003). The *DSM*-oriented scales are designated as Affective Problems, Anxiety Problems, Somatic Problems, Attention-Deficit/Hyperactivity Problems, Oppositional Defiant Problems, and Conduct Problems. CBCL items span diverse behavioral and emotional problems but are not assumed to include every possible problem that might be reported for every child in every society.

CBCL findings for general population samples have been published from more than 25 societies, including societies as different as Puerto Rico (Achenbach et al., 1990), Germany (Döpfner et al., 1997), Turkey (Erol & Simsek, 1997), Hong Kong (Leung et al., 2006), Thailand (Weisz et al., 1987), and Israel (Zilber, Auerbach, & Lerner, 1994). Some of these published studies reported reliability and validity data, factor structures, demographic effects, and bisociety comparisons with data from the United States. Achenbach and Rescorla (2007) have reviewed findings from these and other international studies that tested associations between CBCL scores and demographic variables, psychiatric diagnoses, referral status, longitudinal outcomes, and genetic factors.

To move beyond bisociety comparisons, Crijnen, Achenbach, and Verhulst (1997, 1999) analyzed CBCL data from general population samples in 12 societies (*N* = 13,697). When effects of society, gender, and age were tested using analysis of variance (ANOVA), effect sizes (ES) for society ranged from .01 to .11 for the 1991 versions of the CBCL scales. Across societies, girls scored significantly higher than boys on Somatic Complaints, Anxious/Depressed, and Internalizing, whereas boys scored higher than girls on Attention Problems, Delinquent Behavior (Crijnen et al., 1997, 1999) and Externalizing (ES ≤ .01). However, Crijnen et al. did not report findings regarding how consistent the gender effects were in the various societies,

nor did they test the internal consistency of CBCL scales in each society or the effects of methodological factors on CBCL scores.

THE CURRENT STUDY

The current study was designed to test the hypothesis that the 2001 versions of CBCL scales would be multiculturally robust across a larger and more diverse set of societies than the 12 societies analyzed by Crijnen et al. (1997, 1999) and according to a broader array of statistical tests than those used by Crijnen et al. We tested multicultural robustness with respect to (a) the internal consistency of the CBCL scales; (b) the effects of society, gender, and age on CBCL scores; and (c) the rank ordering of mean item scores. We did not test the multicultural robustness of the CBCL's factor structure, because Ivanova et al. (in press) reported that confirmatory factor analyses (CFAs) of the eight-syndrome model, derived primarily from U.S. data, fit the CBCL data from the 30 non–U.S. societies analyzed in the present study.

To test the CBCL's multicultural robustness, we analyzed data sets from 31 societies: 12 from Western Europe, 5 from Eastern Europe, 6 from Asia, 1 from Africa, 3 from the Middle East, 2 from the Caribbean, plus Australia and the United States. The 31 samples consisted of 9 analyzed by Crijnen and 22 others not included in previous multicultural comparisons. We use the general term *society* because some of our samples were obtained from national surveys of entire countries (e.g., United States), others were obtained from one region or province in a country (e.g., Shandong, China), and still others were obtained from geographical entities that are not countries (e.g., Hong Kong and Puerto Rico). Details of each sample are provided in Table 1.

To address our research questions, we first tested the internal consistency of each CBCL scale within each society and similarities among internal consistencies across all 31 societies. Second, we tested the effects of society, gender, and age on the 2001 versions of the 11 CBCL scales analyzed by Crijnen et al. (Total Problems, Internalizing, Externalizing, and the eight syndromes), as well as on the six *DSM*-oriented scales newly published in 2001. Third, we conducted within-society tests of gender and age differences on every scale that had a significant gender or age effect in the overall analysis. Finally, to assess the effects of various methodological factors on CBCL scores, we tested data collection methods, response rates, and exclusion of referred children as predictors of problem scores.

METHOD

Samples

Following the recommendation of Nunnally and Bernstein (1994), we required a minimum sample size of 300 per society. As presented in Table 1, sample sizes ranged from 513 for Ethiopia to 4,858 for China. For analyses of ages 6 to 16 years,

TABLE 1
Sample Characteristics for 31 Societies

Society	Reference	n	Ages (yrs)	Sampling frame	Response rate (%)	Method used	Referred excluded?
Australia	Sawyer et al., 2001	3,078	6–16	National household sample	86	HI	no
Belgium	Hellinckx et al., 1994	953	6–11	Regional sample in medical clinics	80	M	yes
China	Liu et al., 1999, 2000	4,858	6–15	Regional household sample	92	S	no
Denmark	Bilenberg, 1999	628	6–16	Regional household sample	60	M	no
Ethiopia	Mulatu, 1997	513	11–16	Regional school-based sample	92	HI	no
Finland	Weintraub, 2004	2,093	6–16	Regional school-based sample	77	S	no
France	Fombonne & Vermeersch, 1997	2,133	6–16	Employees of national utility	62	M	yes
Germany	Döpfner et al., 1997	2,184	6–16	National household media sample	82	HI	no
Greece	Roussos et al., 1999	1,113	6–11	National school-based sample	95	S	no
Hong Kong	Leung et al., 2006	2,033	6–16	City school-based sample	91	S	yes
Iceland	Hannesdottir & Einarsdottir, 1995	817	6–16	Regional school-based sample	62	M	no
Iran	Minaei, 2005	1,205	6–16	Regional school-based sample	100	S	yes
Israel	Zilber et al., 1994	1,117	6–16	Jerusalem Israeli-born children	80	HI	no
Italy	Frigerio et al., 2004	1,063	6–16	Regional school-based sample	72	M	yes
Jamaica	Lambert et al., 1994, 1998	671	6–16	Regional school-based sample	91	HI	yes
Japan	Itani et al., 2001	4,645	6–16	Regional school-based sample	91	S	no
Korea	Oh et al., 1997	3,081	6–16	National school-based sample	86	S	no
Lithuania	Zukauskiene et al., 2003	2,920	6–16	National school-based sample	85	S	no
Netherlands	Verhulst et al., 1997	1,715	6–16	National household sample	82	HI	no
Norway	Novik, 1999	939	6–16	Regional household sample	45	M	no
Poland	Wolanczyk, 2003	2,479	6–16	National school-based sample	89	S	no
Portugal	Fonseca et al., 1995	1,120	6–11	Regional school-based sample	85	S	no
Puerto Rico	Achenbach et al., 1990	635	6–16	Island-wide household sample	92	HI	no
Romania	Domuta, 2004	990	6–11	Regional school-based sample	80	S	yes
Russia	Hellinckx et al., 2000	1,998	12–16	National household sample	71	HI	yes
Sweden	Larsson & Frisk, 1999	1,354	6–16	Regional school-based sample	84	S	yes
Switzerland	Steinhausen et al., 1997	2,060	6–16	Regional school-based sample	79	M	no
Taiwan	Yang et al., 2000	834	12–16	Regional school-based sample	88	S	no
Thailand	Weisz et al., 1993, 1987	701	6–16	National school-based sample	83	HI	yes
Turkey	Erol & Simsek, 1997	3,790	6–16	National household sample	84	HI	yes
United States	Achenbach & Rescorla, 2001	1,788	6–16	National household sample	93	HI	no

Note. HI = home interview; M = *Child Behavior Checklist* (CBCL; Achenbach, 1991; Achenbach & Rescorla, 2001) was mailed to parents; S = parents completed CBCLs at school or children conveyed CBCLs to parents.

24 societies met our criterion of more than 80 children in each age group (6–8, 9–11, 12–14, 15–16; $n = 47,987$). For ages 6 to 11, Belgium, Greece, Portugal, and Romania could be added, yielding 28 samples ($n = 29,693$). For ages 12 to 16, Ethiopia, Russia, and Taiwan could be added, yielding 27 samples ($n = 25,815$). For the entire sample spanning ages 6 to 16, the sample size was 55,508. Eleven samples excluded children who had been referred for mental health services. Data were collected by home interviews in 11 samples, by mail in 7 samples, and at schools (or with forms conveyed home from schools) in 13 samples. Crijnen et al. (1997) set 80% as a minimum response rate. However, we included samples with response rates as low as 45% to test the effects of response rates on problem scores. Response rates tended to be lowest when forms were mailed to parents.

The data analyzed for this research consisted of parents' reports of their children's functioning. In each of the 31 societies, conventions for obtaining informed consent required by the investigator's research institution were followed. In the United States, each parent gave signed written consent. Data were identified only by participant numbers.

Measure

Foreign language versions of the CBCL, obtained using the translation and back-translation process described above, were used for all societies except Australia, Jamaica, and the United States. Each problem item was rated on a 3-point scale (0 = *not true* [*as far as you know*], 1 = *somewhat or sometimes true*, and 2 = *very true or often true*), based on the preceding 6 months. In 27 of the 31 societies, the 1983/1991 edition of the CBCL was used (Achenbach, 1991; Achenbach & Edelbrock, 1983), whereas Iran, Lithuania, Romania, and the United States used the 2001 edition (Achenbach & Rescorla, 2001). Because Items 2, 4, 5, 28, 78, and 99 were replaced with new problem items when the CBCL was revised in 2001, these 6 items plus the 2 open-ended items were omitted from all analyses. Thus, for each sample, we analyzed 112 items that were common to all societies.

The 2001 versions of the syndrome scales (minus the six new items on those scales) were scored from each data set. Although different factor analytic procedures and different samples were used to derive the CBCL 1991 and CBCL 2001 versions of the eight syndromes, the composition of the 2001 versions of the syndrome scales is similar but not identical to the composition of the 1991 scales. The six *DSM*-oriented scales were also scored from each data set. For each scale, we analyzed raw scores, which consisted of the sum of 0, 1, and 2 ratings obtained by a child for that scale.

Data Analysis

So that data from all 31 societies could be used, we conducted parallel analyses for ages 6 to 16 (24 societies), ages 6 to 11 (28 societies), and ages 12 to 16 (27 societies). With society, gender, and age (6–8, 9–11, 12–14, 15–16) as factors, we used an ANOVA for Total Problems and three separate multivariate analyses of variance (MANOVA) for Internalizing and Externalizing, the eight syndromes, and the six *DSM*-oriented scales. In subsidiary analyses, we conducted within-society Gender × Age ANOVAs for each scale that had significant gender or age effects in the Society × Gender × Age analyses. Effect sizes, measured by eta squared, were characterized using Cohen's (1988) criteria (*small* = .01 to .06, *medium* = .06 to .14, *large* ≥ .14). This is the metric that has been used in most studies reporting effect sizes for the CBCL. We focus on results for ages 6 to 16 in 24 societies to report tests of age effects across the full range from 6 to 16. Findings from the analyses for ages 6 to 11 in 28 societies and 12 to 16 in 27 societies are presented only where they add substantively to the main findings.

For the following analyses, we used the entire sample of 55,508 to (a) test internal consistency reliability (alpha coefficients) for each CBCL scale in each society; (b) obtain means and standard deviations for the 17 scales, which we calculated both weighted and unweighted by sample sizes; (c) compute correlations of the mean item scores from each society with the mean item scores from each other society; and (d) conduct multilevel modeling of the effects of response rate, data collection method, and exclusion of referred children on Total Problems scores. Because of the high statistical power, we set alpha at $p \leq .001$ for ANOVAs, MANOVAs, and bisociety correlations.

RESULTS

Internal Consistency of CBCL Scales in Different Societies

Alpha coefficients for the 17 CBCL scales are summarized in Table 2. Averaged across the analyses for each of the 31 societies, mean alphas for Total Problems, Internalizing, and Externalizing were .93, .83, and .87, respectively. For each of the 31 societies, alphas for Total Problems were ≥ .90, while the alphas for Internalizing and Externalizing were ≥ .72 and ≥ .80, respectively. Mean alphas ranged from .58 to .84 for the eight empirically based syndromes and from .58 to .75 for the six *DSM*-oriented scales. The greater variability of alphas for empirically based syndromes and *DSM*-oriented scales than for Total Problems, Internalizing, and Externalizing could result in part from the smaller number of items in the former scales.

When the 17 alphas obtained for each society were correlated with the 17 alphas for each of the other 30 societies, all 465 bisociety *r*s were significant at $p < .001$. Bisociety *r*s for alphas ranged from .57 (Ethiopia with Iceland) to .99 (Puerto Rico with Portugal, Finland with Switzerland). The mean of the 465 bisociety *r*s for alphas was .88. Ethiopia had the lowest mean bisociety *r* (.77), whereas Poland and Israel had the highest (.92). Thus, the 17 CBCL scales manifested similar internal consistencies across the 31 societies.

TABLE 2

Internal Consistency Alpha Coefficients for CBCL Scales Across 31 Societies

CBCL scale	Minimum	Maximum	M
Total problems	.90	.96	.93
Internalizing	.72	.88	.83
Externalizing	.80	.91	.87
Anxious/depressed	.51	.80	.72
Withdrawn/depressed	.53	.75	.66
Somatic complaints	.51	.77	.65
Social problems	.50	.80	.68
Thought problems	.43	.70	.58
Attention problems	.55	.82	.71
Rule-breaking behavior	.44	.79	.62
Aggressive behavior	.74	.90	.84
DSM Affective problems	.44	.75	.62
DSM Anxiety problems	.39	.68	.58
DSM Somatic problems	.45	.73	.59
DSM Attention-Deficit/Hyperactivity Disorder problems	.54	.79	.70
DSM Oppositional Defiant problems	.54	.79	.70
DSM Conduct problems	.62	.85	.75

Note. CBCL = *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001); *DSM = Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition* (American Psychiatric Association, 1994).

Descriptive Statistics for Problem Scores

Columns (A), (B), and (C) in Table 3 display the lowest society mean, the highest society mean, and the mean for the full sample ($N = 55,508$) for the 17 CBCL scales. The *full sample* means in Column (C) are weighted by within-society sample size, because the larger samples contributed more scores than the smaller samples. Column (D) in Table 3 contains the mean for each scale obtained by averaging the 31 society means, denoted as *averaged* means. These averaged means are unweighted by sample size, because each of the 31 society means contributed equally to the overall mean, regardless of the size of the sample from which it was obtained. As shown in Table 3, the full sample mean of 21.3 for Total Problems score was slightly smaller than the averaged mean of 22.5, because Japan and China, the societies with the largest samples, had low mean Total Problems scores. Most of the other 16 CBCL scales (Internalizing and Externalizing, the eight syndromes, and the six *DSM*-oriented scales) showed this same pattern of the full sample mean being slightly smaller than the averaged mean.

Column (E) contains the average of the 31 standard deviations for each scale, whereas Column (F) contains the standard deviations of the 31 society means for the 17 scales. For most of the 17 scales, the standard deviations in Column (E) are almost three times larger than the standard deviations in Column (F)—for example, 16.9 vs. 5.7 for Total Problems. This pattern of findings indicates much greater within-society variation than between-society variation in scores. Figure 1 displays the mean Total Problems score for each of the 31 societies.

Although Total Problems scores could range from 0 to 224, the 31 means in Figure 1 ranged only from 13.1 for Japan to 34.7 for Puerto Rico. Nineteen societies scored within 1 *SD* (5.7) of the averaged mean of 22.5 (16.8–28.1). These 19 societies included 10 from Europe, 3 from the Middle East, 3 from Asia, plus Jamaica, the United States, and Australia. Six societies scored below this range (Japan, Sweden, Norway, China, Germany, and Iceland), and another six scored above this range (Puerto Rico, Portugal, Ethiopia, Greece, Lithuania, and Hong Kong).

Mean Score Comparisons

Table 4 displays significant effect sizes (η^2) for society, gender, and age. The first ES in each triplet contains the main analyses of the study (i.e., those for ages 6–16 in 24 societies), whereas the second and third effect sizes in each triplet contain findings

TABLE 3
Ranges, Means, and Standard Deviations Across 31 Societies

CBCL scale	(A) Minimum society *M*	(B) Maximum society *M*	(C) Full sample *M*	(D) Averaged *M*	(E) Averaged *SD*	(F) *SD* of 31 societies *M*
Total problems	13.1	34.7	21.3	22.5	16.8	5.7
Internalizing	3.5	12.8	6.2	6.6	5.6	2.1
Externalizing	3.5	10.4	6.2	6.5	6.0	1.5
Anxious/depressed	1.3	5.3	3.0	3.2	2.9	1.0
Withdrawn/depressed	.7	4.0	1.6	1.8	1.9	.7
Somatic complaints	.9	4.4	1.6	1.7	2.0	.8
Social problems	1.2	4.1	2.4	2.5	2.5	.7
Thought problems	.8	2.3	1.4	1.5	1.9	.4
Attention problems	1.4	4.2	2.6	2.7	2.5	.8
Rule-breaking behavior	.8	2.8	1.5	1.6	1.9	.5
Aggressive behavior	2.9	6.6	4.7	5.0	4.6	1.1
DSM Affective problems	1.0	3.7	1.9	1.9	2.1	.7
DSM Anxiety problems	.7	3.1	1.6	1.7	1.7	.6
DSM Somatic problems	.4	2.7	.9	1.5	1.4	.5
DSM Attention-Deficit/ Hyperactivity Disorder problems	1.2	3.3	2.1	2.1	2.0	.6
DSM Oppositional Defiant problems	1.6	2.7	2.2	2.2	1.9	.4
DSM Conduct problems	.6	2.5	1.4	1.5	2.3	.5

Note. N = 55,508. CBCL = *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001); *DSM* = *Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition* (American Psychiatric Association, 1994). Full sample *M* = mean across all participants without regard to society; Averaged *M* = mean of the 31 society means; Averaged *SD* = mean of the 31 society standard deviations; *SD* of 31 societies; *M* = standard deviation for the 31 society means.

for the additional analyses (ages 6–11 in 28 societies and ages 12–16 in 27 societies). Table 5 displays the number of societies with a significant gender effect as well as the direction of the gender effect in the within-society ANOVAs for each scale for each set of analyses. No significant gender effects were found in the opposite direction for any scale within any society.

Total Problems. As shown in Table 4, the effect size for society was .08 for Total Problems for ages 6 to 16. Gender, age, and all interactions were significant, but the effect sizes were all < .01. Overall, boys scored significantly higher than girls. However, within-society gender differences were significant in only 2 of 24 societies (see Table 5). Associations between age and Total Problems showed no consistent pattern.

Internalizing and Externalizing. As shown in Table 4, the society effect size was .08 for Internalizing scores for ages

6 to 16. Gender, age, and all interaction effect sizes related to Internalizing were < .01. Girls scored higher than boys overall and in each society, but the within-society gender difference was significant in only 8 of 24 societies. Mean Internalizing scores tended to increase slightly with age, but the age effect size was significant in only 5 of 24 societies. The effect size for society was .05 for Externalizing for ages 6 to 16, and all other effect sizes related to Externalizing were < .01. Boys scored higher than girls overall and in each society, significantly so in 12 of 24 societies. Externalizing scores tended to decrease with age, significantly so in 8 of 24 societies.

A noteworthy finding from the additional analyses for ages 6 to 11 and ages 12 to 16 is that girls did not score significantly higher than boys on Internalizing in any society at ages 6 to 11, but they obtained significantly higher scores than boys in 10 of 27 societies at ages 12 to 16 (see Table 5). Conversely, boys scored significantly higher than girls on Externalizing in 19 of
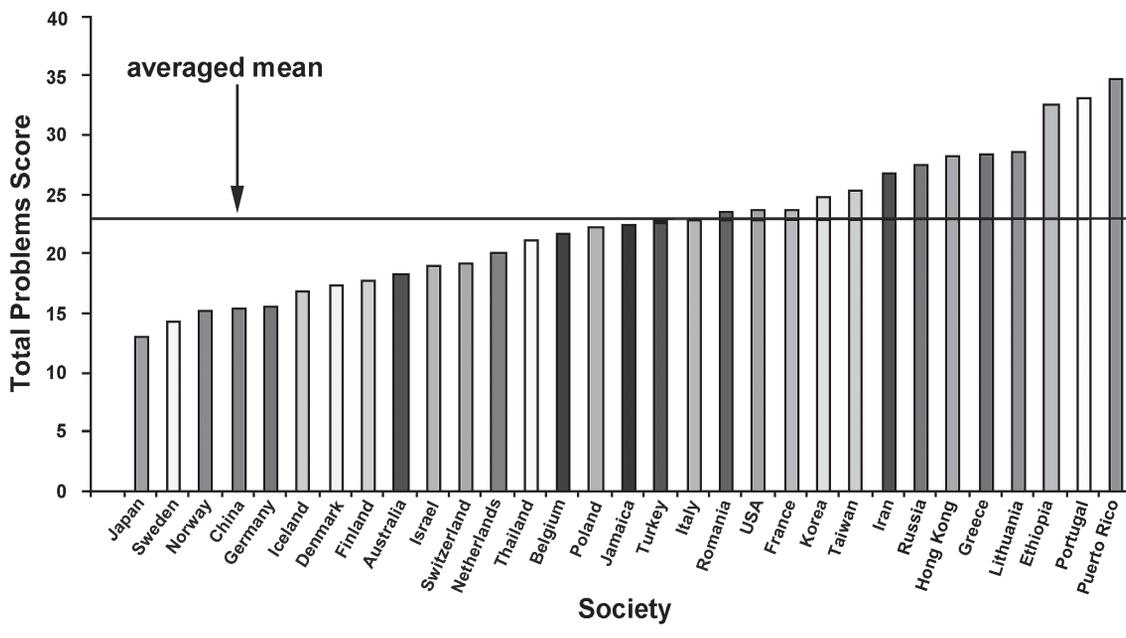
FIGURE 1. Mean Total Problems scores on the *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001) for sample of children (*N* = 55,508) from 31 societies.

28 societies at ages 6 to 11 but in only 5 of 27 societies at ages 12 to 16.

Syndromes. In the 24-society ages 6 to 16 MANOVA, the effect sizes for society ranged from .04 for Rule-Breaking Behavior to .09 for Anxious/Depressed. The six significant gender effect sizes were ≤ .01. As shown in Table 5, girls scored significantly higher than boys on Anxious/Depressed in 9 societies and on Somatic Complaints in 11 societies. Boys scored significantly higher than girls on Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior in 17, 14, and 11 societies, respectively. Age and Age × Society interactions were significant for all syndromes, but all effect sizes were ≤ .01. The most consistent increases with age were for Withdrawn/Depressed and Somatic Complaints, whereas the most consistent decreases with age were for Social Problems.

A noteworthy finding from the additional analyses for ages 6 to 11 and ages 12 to 16 is that the effect size for society on Somatic Complaints was considerably smaller at ages 6 to 11 (ES = .05) than at ages 12 to 16 (ES = .14; see Table 4). Also notable was that girls scored higher than boys on Anxious/Depressed and Somatic Complaints in more societies at ages 12 to 16 than at ages 6 to 11 (see Table 5). Conversely, on Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior, boys scored significantly higher than girls in fewer societies at ages 12 to 16 than at 6 to 11.

DSM-Oriented Scales. In the 24-society ages 6 to 16 MANOVA for *DSM*-oriented scales, effect sizes for society ranged from .03 for Conduct Problems to .08 for Anxiety Problems. Effect sizes for gender were ≤ .01. Girls scored significantly higher than boys on Affective Problems, Anxiety Problems, and Somatic Problems, with the most consistency across societies for Somatic Problems (see Table 5). Boys scored significantly higher than girls on Attention-Deficit/Hyperactivity Problems, Oppositional Defiant Problems, and Conduct Problems, with the most consistency across societies for Attention-Deficit/Hyperactivity Problems and Conduct Problems. Age was significant for all *DSM*-oriented scales except for Affective Problems, with significant Age × Society interactions for all six scales. All ESs were < .01. Attention-Deficit/Hyperactivity Problems generally decreased with age, whereas Somatic Problems generally increased with age.

A noteworthy finding from the additional ages 6 to 11 and ages 12 to 16 analyses was that the effect sizes for society on Somatic Problems were much smaller at ages 6 to 11 (ES = .04) than at ages 12 to 16 (ES = .12; see Table 4). Additionally, girls scored higher than boys in more societies at older than at younger ages on the three *DSM*-oriented scales tapping Internalizing kinds of problems, whereas boys scored higher than girls in more societies at younger than at older ages on Attention-Deficit/Hyperactivity Problems, Oppositional Defiant Problems, and Conduct Problems (see Table 5).

Mean Item Scores

We computed *r*s between the 112 mean item scores obtained in each society and the 112 mean item scores in each other soci-

Significant Effect Sizes (η^2) for Society, Gender, and Age for CBCL Scores for Ages 6 to 16 (24 Societies), Ages 6 to 11 (28 Societies), and Ages 12 to 16 (27 Societies)

	Society			Gender			Age		
CBCL scale	**Ages 6 to 16**	**Ages 6 to 11**	**Ages 12 to 16**	**Ages 6 to 16**	**Ages 6 to 11**	**Ages 12 to 16**	**Ages 6 to 16**	**Ages 6 to 11**	**Ages 12 to 16**
Total problems	.08	.09	.09	< .01	< .01	ns	< .01	ns	ns
Internalizing	.08	.09	.10	< .01	< .01	< .01	< .01	< .01	ns
Externalizing	.05	.06	.05	< .01	.01	< .01	< .01	< .01	ns
Anxious/depressed	.09	.12	.09	< .01	< .01	< .01	< .01	ns	ns
Withdrawn/depressed	.06	.06	.08	ns	ns	ns	< .01	< .01	< .01
Somatic complaints	.05	.08	.14	< .01	< .01	< .01	< .0	< .01	ns
Social problems	.06	.07	.06	ns	ns	ns	.01	< .01	< .01
Thought problems	.05	.04	.06	< .01	< .01	ns	< .01	ns	ns
Attention problems	.06	.07	.06	.01	.01	< .01	< .01	< .01	ns
Rule-breaking behavior	.04	.05	.06	< .01	.01	< .01	< .01	ns	< .01
Aggressive behavior	.05	.05	.05	< .01	< .01	< .01	< .01	< .01	< .01
DSM Affective problems	.06	.06	.08	< .01	ns	< .01	ns	ns	ns
DSM Anxiety problems	.08	.09	.08	< .01	< .01	< .01	< .01	ns	< .01
DSM Somatic problems	.04	.04	.12	< .01	< .01	< .01	< .01	< .01	ns
DSM Attention-Deficit/ Hyperactivity Disorder problems	.05	.07	.05	< .01	.01	< .01	< .01	< .01	< .01
DSM Oppositional Defiant problems	.04	.04	.04	< .01	< .01	< .01	< .01	< .01	< .01
DSM Conduct problems	.03	.04	.04	.01	.02	< .01	< .01	ns	ns

Note. All ages in years. For ages 6 to 16, *n* = 47,987 (24 societies), for ages 6 to 11, *n* = 29,693 (28 societies), and for ages 12 to 16, *n* = 25,815 (27 societies). CBCL = *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001); *DSM* = *Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition* (American Psychiatric Association, 1994).

ety. All bisociety *r*s were significant, with the range being .37 (Ethiopia with Iceland) to .94 (United States with Australia; Germany with Switzerland). Ethiopia had the lowest mean *r* (.47), whereas Belgium had the highest (.80). The mean *r* across all 31 societies was .74 (25th percentile = .71, 75th percentile = .78).

As displayed in Table 6, the 10 items with the highest mean scores when data from all 31 societies were pooled were attention, oppositional, and self-control problems, plus shyness, perfectionism, and jealousy.

The 10 items with the lowest mean scores included gender problems, suicide attempts, seeing things that aren't there, using alcohol or drugs, running away from home, vandalism, stealing, and daytime enuresis.

Effects of Response Rates, Data Collection Methods, and Exclusion of Referred Children

For the complete sample, we used multilevel modeling implemented by Mplus (Muthén & Muthén, 2001, pp. 205–209) to test effects of response rate, data collection method, and exclusion of referred children. For this analysis, children were nested within countries. Response rate, data collection method, and exclusion of referred children served as Level 2 predictors. Low response rates significantly predicted low Total Problems, Internalizing, and Externalizing scores at *p* < .01, with unstandardized parameter estimates of 3.49, 2.68, and 2.87, respectively. Furthermore, when we correlated Total Problems scores with response rates across the 31 societies, the *r* of .38 (*p* < .05)

TABLE 5
Number of Societies With Significant Within-Society Gender Effects by CBCL Scale for
Ages 6 to 16, Ages 6 to 11, and Ages 12 to 16

CBCL scale	Ages 6 to 16 (24 societies)	Ages 6 to 11 (28 societies)	Ages 12 to 16 (27 societies)
Total problems	2 B > G	6 B > G	ns
Internalizing	8 G > B	0	10 G > B
Externalizing	12 B > G	19 B > G	5 B > G
Anxious/depressed	9 G > B	1 G > B	9 G > B
Withdrawn/depressed	ns	ns	ns
Somatic complaints	11 G > B	3 G > B	13 G > B
Social problems	ns	ns	ns
Thought problems	2 B > G	1 B > G	ns
Attention problems	17 B > G	21 B > G	14 B > G
Rule-breaking behavior	14 B > G	17 B > G	10 B > G
Aggressive behavior	11 B > G	17 B > G	3 B > G
DSM Affective problems	4 G > B	ns	4 G > B
DSM Anxiety problems	4 G > B	1 G > B	4 G > B
DSM Somatic problems	10 G > B	4 G > B	10 G > B
DSM Attention-Deficit/Hyperactivity Disorder problems	17 B > G	19 B > G	8 B > G
DSM Oppositional Defiant problems	10 B > G	14 B > G	3 B > G
DSM Conduct problems	18 B > G	22 B > G	13 B > G

Note. All ages in years. For ages 6 to 16, *n* = 47,987 (24 societies), for ages 6 to 11, *n* = 29,693 (28 societies), and for ages 12 to 16, *n* = 25,815 (27 societies). CBCL = *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001); *DSM* = *Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition* (American Psychiatric Association, 1994). B > G indicates that boys obtained higher scores than girls; G > B indicates that girls obtained higher scores than boys; *ns* indicates that the gender effect was not significant in the Society × Gender × Age analysis.

indicated a moderate tendency for higher scores to be associated with higher response rates. Data collection method (–.75, –.64, and –.80) and exclusion of referred children (–.39, –.84, –.50) were not significant (*p* > .10) predictors of problem scores. However, a one-way ANOVA on Total Problems indicated that the 11 societies using home interviews tended to have the highest scores, the 7 using mail surveys had the lowest, and the 13 using school-based methods were in between (*p* < .001).

DISCUSSION

Results of the study provided strong support for the multicultural robustness of the CBCL across 31 diverse societies in terms of internal consistency reliability, mean scale scores across societies, gender and age effects, and mean item scores. Ivanova et al.'s (in press) CFAs of the same samples demonstrated the multicultural robustness of the syndromes scored from the

CBCL by showing that the U.S.-based eight-syndrome structure fits CBCL data from the other 30 societies.

Internal Consistency of CBCL Scales

Across the 31 societies, the mean bisociety *r* of .88 (range = .77–.92) indicated a high degree of similarity in patterns of internal consistency across CBCL scales. In all societies, mean alphas for Total Problems, Internalizing, and Externalizing were high. The 31 societies were similar with respect to which scales had the highest alphas (Aggressive Behavior and *DSM*-oriented Conduct Problems) and the lowest alphas (Thought Problems, *DSM*-oriented Anxiety Problems, *DSM*-oriented Somatic Problems). The consistency of the alphas across these diverse societies complements Ivanova et al.'s (in press) CFA findings of good fit for the eight-syndrome model in the same societies analyzed for the present article.

TABLE 6
CBCL Items With the 10 Highest and 10 Lowest Scores Across 31 Societies

Highest mean score items	Lowest mean score items
3. Argues a lot	60. Plays with sex parts too much
19. Demands a lot of attention	106. Vandalism
10. Can't sit still, restless, or hyperactive	70. Sees things that aren't there
8. Can't concentrate, can't pay attention for long	67. Runs away from home
32. Feels he/she has to be perfect	105. Uses alcohol or drugs for nonmedical purposes
86. Stubborn, sullen, or irritable	59. Plays with sex parts in public
27. Easily jealous	73. Sexual problems
22. Disobedient at home	18. Deliberately harms self or attempts suicide
63. Prefers being with older kids	82. Steals outside the home
77. Too shy or timid	107. Wets self during the day

Note. Items are listed in descending order of mean scores. CBCL = *Child Behavior Checklist* (Achenbach, 1991; Achenbach & Rescorla, 2001).

Mean Scale Scores

Our results indicated considerable consistency across 31 societies in CBCL mean scores. Nineteen societies that differ widely in geographical region, political and economic systems, size, population, ethnicity/race, and predominant religion scored within 1 *SD* of the overall mean of 22.5 on the Total Problems scale. The 19 societies with similar scores included several that are typically considered to be highly *individualistic* in orientation (e.g., United States, Australia, the Netherlands; Triandis, 1989), societies typically considered to be more *collectivistic* in orientation (e.g., Thailand and Korea), and societies that previously had collectivistic regimes for decades (e.g., Poland and Romania). These cultural factors did not seem to have discernible effects on mean Total Problems scores.

Six societies (Japan, China, Sweden, Norway, Germany, and Iceland) scored >1 *SD* below the overall mean on Total Problems, whereas six other societies (Puerto Rico, Portugal, Ethiopia, Greece, Lithuania, and Hong Kong) scored >1 *SD* above the overall mean. Neither the six low-scoring societies nor the six high-scoring societies have any obvious common denominator, as they differ in geographic region, ethnicity, religion, size, population, and economic/political system.

One reason for particularly low scores in some societies may be parents' reluctance to report problems. For instance, people in Asian societies are more concerned with self-presentation ("saving face") than people in Euro-American societies (Yabuuchi, 2004). This tendency might account for the particularly low scores found in China and Japan, but it does not explain why scores in Taiwan, Hong Kong, Thailand, and Korea were not particularly low. Additionally, although our data suggested that relatively low response rates may have contributed to low problem scores in Norway and Iceland, they cannot explain the low scores in Japan and China (response rates > 90%). Conversely, high scores may reflect relatively low thresholds for reporting problems. Perhaps parents in Hong Kong, Greece, Lithuania, Ethiopia, Portugal, and Puerto Rico have high expectations for how children should behave. If children do not conform to these expectations, parents regard them as displaying problems.

Effect sizes for society ranged from .03 to .14, with 47 of the 51 ESs < .10. Societies differed more on scores for Internalizing kinds of problems than on scores for Externalizing kinds of problems. This suggests more variation among societies in the expression, perception, and reporting of anxiety, depression, and somatic complaints than of aggressive and rule-breaking behavior.

Although most of the effect sizes for society were < .10, they were consistently larger than effect sizes for gender or age. There were thus some important variations across societies in mean scores on CBCL scales. However, no single factor that we tested statistically (such as response rate) or inspected visually (such as geographic region, ethnicity, religion, or economic/political system) appears to account for particularly low scores or particularly high scores. Additional research using measures of cultural values as well as the CBCL is clearly needed to ad-

dress these issues. What is clear from our results, however, is that multicultural norms for the CBCL need to include low, medium, and high scoring brackets.

Age and Gender Differences

Effect sizes for age differences never exceeded .01 and were rather inconsistent across societies. The most consistent age trends across societies were increases with age for Withdrawn/Depressed and decreases with age for Social Problems and *DSM*-oriented Attention-Deficit/Hyperactivity Problems. These age effects suggest that adolescents in many societies may become more withdrawn from their parents as they become increasingly engaged with their peers. Additionally, the findings indicate that parents in many societies observe increases with age in depressive feelings and decreases with age in inattentive, hyperactive, and impulsive behavior.

Significant gender effects were found on all 17 CBCL scales except Withdrawn/Depressed and Social Problems. Although 41 of the 42 significant effect sizes for gender were ≤ .01 and the gender differences found in the overall analysis were not always significant at $p ≤ .001$ in every society, no significant within-society gender differences in the opposite direction to the general trend were found for any scale within any society. Girls in most societies tended to score higher on Internalizing kinds of problems, especially at ages 12 to 16. By contrast, boys in most societies tended to score higher on Externalizing kinds of problems, especially at ages 6 to 11. The cross-society consistency in gender differences across age in Internalizing versus Externalizing problems is one of the major findings of this study. Although age and gender effects were generally small, they manifested sufficient consistency across societies to support the long-standing practice of norming the CBCL by age group and gender.

Mean Item Scores

Although it cannot be known whether CBCL items have exactly the same meaning for parents in all societies, the mean bisociety *r* of .74 indicated that parents' ratings in these 31 diverse societies were quite consistent with respect to which CBCL items tended to receive high, medium, and low scores. Furthermore, items with high versus low scores seemed congruent with commonsense notions about children's problems. For example, items with the highest mean scores included oppositional behavior, shyness, and difficulties with attention and self-control, whereas items with the lowest mean scores included seeing things that aren't there, running away from home, and daytime enuresis.

Response Rates, Data Collection Methods, and Exclusion of Referred Children

Our multilevel analysis supported the hypothesis that higher response rates are somewhat associated with higher problem scores, as did the *r* of .38 between Total Problems and response rate. Additionally, problem scores tended to be lower when data were obtained by mail and higher when obtained by home interviews. Exclusion of referred children was not associated with lower problem scores, perhaps because referred children were not numerous enough for their inclusion versus exclusion to have much impact on overall scores.

Limitations

The present study used an *etic* approach in that the same standardized assessment instrument was used to obtain parents' reports in 31 societies. A possible limitation of the study was that it did not use an *emic* approach, whereby the meanings of the CBCL items to parents would be explored within each society. Additionally, because our items do not tap all possible problems that parents might report for all children in all societies, use of additional problem items might yield different findings. Our study does, however, provide systematic comparisons of scores on 112 problem items and 17 scales for 55,508 children in 31 societies that differ widely in ethnicity, geographical region, religion, economic/political systems, and historical experience.

Another limitation of our study was that it included only parents' reports. Parallel studies of teacher reports in 21 societies (Rescorla et al., 2007) and self-reports in 24 societies (Rescorla et al., 2007) yielded results consistent in many ways with those presented here for parents' reports. For both teachers' reports and self-reports, ESs for society were significant but of small to moderate size, most societies scored within 1 *SD* of the overall mean, and bisociety *r*s for mean item scores were high. However, gender differences varied in important ways across the three studies. For example, girls did not obtain consistently higher scores than boys on any scales in teachers' reports, contrary to the pattern found for parents' reports and self-reports. In addition, girls scored as high as boys on the Attention Problems syndrome and the *DSM*-oriented Attention-Deficit/Hyperactivity Problems scale according to self-reports, whereas boys obtained higher scores than girls on these scales according to parents' and teachers' reports.

Conclusions and Implications

To help children from diverse backgrounds, assessment instruments need to be economical, easily administered, readily interpreted by different kinds of professionals, and multiculturally robust. Establishing multicultural robustness involves demonstrating that an instrument yields comparable data in many kinds of analyses across multiple societies. Taken together, the present study and Ivanova et al.'s (in press) CFA study provide new and converging evidence for the multicultural robustness of the CBCL. Our comparison of parent-reported problems in 31 societies that differ markedly in geographical region, political/economic systems, size, population, history, ethnic and racial composition, and religion yielded many striking similarities across societies. On balance, our study demonstrates that

within-society variation in parent-reported problem scores greatly exceeded between-society variation.

About the Authors

LESLIE RESCORLA, PhD, is a professor of psychology and the director of the Child Study Institute at Bryn Mawr College in Philadelphia, Pennsylvania. Her research interests focus on language delay in young children; empirically based assessment of emotional and behavioral problems in children, adolescents, and adults; and longitudinal patterns of school achievement; **THOMAS ACHENBACH, MASHA Y. IVANOVA,** and **LEVENT DUMENCI,** University of Vermont, Burlington; **FREDRIK ALMQVIST** and **SHEILA WEINTRAUB,** Helsinki University, Helsinki, Finland; **NIELS BILENBERG,** University of Southern Denmark, Odense, Denmark; **HECTOR BIRD,** Columbia University, New York City; **WEI CHEN,** National Taiwan University, Taipei, Taiwan; **ANCA DOBREAN,** Babes-Bolyai University, Cluj, Romania; **MANFRED DÖPFNER,** Köln University, Cologne, Germany; **NESE EROL,** Ankara University, Ankara, Turkey; **ERIC FOMBONNE,** McGill University, Montreal, Canada; **ANTONIO FONSECA,** University of Coimbra, Coimbra, Portugal; **ALESSANDRA FRIGERIO,** Istituto Scientifico Eugenio Medea, Lecco, Italy; **HANS GRIETENS,** University of Leuven, Leuven, Belgium; **HELGA HANNESDOTTIR,** University Hospital, Reykjavik, Iceland; **YASUKO KANBAYASHI,** Chuo University, Tokyo, Japan; **MICHAEL LAMBERT,** University of Missouri, Columbia; **BO LARSSON,** Norwegian University of Science and Technology, Trondheim, Norway; **PATRICK LEUNG,** Chinese University of Hong Kong; **XIANCHEN LIU,** University of Pittsburgh, Pittsburgh, Pennsylvania; **ASGHAR MINAEI,** Research Institute of Exceptional Children, Tehran, Iran; **MESFIN S. MULATU,** The MayaTech Corporation, Silver Spring, Maryland; **TORUNN S. NOVIK,** Buskerud Hospital, Drammen, Norway; **KYUNG-JA OH,** Yonsei University, Seoul, South Korea; **ALEXANDRA ROUSSOS,** Attiki Child Psychiatric Hospital, Athens, Greece; **MICHAEL SAWYER,** University of Adelaide, Adelaide, Australia; **ZEYNEP SIMSEK,** University of Harran, Harran, Turkey; **HANS-CHRISTOPH STEINHAUSEN** and **CHRISTA WINKLER METZKE,** University of Zurich, Zurich, Switzerland; **JOHN WEISZ,** Harvard University, Cambridge, Massachusetts; **TOMASZ WOLANCZYK,** Medical University of Warsaw, Poland; **HAO-JAN YANG,** Chung Shan Medical University, Taichung, Taiwan; **NELLY ZILBER,** Falk Institute for Mental Health Studies and the French Research Center, Jerusalem, Israel; **RITA ZUKAUSKIENE,** Mykolas Romeris University, Vilnius, Lithuania; **FRANK VERHULST,** Erasmus Medical Center–Sophia Children's Hospital, Rotterdam, The Netherlands. Address: Leslie Rescorla, Bryn Mawr College, 101 N. Merion Avenue, Bryn Mawr, PA 19010; e-mail: lrescorl@brynmawr.edu

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/ 4–18 and 1991 profile*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., Bird, H. R., Canino, G. J., Phares, V., Gould, M., & Rubio-Stipec, M. (1990). Epidemiological comparisons of Puerto Rican and U.S. mainland children: Parent, teacher, and self reports. *Journal of the American Academy of Child and Adolescent Psychiatry, 29,* 84–93.

Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2003). *DSM*-oriented and empirically based approaches to constructing scales from the same item pools. *Journal of Clinical Child and Adolescent Psychology, 32,* 328–340.

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior profile*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

Achenbach, T. M., & Rescorla, L. R. (2007). *Multicultural understanding of child and adolescent psychopathology: Implications for mental health assessment*. New York: Guilford Press.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bilenberg, N. (1999). The Child Behavior Checklist (CBCL) and related material: Standardization and validation in Danish population and clinically based samples. *Acta Psychiatrica Scandivica, Supplementum, 100,* 398, 1–52.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1997). Comparisons of problems reported by parents of children in 12 cultures: Total problems, externalizing, and internalizing. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 1269–1277.

Crijnen, A. A. M., Achenbach, T. M., & Verhulst, F. C. (1999). Comparisons of problems reported by parents of children in twelve cultures: The CBCL/4-18 syndrome constructs. *The American Journal of Psychiatry, 156,* 569–574.

Domuta, A. (2004). *Evaluarea sindromului ADHD la varsta prescolara. Teza de doctorat*. Cluj Napoca, Romania: Universitatea Baves-Bolyai.

Döpfner, M., Plück, J., Berner, W., Fegert, J. M., Hüss, M., Schmeck, K., et al. (1997). Psychishe auffälligkeiten von kindern und jugendlichen in Deutschland. *Zeitschrift fur Kinder- und Jugendpsychiatrie und Psychotherapie, 25,* 218–233.

Erol, N., & Simsek, Z. (1997). Mental health profiles of Turkey: Behavioral and emotional problems and competencies in Turkish children (in Turkish). In N. Erol, C. Kilic, M. Ulusoy, M. Kececi, & Z. Simsek (Eds.), *Mental health profiles in Turkey: A preliminary report* (pp. 12–33). Ankara, Turkey: Aydogdu.

Fombonne, E., & Vermeersch, S. (1997). *Les enfants de la Cohorte GAZEL: I–Prévalence des contacts avec le système médico-éducatif pour des motifs psychologiques, et facteurs associés* [Children of the GAZEL Cohort: Prevalence of contacts with medical-educational systems for psychological causes, and associated factors]. *Revue Epidémiologique et Santé Publique, 45,* 29–40.

Fonseca, A. C., Simoes, A., Rebelo, J. A., Ferreira, J. A., Cardoso, F., & Temudo, P. (1995). Hyperactivity and conduct disorder among Portuguese children and adolescents: Data from parents' and teachers' reports. In J. Sergeant (Ed.), *Eunethydis: European approaches to hyperkinetic disorder* (pp. 115–129). Zürich: Fotorotar.

Frigerio, A., Cattaneo, C., Cataldo, M. G., Schiatti, A., Molteni, M., & Battaglia, M. (2004). Behavioral and emotional problems among Italian children and adolescents aged 4 to 18 years as reported by parents and teachers. *European Journal of Psychological Assessment, 20,* 124–133.

Geisinger, K. (1994). Cross-cultural normative assessment: Transla-

tion and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6,* 304–312.

Hannesdottir, H., & Einarsdottir, S. (1995). Icelandic Child Mental Health Study. An epidemiological study of Icelandic children 2–18 years of age using the Child Behavior Checklist as a screening instrument. *European Child & Adolescent Psychiatry, 4,* 237–248.

Hellinckx, W., Grietens, H., & De Munter, A. (2000). Parent-reported problem behavior in 12–16-year-old American and Russian children: A cross-national comparison. In N. N. Singh, J. P. Leung, & A. N. Singh (Eds.), *International perspectives on child and adolescent mental health* (pp. 205–222). Oxford: Elsevier.

Hellinckx, W., Grietens, H., & Verhulst, F. (1994). Competence and behavioral problems in 6- to 12-year-old children in Flanders (Belgium) and Holland: A cross-national comparison. *Journal of Emotional and Behavioral Disorders, 2,* 130–142.

Itani, T., Kanbayashi, Y., Nakata, Y., Kita, M., Kuramoto, H., Negishi, T., et al. (2001). Standardization of the Japanese version of the Child Behavior Checklist/4-18. *Psychiatrica et Neurologia Paediatrica Japonica, 41,* 243–252.

Ivanova, M. Y., Achenbach, T. M., Dumenci, L., Rescorla, L. A., Almqvist, F., Bilenberg, N., et al. (in press). Testing the 8-syndrome structure of the CBCL in 30 societies. *Journal of Consulting and Clinical Psychology.*

Lambert, M. C., Knight, F., Taylor, R., & Achenbach, T. M. (1994). Epidemiology of behavioral and emotional problems among children of Jamaica and the United States: Parent reports for ages 6–11. *Journal of Abnormal Child Psychology, 22,* 113–128.

Lambert, M. C., Lyubansky, M., & Achenbach, T. M. (1998). Behavioral and emotional problems among adolescents of Jamaica and the United States: Parent, teacher, and self-reports for ages 12 to 18. *Journal of Emotional and Behavioral Disorders, 6,* 180–187.

Larsson, B., & Frisk, M. (1999). Social competence and emotional/behaviour problems in 6–16-year-old Swedish school children. *European Child & Adolescent Psychiatry, 8,* 24–33.

Leung, P., Kwong, S. L., Tang, C. P., Ho, T. P., Hung, S. F., Lee, C. C., et al. (2006). Test-retest reliability and criterion validity of the Chinese version of CBCL, TRF, and YSR. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47,* 970–973.

Liu, X., Guo, C., Okawa, M., Zhai, J., Li, Y., Uchiyama, M., et al. (2000). Behavioral and emotional problems among Chinese children of divorce. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 896–903.

Liu, X., Kurita, H., Guo, C., Miyake, Y., Ze, J., & Cao, H. (1999). Prevalence and risk factors of behavioral and emotional problems among Chinese children aged 6 through 11 years. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 708–715.

Minaei, A. (2005). *Manual of ASEBA school-age forms for Iranian children.* Tehran: Research Institute for Exceptional Children.

Mulatu, M. S. (1997). *Stress, coping, and adaptation of Ethiopian adolescents: Testing the applicability of the transactional model of stress and coping.* Unpublished doctoral dissertation, Queen's University, Ontario, Canada.

Muthén, L. K., & Muthén, B. O. (2001). *Mplus: User's guide.* Los Angeles: Muthén & Muthén.

Novik, T. S. (1999). Validity of the child behavior checklist in a Norwegian sample. *European Child & Adolescent Psychiatry, 8,* 247–254.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Oh, K.-J., Lee, H., Hong, K. E., & Ha, E. H. (1997). *Manual for the Korean Child Behavior Checklist.* Seoul: Choongang Juksung Press.

Rescorla, L. A., Achenbach, T. M., Ginzburg, S., Ivanova, M. Y., Dumenci, L., Almqvist, F., et al. (2007). Consistency of teacher-reported problems for students in 21 countries. *School Psychology Review, 36,* 91–110.

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., et al. (2007). Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology, 75,* 351–358**.**

Roussos, A., Karantanos, G., Richardson, C., Hartman, C., Karajiannis, D., Kyprianos, S., et al. (1999). Achenbach's Child Behavior Checklist and Teacher's Report Form in a normative sample of Greek children 6–12 years old. *European Child & Adolescent Psychiatry, 8,* 165–172.

Sawyer, M. G., Arney, F. M., Baghurst, P. A., Clark, J. J., Graetz, R. J., Kosky, R. J., et al. (2001). The mental health of young people in Australia: Key findings from the child and adolescent component of the National Survey of Mental Health and Well-being. *The Australian and New Zealand Journal of Psychiatry, 35,* 806–814.

Steinhausen, H. C., Metzke, C. W., Meier, M., & Kannenberg, R. (1997). Behavioral and emotional problems reported by parents for ages 6 to 17 in a Swiss epidemiological study. *European Child & Adolescent Psychiatry, 6,* 136–141.

Triandis, H. C. (1989). The self and social behavior in different cultural contexts. *Psychological Review, 96,* 506–520.

Verhulst, F. C., van der Ende, J., Ferdinand, R. F., & Kasius, M. C. (1997). The prevalence of *DSM-III-R* diagnoses in a national sample of Dutch adolescents. *Archives of General Psychiatry, 54,* 329–336.

Weintraub, S. (2004). *The Finnish standardization of the ASEBA scales.* Unpublished manuscript, University of Helsinki, Finland.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Achenbach, T. M., & Eastman, K. L. (1993). Behavioral and emotional problems among Thai and American adolescents: Parent reports for ages 12–16. *Journal of Abnormal Psychology, 102,* 395–403.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., Weiss, B., Achenbach, T. M., & Walter, B. R. (1987). Epidemiology of behavioral and emotional problems among Thai and American children: Parent reports for ages 6–11. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 890–897.

Wolanczyk, T. (2003). *Zaburzenia emocjonalne i behawioralne u dzieci i modziezy w Polsce* [Behavioral and emotional problems of children and adolescents in Poland]. Warsaw, Poland: Akademia Medyczna w Warszawie.

Yabuuchi, A. (2004). Face in Chinese, Japanese, and U.S. American cultures. *Journal of Asian Pacific Communication, 14,* 261–297.

Yang, H. J., Soong, W. T., Chinag, C. N., & Chen, W. J. (2000). Competence and behavioral/emotional problems among Taiwanese adolescents as reported by parents and teachers. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 232–239.

Zilber, N., Auerbach, J., & Lerner, Y. (1994). Israeli norms for the Achenbach Child Behavior Checklist: Comparison of clinically-referred and non-referred children. *The Israel Journal of Psychiatry and Related Sciences, 31,* 5–12.

Zukauskiene, R., Ignataviciene, K., & Daukantaite, D. (2003). Subscale scores of the Lithuanian version of the CBCL. *European Child & Adolescent Psychiatry, 12,* 136–143.