European Child & Adolescent Psychiatry (2019) 28:1107–1115
https://doi.org/10.1007/s00787-018-01268-3

ORIGINAL CONTRIBUTION

Effects of society and culture on parents' ratings of children's mental health problems in 45 societies

Leslie A. Rescorla[1] · Robert R. Althoff[2] · Masha Y. Ivanova[2] · Thomas M. Achenbach[2]

Received: 18 August 2018 / Accepted: 20 December 2018 / Published online: 18 January 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

To improve international needs assessment for child mental health services, it is necessary to employ standardized assessment methods that can be easily administered and scored, can be interpreted by practitioners and researchers with various kinds of training, and that perform similarly across many societies. To this end, we tested the effects of both society and culture on parents' ratings of children's problems. We used hierarchical linear modeling as well as analyses of variance to analyze parents' Child Behavior Checklist ratings of 72,493 6- to 16-year-olds from 45 societies. The 45 societies were nested within 10 culture clusters based on the Global Leadership and Organizational Behavior Effectiveness (GLOBE) taxonomy. Societal differences accounted for 3.8–10.7% of variance in various kinds of problems, while differences between culture clusters (e.g., Anglo vs. Confucian) accounted for 0.1–10.0%. By contrast, differences associated with parents' ratings of individual children accounted for 85.5–93.3% of variance. Averaged across 17 problem scales, society plus culture cluster accounted for about 10% of the variance in parents' ratings of children's problems, whereas individual differences and other possible variables accounted for about 90%. These findings indicate that parents' standardized ratings can be used to assess effects associated with individual differences in child and adolescent psychopathology, over and above differences associated with societies and culture clusters.

Keywords International epidemiology · Individual differences in child psychopathology · Cross-cultural · CBCL

Introduction

Needs assessment for child mental health services is a major public health goal around the world. Because mental health research has developed primarily in a few rather similar societies, extension of such research to additional societies requires standardized assessment methods that are applicable across many societies. (We define society as a geopolitically demarcated population having a dominant language). To be useful for improving public health around the world, assessment instruments need to be easily administered and scored. Practitioners and researchers from diverse backgrounds should also be able to interpret the results. To extend standardized assessment of child mental health internationally, assessment instruments need to be tested across many societies [2, 3].

To test societal similarities and differences in results obtained from standardized child mental health assessment instruments, Rescorla et al. [4] compared parents' 0–1–2 ratings of behavioral and emotional problems on the Child Behavior Checklist/6–18 (CBCL/6–18) [5] for population samples from 31 societies ($N = 55{,}508$). When Q correlations were computed between mean item ratings for each pair of societies and the bi-society correlations were then averaged, the mean correlation was 0.74, indicating that parents' ratings in these 31 societies were quite consistent with respect to which CBCL/6–18 items tended to receive high, medium, or low ratings. Analyses of variance (ANOVAs) yielded effect sizes (ESs) for society ranging from 3 to 14% of the variance in scores across the syndromes, DSM-oriented scales, and broad-spectrum scales (Internalizing, Externalizing, and Total Problems) scored from the problem items. ESs for gender were ≤ 1%, with girls generally

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00787-018-01268-3) contains supplementary material, which is available to authorized users.

✉ Leslie A. Rescorla
lrescorl@brynmawr.edu

[1] Bryn Mawr College, Bryn Mawr, USA

[2] University Vermont, Burlington, USA

scoring higher on Internalizing problems and boys generally scoring higher on Externalizing problems. ESs for age were also ≤ 1% and varied across types of problems. Total Problems scores for 19 of 31 societies were within one standard deviation (SD=5.7) of the overall mean of 22.5. 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 seemed to have any obvious common denominator, as they differed in geographic region, ethnicity, religion, population, and economic/political system.

In 2012, Rescorla et al. [3] added 11 new societies to the 31 previously studied [4] to compare parents' ratings on the CBCL/6–18 for population samples from 42 societies. ANOVAs yielded societal differences accounting for 3–10% of the variance in the scores across the various scales, plus small age and gender effects similar to those reported by Rescorla et al. [4]. The average of the mean scores from the 42 societies was computed to obtain the omnicultural mean for each scale. The SD of the 42 societal mean scores around the omnicultural mean was much smaller than the SD of scores around the mean for each society for every CBCL/6–18 scale [3]. In other words, less of the variance in problem scores was attributable to differences between societies than to individual differences within societies. Seven of the 42 societies (Japan, Sweden, Norway, China, Germany, Serbia, and Iceland) scored > 1 SD (6.7) below the overall mean of 24.4 on Total Problems, whereas nine other societies (Brazil, Colombia, Puerto Rico, Algeria, Portugal, Ethiopia, Pakistan, Uruguay, and Tunisia) scored > 1 SD above the overall mean. Neither the low-scoring societies nor the high-scoring societies seemed to have any obvious common denominator.

The 42 societies compared by Rescorla et al. [3] are located in many regions of the world, including Africa, Asia, North and South America, Europe, and the Middle East, but the authors did not analyze their data by these broad geographical regions. However, if geographical regions such as Europe were subdivided (e.g., into Western, Eastern, and Southern Europe), cultural effects might be detectible that were not apparent across societies.

Culture clusters

Much research on cultural differences has focused on people in Western versus other societies. A longstanding hypothesis is that people in Western societies tend to be individualistic, whereas people in Eastern societies tend to be collectivistic [6]. However, within the broad category of Western (or Eastern) cultures, societies differ widely in political and economic systems, historical experience, population, ethnicity/race, and religion (e.g., Finland vs. Italy, Japan vs. Vietnam). Some Western (and Eastern) societies may, therefore, be more individualistic (or collectivistic) than others.

Furthermore, regional populations may differ within societies (e.g., Prussians vs. Bavarians, Northern vs. Southern Italians). For example, Talhelm et al. [7] hypothesized that within prototypically collectivist China, people from wheat farming regions would be less collectivistic than people from rice farming regions because wheat farming involves independence among farmers whereas rice farming involves interdependence. In support of their hypothesis, Talhelm et al. found that Chinese students from wheat regions obtained higher test scores for individualism and lower scores for holistic thinking than students from rice regions. These findings suggest that important differences may exist within societies commonly viewed as being culturally homogeneous, but testing within-society cultural differences was beyond the scope of our study.

Given the numerous ways in which cultural groups can be defined, both within and across societies, it is useful to consider cultural groupings that have been widely studied in previous research when testing effects of culture on scales for assessing mental health problems. A well-researched set of cultural groupings, termed culture clusters, was developed by the Global Leadership and Organizational Behavior Effectiveness study (GLOBE) [1]. Building on previous research and theory [8, 9], over 200 scholars from 69 societies formulated nine dimensions: Performance Orientation, Assertiveness, Future Orientation, Humane Orientation, Institutional Collectivism, In-Group Collectivism, Gender Egalitarianism, Power Distance, and Uncertainty Avoidance. Using data from more than 17,000 participants in 62 societies, GLOBE researchers [1] used these nine dimensions to classify societies into 10 culture clusters. For example, China, Hong Kong, Japan, Singapore, South Korea, and Taiwan were grouped into the Confucian Asia cluster, based on culturally similar practices and values.

Stankov [10] used the GLOBE taxonomy to analyze scores from English language personality tests completed by college students in 45 societies. Using hierarchical linear modeling (HLM), Stankov estimated the percentage of variance in personality test scores attributable to society, culture cluster, and individual differences. For Neuroticism, the personality trait most closely associated with psychopathology, 2.0% of the variance in students' scores was attributable to society, 2.7% to culture cluster, and 95.3% to individual differences within societies. However, all participants were college students and they all completed the tests in English, regardless of their native language. It is, therefore, possible that the effects of society and culture cluster might be greater with more representative samples of people assessed in their native languages.

Purpose of this study

Rescorla et al. [3] found that society accounted for significant but modest effects on parents' ratings of children's mental health problems, with even smaller effects of gender and age. However, they did not test effects of cultural groupings on CBCL/6–18 scores. To address this gap, we tested the hypothesis that culture cluster would significantly predict variance in children's CBCL/6–18 scale scores, over and above effects of society. Based on Stankov's findings [10], we predicted that society and culture cluster would account for relatively small portions of the total variance, with individual differences in children's mental health problems as well as untested other factors accounting for the rest. We used HLM to take account of the nesting of societies within culture clusters. To further examine the effects of culture cluster and society on parent-reported mental health problems, we tested society and gender as predictors in one ANOVA and culture cluster and gender in a second ANOVA, using the same samples from 45 societies.

Method

Participants

Parents' ratings of population samples totaling 72,493 6- to 16-year-olds were obtained by indigenous researchers in 45 societies. Many of the societies were countries, but some (e.g., Hong Kong, Flanders, Puerto Rico) were not. We added three societies to the 42 analyzed by Rescorla et al. [3], namely Indonesia, Kenya, and Vietnam. Using the GLOBE [1] culture cluster taxonomy, we classified our 45 societies into the following 10 culture clusters: Anglo ($N=5537$, 3 societies); Confucian Asia ($N=18{,}353$, 7 societies); Eastern Europe ($N=13{,}693$, 8 societies); Germanic Europe ($N=6912$, 4 societies); Latin America ($N=5433$, 5 societies); Latin Europe ($N=5433$, 4 societies); Middle East ($N=5361$, 3 societies); Nordic Europe ($N=5831$, 5 societies); Southern Asia ($N=2814$, 4 societies); and Sub-Saharan Africa ($N=1535$, 2 societies) (see Table 1). Cases that were missing ratings for >8 problem items were excluded from the analyses. The mean of the completion rates shown in Table 1 was 83%. Indigenous researchers followed their institutions' procedures for protection of human participants.

Measure

The CBCL/6–18 has 118 items describing behavioral, emotional, social, and thought problems that parents rate on a Likert scale where 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 [5]. Because the pre-2001 edition of the CBCL was used for the samples assessed prior to 2000, the six items that were replaced on the 2001 edition (items *2*, 4, 5, 28, 78, and 99) were omitted from all analyses, consistent with Rescorla et al. [3]. The 2001 version of each scale was scored by summing the 0–1–2 ratings of the items comprising the scale (minus the six excluded items). The English language CBCL/6–18 was used in Australia, Jamaica, and the US, while translations were used in the other 42 societies. Indigenous interviewers administered the CBCL/6–18 aloud or gave parents a paper copy to complete.

We analyzed scores for all eight empirically based syndromes: Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, and Aggressive Behavior. We also analyzed three broad-spectrum scales: Total Problems (sum of 0–1–2 ratings for 112 problem items); Internalizing (sum of 0–1–2 ratings for items comprising the Anxious/Depressed, Withdrawn/Depressed, and Somatic Complaints syndromes); and Externalizing (sum of 0–1–2 ratings for items comprising the Rule-Breaking Behavior and Aggressive Behavior syndromes). The Internalizing and Externalizing scales comprise syndromes found to co-occur in second-order factor analyses of the eight syndromes [5]. Lastly, we analyzed the six DSM-oriented scales (Depressive Problems, Anxiety Problems, Somatic Problems, Attention Deficit Hyperactivity Problems, Oppositional Defiant Problems, and Conduct Problems) comprising items that an international panel of expert clinicians judged to be consistent with diagnostic categories of the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (DSM-5) [11, 12].

All CBCL scales analyzed for the present study have demonstrated high 8-day test–retest reliability correlations (ranging from 0.80 for DSM-Anxiety Problems to 0.94 for Total Problems) and high internal consistency reliabilities (Cronbach's α coefficients ranging from 0.78 for Somatic Complaints and Thought Problems to 0.97 for Total Problems) [5]. All scales except Somatic Complaints and DSM-Somatic Problems have also been found to discriminate well between demographically matched samples of children referred for mental health services versus non-referred children (significant effects for referral status ranged from 20% of variance for Anxious/Depressed to 39% for DSM-Conduct Problems) in multiple regressions that controlled for demographic effects [5].

Data analyses

Each scale was analyzed via HLM 7 [13]. Prior to the HLM analyses, we transformed all scale scores with the Box-Cox

Table 1 CBCL/6–18 samples from 45 societies grouped by culture cluster

Society	Reference[a]	*N*	Age range	Sampling frame	Method[b]	% boys	Comp. rate (%)
Anglo (*N* = 5537)							
Australia	Sawyer et al. (2001)	3078	6–16	National household	HI	50	86
Jamaica	Lambert et al. (1994, 1998)	671	6–16	Regional school-based	HI	50	91
U.S.	Achenbach and Rescorla (2001)	1788	6–16	National household	HI	53	93
Confucian Asia (*N* = 18,353)							
China	Liu et al. (1999, 2000)	4858	6–15	Regional household	S	51	92
Hong Kong	Leung et al. (2006)	2033	6–16	Regional school-based	S	49	91
Japan	Itani et al. (2001)	4645	6–16	Regional school-based	S	48	91
Korea	Oh et al. (1997)	3081	6–16	National school-based	S	50	86
Singapore	Woo et al. (2007)	1605	6–11	National school-based	S	48	60
Taiwan	Yang et al. (2000)	834	12–16	Regional school-based	S	48	88
Vietnam	Weiss et al. (2014)	1315	6–16	National household	HI	50	99
Eastern Europe (*N* = 13,693)							
Croatia	Rudan et al. (2005)	2373	6–16	National school-based	S	48	98
Greece	Roussos et al. (1999)	1113	6–11	National school-based	S	50	95
Kosovo	Shahini (unpublished data, 2012)	1332	6–16	National school-based	S	51	67
Lithuania	Zukauskiene et al. (2003)	2920	6–16	National school-based	S	49	85
Poland	Wolanczyk (2003)	2479	6–16	National school-based	S	49	89
Romania	Domuta (2004)	990	6–11	Regional school-based	S	50	80
Russia	Hellinckx et al. (2000)	1998	12–16	National household	HI	50	71
Serbia	Markovic et al. (unpublished data, 2011)	488	6–11	Regional school-based	S	50	70
Germanic Europe (*N* = 6912)							
Flanders	Hellinckx et al. (1994)	953	6–11	Regional clinics	M	50	80
Germany	Döpfner et al. (1997)	2184	6–16	National household	HI	50	82
Netherlands	Verhulst et al. (1997)	1715	6–16	National household	HI	51	82
Switzerland	Steinhausen et al. (1997)	2060	6–16	Regional school-based	M	51	79
Latin America (*N* = 7024)							
Brazil	Rocha et al. (unpublished data, 2010)	1228	6–11	Regional convenience	O	51	NA
Colombia	Hewitt et al. (unpublished data, 2012)	1491	6–16	Regional school-based	S	53	95
Peru	Pomalima et al. (unpublished data, 2009)	2296	6–16	Regional household	HI	49	99
Puerto Rico	Achenbach et al. (1990)	635	6–16	Island-wide household	HI	50	92
Uruguay	Viola et al. (2011)	1374	6–11	National school-based	HI	49	79
Latin Europe (*N* = 5433)							
France	Fombonne and Vermeersch (1997)	2133	6–16	National gas company	M	52	62
Israel	Zilber et al. (1994)	1117	6–16	Regional household	HI	51	80
Italy	Frigerio et al. (2004)	1063	6–16	Regional school-based	M	48	72
Portugal	Fonseca (1995)	1120	6–11	Regional school-based	S	50	85
Middle East (*N* = 5361)							
Algeria	Petot et al. (2008)	384	11–16	Regional school-based	S	42	87
Tunisia	Chahed (2010)	1187	6–16	Regional school-based	S	41	85
Turkey	Erol and Simsek (1997)	3790	6–16	National household	HI	52	84
Nordic Europe (*N* = 5831)							
Denmark	Bilenberg, 1999	628	6–16	Regional household	M	45	60
Finland	Weintraub (unpublished data, 2004)	2093	6–16	Regional school-based	S	49	77
Iceland	Hannesdottir and Einarsdottir (1995)	817	6–16	Regional school-based	M	51	62
Norway	Novik (1999)	939	6–16	Regional household	M	48	45
Sweden	Larsson and Frisk (1999)	1354	6–16	Regional school-based	S	48	84
Southern Asia (*N* = 2814)							
Indonesia	Hartini et al. (2015)	308	6–14	Regional school-based	S	47	NA
Iran	Minaei (2005)	1205	6–16	Regional school-based	S	54	100

Table 1 (continued)

Society	Reference[a]	*N*	Age range	Sampling frame	Method[b]	% boys	Comp. rate (%)
Pakistan	Haider (unpublished data, 2011)	600	6–14	Regional school-based	S	50	87
Thailand	Weisz et al. (1987, 1993)	701	6–16	National school-based	HI	50	83
Sub-Saharan Africa ($N = 1535$)							
Ethiopia	Mulatu (1997)	513	11–16	Regional school-based	HI	49	92
Kenya	Malai et al. (2018)	1022	6–18	Regional household	HI	50	96

CBCL/6–18 Child Behavior Checklist for Ages 6–18

[a]The source document or publication for each sample is listed alphabetically by author in ESM Appendix 1

[b]Method codes: *M* forms mailed to parents, *S* forms completed by parents at school or sent home from school, *HI* home interview, *O* mixture of methods, *Comp. rate* completion rate, *NA* completion rate unknown

transformation [14] to correct their positive skew. The results were similar to those obtained with the untransformed scores. First, a null model was created using the CBCL/6–18 scale as the outcome variable at Level 1, society as the Level 2 variable, and culture cluster as the Level 3 variable. Full maximum likelihood estimation of variance components was performed with 100 iterations and a threshold of 1×10^{-6} percent change to stop iterating. Intercepts that were significantly associated with the Level-2 and Level-3 variables at $p < 0.05$ were considered significant for the model. The proportion of variance accounted for by each level was then estimated via the Interclass Correlation (ICC), calculated as the variance component accounted for at each level divided by the total variance.

In the next step, we performed a 3-level random coefficients regression model for each outcome variable to add child gender as a predictor at Level 1. We did not analyze age because not all 45 societies included the full age range of 6–16. Our 3-level random coefficients regression model allowed for modeling random effects associated with society and culture cluster on the outcome variables by examining the Level 2 and Level 3 error terms. A likelihood ratio test was performed within HLM to evaluate the difference between the null model and models with the added Level 1 predictor of gender entered as a binary variable. If the −2 log likelihood (−2 LL) was different at $p < 0.05$, the addition of gender was considered significant to the model. The proportion of variance accounted for by gender was also estimated by ICCs, computed as above.

To obtain more information about how society and culture cluster were associated with children's problem scale scores, we performed society × gender and culture × gender ANOVAs on Total Problems, Internalizing, and Externalizing scores for the 72,493 children in our samples. Our large samples warranted using $p \leq 0.001$ to determine significance. ANOVA ESs (η^2) were interpreted according to Cohen's [15] conventions: small = 1–6%, medium = 6–13%, large ≥ 14%. Finally, to test whether economic status was a better predictor of scale scores than culture cluster, we conducted an HLM analysis using the World Bank's estimates of Purchasing Power Parity per Capita (PPP) [16].

Results

As shown in Table 2, the mean ESs across scales for the variance components derived by HLM were 6.1% for society, 4.2% for culture cluster, and 89.8% for effects associated

Table 2 Percent variance associated with individual differences, society, and culture cluster in hierarchical linear models of parents' CBCL/6–18 ratings of 72,493 children in 45 societies

Scale	Individual differences	Society	Culture cluster
Broad-spectrum scales			
Internalizing problems	85.5	7.6	6.8
Externalizing problems	91.9	4.8	3.4
Total problems	86.9	6.8	6.4
Syndromes			
Anxious/Depressed	85.6	6.5	8.0
Withdrawn/Depressed	89.8	6.4	3.9
Somatic complaints	89.0	9.9	1.2
Social problems	88.7	4.8	6.5
Thought problems	91.8	5.5	2.6
Attention problems	89.0	6.7	4.2
Rule-breaking behavior	91.8	4.3	3.8
Aggressive behavior	93.3	4.7	2.0
DSM-oriented scales			
Depressive problems	90.0	6.1	3.9
Anxiety problems	86.3	3.8	10.0
Somatic problems	89.2	10.7	0.1
Attention deficit problems	91.2	5.2	3.6
Oppositional defiant problems	93.1	4.9	2.00
Conduct problems	92.9	4.6	2.6
Average across 17 scales	89.8	6.1	4.2

All effects were significant at $p < 0.001$ in HLM analyses

with individual differences and factors not measured in the model. These results suggest that society was a slightly better predictor of scale scores than culture cluster, although HLM did not directly test this difference. Clearly, most of the variance in parents' ratings of children's problems was not accounted for by either society or culture, as together they accounted for only 10.3% of the variance. Adding gender into the random coefficient regression model yielded significant fixed effects for most scales, but these effects were consistently very small.

Might particular kinds of problems show bigger effects of society or culture cluster? The 10-item Somatic Complaints syndrome and the 7-item DSM-Somatic Problems scale (which have seven items in common) showed the largest effects of society (9.9% and 10.7%, respectively). Interestingly, these two scales tapping somatic concerns showed the smallest effects of culture cluster (1.2% and 0.1%). In contrast, the 13-item Anxious/Depressed syndrome and the 9-item DSM-Anxiety Problems scale (also with seven items in common) showed the largest effects of culture cluster (8.0% and 10.0%, respectively). DSM-Anxiety Problems showed the smallest effect of society (3.8%). The 6.5% effect of society on Anxious/Depressed was near the mean of 6.1% across all 17 scales. Internalizing showed larger effects than Externalizing for both society (7.6% vs. 4.8%) and culture cluster (6.8% vs. 3.4%).

The 10 (culture cluster) × 2 (gender) ANOVA on Total Problems yielded ESs of 6.2% for culture cluster and < 1% for gender. In contrast, the ANOVA for Total Problems using society rather than culture cluster yielded ESs of 11.4% for society and < 1% for gender. Similar results were obtained for Internalizing and Externalizing, with Internalizing consistently showing larger society (11.6%) and culture cluster effects (6.3%) than Externalizing (7.7% and 4.3%). The gender effect was < 1% for both scales in both the society and the culture cluster analyses. Student–Neuman–Keuls (SNK) post hoc tests for Total Problems indicated significant differences between most culture clusters: Nordic (M = 16.4, SD = 14.2) < Germanic and Confucian Asian (respectively, M = 18.7, SD = 15.8 and M = 19.3, SD = 17.8) < Anglo (M = 20.6, SD = 18.5) < Eastern Europe and Latin Europe (respectively, M = 23.9, SD = 18.6 and M = 24.5, SD = 17.0) < Middle East (M = 25.3, SD = 18.4) < South Asia (M = 27.2, SD = 19.6) < Sub-Saharan Africa (M = 29.0, SD = 21.8) < Latin America (M = 33.3, SD = 21.6).

Figure 1 presents the effects of culture cluster on Internalizing and Externalizing scores, with the 10 culture clusters arranged in ascending order for mean Internalizing score. SNK post hoc tests indicated significant differences between each of the four culture clusters with the lowest Internalizing scores. The other six culture clusters were grouped into three pairs (e.g., Eastern Europe and Latin Europe), each pair differing significantly from all other clusters. As shown in Fig. 1, the rank-ordering of culture clusters for Externalizing differed somewhat from the rank-ordering for Internalizing. The most notable differences were for the Middle East and Sub-Saharan Africa, which had among the highest Internalizing scores but Externalizing scores in the middle of the ranking. SNK post hoc tests indicated that each culture cluster differed significantly from all others except for two pairs of clusters (Middle East and Sub-Saharan Africa, and Latin and Eastern Europe), which did not show significant within-pair differences.

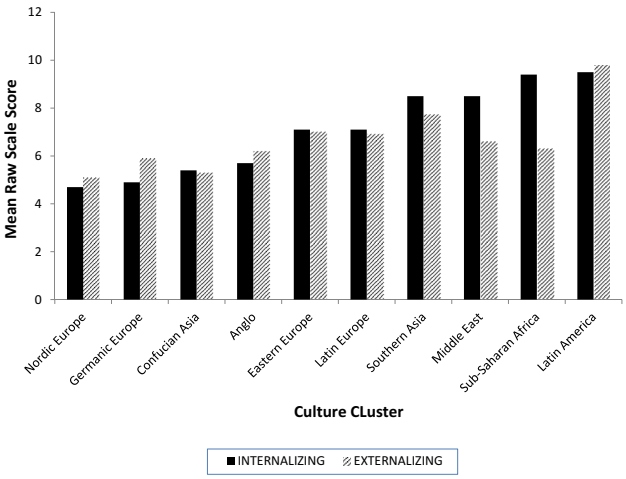

Fig. 1 Mean Scale scores by culture cluster for Internalizing and Externalizing

Although GLOBE assignments of societies to culture clusters were based on values and practices differentially endorsed by more than 17,000 participants in 62 societies, perhaps the rank-ordering of culture clusters for CBCL/6–18 scores might reflect economic development. Numerous within-society studies have found small negative associations between CBCL/6–18 problem scores and socioeconomic status (SES) [17]. In our study, the societies with the lowest Total Problems scores came from Nordic, Germanic, Confucian Asian, and Anglo culture clusters, which include societies with the most advanced economies. Conversely, the societies with the highest CBCL Total Problems scores came from culture clusters with economically less advanced societies, namely the Middle East, South Asia, Sub-Saharan Africa, and Latin America. When we replaced culture cluster with the World Bank's PPP index [16] in an HLM analysis of Total Problems, we found that society accounted for 10.4% of the variance, PPP accounted for 2.4%, and individual differences/other variables accounted for 87.2%. These results indicate that PPP added little to the variance explained by society. In contrast, society and culture cluster explained rather similar amounts of variance (6.1% and 4.2%, respectively) but the total variance explained by society + PPP (12.4%) was slightly greater than the total variance

explained by society + culture cluster (10.3%). In both HLM analyses, most of the variance in problem scores was associated with individual differences and other variables.

Discussion

In the Rescorla et al. [3] study, society accounted for about 10% of the variance in Total Problems scores. In the current study, the combined effects of society plus culture cluster were also about 10%. This suggests that our previously reported effects of society included culture cluster effects, but that about the same amount of variance was attributable to individual differences and other variables regardless of whether only society or both culture cluster and society were included in the models. Furthermore, the ANOVAs in the current study for Total Problems, Internalizing, and Externalizing indicated that, when tested alone, society had a larger effect than culture cluster (11% vs. 6%, 12% vs. 6%, and 8% vs. 4%, respectively).

Both Rescorla et al. [3] and the current study show that for standardized assessment of children's mental health problems via parents' ratings, the effects associated with society and/or culture cluster were significant and small-to-medium in size by Cohen's (1988) benchmarks. Across the 17 problem scales, the effect sizes for society ranged from 3.8 to 10.7%, whereas those for culture cluster ranged from 0.1 to 10.0%. Perhaps our most notable findings regarding effects of society and culture cluster on particular kinds of problems were that the results for the two somatic scales (Somatic Complaints and DSM-Somatic Problems) contrasted markedly with the results for the two anxiety scales (Anxious/Depressed and DSM-Anxiety Problems). Specifically, the two somatic scales showed the largest effects of society (9.9% and 10.7%) and the smallest effects of culture cluster (1.2% and 0.1%), suggesting that parents in different societies within the same culture cluster varied widely in their endorsement of somatic problems in their children. By contrast, the two anxiety scales showed the largest effects of culture cluster (8.0% and 10.0%), with DSM-Anxiety Problems (but not Anxious/Depressed) showing the smallest effect of society (3.8% vs. 6.5%). Perhaps the difference in societal ESs for the two anxiety scales reflects inclusion of depressive items in the Anxious/Depressed syndrome but not in the DSM-Anxiety Problems scale. It may be that parents in different societies within a culture cluster are more similar in their ratings of anxiety problems than in their ratings of depressive problems, leading to a smaller societal effect for DSM-Anxiety Problems than for Anxious/Depressed. Given the narrow ranges of effect sizes for society and culture cluster, we interpret these findings cautiously, but they might warrant further examination.

Although society and culture cluster clearly have important, and sometimes contrasting effects, on parents' ratings of children's problems, these effects of society and culture cluster nevertheless explained only a small percent of the variance, thus indicating great heterogeneity within societies and culture clusters in parents' ratings of children's problems. The percent of variance associated with unmeasured variables leading to individual differences was very large across all 17 problem scales (85.5–93.3%), thus leaving most of the variance accounting for individual differences unexplained.

A limitation of our HLM analyses is that the effects associated with individual differences included residual variance that may be attributable to other variables. We could not analyze other possible influences on parents' ratings of children's problems because the indigenous investigators who obtained data for 72,493 children could not uniformly assess additional variables that do not lend themselves to standardized assessment across such diverse societies. Examples of such variables include each child's family, school, and neighborhood. Had we been able to add child-specific variables to our models, this might have reduced the estimated effects of individual differences. However, the relatively small effects attributed to society and culture cluster would probably not increase if more of the residual individual variance were explained by child-specific variables. This assumption is supported by our findings of little change in effects associated with society and culture cluster when gender was added to the models. That is, even though gender effects were very consistently found in this study, as well as in the Rescorla et al. study [3] (e.g., girls scoring higher on internalizing scales and boys scoring higher on externalizing and attention problems scales), these gender effects were also consistently very small. Moreover, adding gender effects to our HLM models did not affect the magnitude of the effects for society and culture cluster.

Additional limitations of our study are that our 45 samples varied in representativeness and response rate; that we could not test the effects of regions within societies; and that some of the samples used pre-2001 forms, necessitating omission of the six items changed in 2001. Additionally, our data were collected over about two decades, but Rescorla et al. [18] reported a negligible though significant correlation ($r = 0.12$) between data collection year and CBCL/6–18 scores for 67,806 children in 42 of our 45 societies, suggesting that data collection year minimally affected our results.

Despite these limitations, the very large sample size, the large number of societies, and their range with respect to region, political/economic systems, history, population, ethnicity/race, and religion justify confidence in our findings of relatively small effects associated with society and culture cluster and very large effects associated with individual differences within societies and culture clusters on parents'

ratings of children's problems. Our study could not include social and familial variables such as poverty, discrimination, and family conflict/cohesion, which other studies have shown to be associated with children's problems. If such variables had been included in our models, they would most likely have explained some of the variance associated with individual differences.

Clinical implications

HLM analyses revealed that much more variance in parents' ratings of mental health problems for 72,493 children in 45 societies was associated with individual differences than with societal or culture cluster differences. Variance associated with individual differences doubtless reflects a host of genetic and environmental factors that combine in complex ways to make children different from one another, none of which our study could measure. However, our findings indicate that standardized parent ratings are robust to effects associated with society and culture cluster, suggesting that they can be used in diverse societies to assess parents' perceptions of children's problems. To advance this line of research beyond parent ratings, future research should test effects of society and culture cluster on self-ratings by adolescents and adults, as well as on teachers' ratings of students and collaterals' ratings of adults. Gender, age, and societal differences can be accommodated in clinical practice by using norms based on gender and age group for children from low-, medium-, or high-scoring societies [19].

The evidence supporting the generalizability of data obtained from applying the same standardized assessment procedure in 45 societies does not minimize the influence of socioeconomic and family factors on the assessed problems. Additionally, although the problems we analyzed in this and previous studies yielded very consistent results across societies from many regions of the world, it is possible that problems other than those we assessed might show larger effects of society and culture cluster. It would, therefore, be worth conducting emic studies to identify problems that might be more specific to particular populations.

Funding Dr. Althoff receives grant or research support from the National Institute of Mental Health (MH082116), the National Institute of Health Center of Biomedical Research Excellence award (P20GM103644), and the Klingenstein Third Generation Foundation. This research was also supported by the non-profit University of Vermont Research Center for Children, Youth, and Families.

Compliance with ethical standards

Ethical approval No animals were used in this research. All procedures involving human participants were in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

Conflict of interest The first, third, and fourth authors receive salary and/or other compensation from the non-profit University of Vermont Research Center for Children, Youth, and Families, which publishes the instrument used in this research. The second author declares that he has no conflict of interest.

References

1. House RJ, Hanges PJ, Javidan M et al (2004) Culture leadership and organizations: the GLOBE Study of 62 societies. Sage Publications, Thousand Oaks
2. Geisinger K (1994) Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. Psychol Assess 6:304–312
3. Rescorla L, Ivanova MY, Achenbach TM et al (2012) International epidemiology of child and adolescent psychopathology II: integration and applications of dimensional findings from 44 societies. J Am Acad Child Adolesc Psychiatry 51:1273–1283. https://doi.org/10.1016/j.jaac.2012.09.012
4. Rescorla LA, Achenbach TM, Ivanova MY et al (2007) Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. J Emot Behav Disord 15(3):130–142. https://doi.org/10.1177/10634266070150030101
5. Achenbach TM, Rescorla LA (2001) Manual for the ASEBA school-age forms & profiles. University of Vermont Research Center for Children Youth and Families, Burlington
6. Nisbett RE, Peng K, Choi I, Norenzayan A (2001) Culture and systems of thought: holistic versus analytic cognition. Psychol Rev 108:291–310
7. Talhelm T, Zhang X, Oishi S et al (2014) Large-scale psychological differences within China explained by rice versus wheat agriculture. Science 344:603–608
8. Hofstede G (1980) Culture's consequences. Sage Publications, Beverly Hills
9. Triandis HC (1995) Individualism and collectivism. Westview Press, Boulder
10. Stankov L (2011) Individual, country and societal cluster differences on measures of personality, attitudes, values, and social norms. Learn Individ Differ 21:55–66. https://doi.org/10.1016/j.lindif.2010.09.002
11. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Washington DC. https://doi.org/10.1007/s00787-004-2004-4
12. Achenbach TM (2014) DSM-oriented guide for the Achenbach System of Empirically Based Assessment (ASEBA). University of Vermont Research Center for Children Youth and Families, Burlington
13. Raudenbush SW, Bryk AS, Cheong YF et al (2011) HLM 7: linear and nonlinear modeling. Scientific Software International Inc, Lincolnwood
14. Box GEP, Cox DR (1964) An analysis of transformations (with discussion). J R Stat Soc Ser B Stat Methodol 26:211–252
15. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Academic Press, New York
16. World Bank (2015) Purchasing power parities and the real size of world economies. World Bank, Washington DC
17. Achenbach TM, Rescorla LA (2007) Multicultural understanding of child and adolescent psychopathology: implications for mental health assessment. Guilford Press, New York

18. Rescorla LA, Blumenthal MC, Ivanova MY, Achenbach TM (2018) International comparisons of the dysregulation profile based on reports by parents, adolescents, and teachers. J Clin Child Adolesc Psychol. https://doi.org/10.1080/15374416.2018.1469090

19. Achenbach TM, Rescorla LA (2007) Multicultural supplement to the manual for the ASEBA school-age forms & profiles. University of Vermont Research Center for Children Youth and Families, Burlington