International Perspectives in Psychology: Research, Practice, Consultation
2016, Vol. 5, No. 2, 91–109

2157-3883/16/$12.00 http://dx.doi.org/10.1037/ipp0000046

Problems and Adaptive Functioning Reported by Adults in 17 Societies

Leslie A. Rescorla
Bryn Mawr College

Thomas M. Achenbach, Masha Y. Ivanova, Lori V. Turner, and Robert R. Althoff
University of Vermont

Hervör Alma Árnadóttir
University of Iceland

Alma Au
Hong Kong Polytechnic University

Monica Bellina
Eugenio Medea Scientific Institute

J. C. Caldas
Instituto Superior de Ciencias da Saude-Norte

Yi-Chuen Chen
National Chung Cheng University

Ladislav Csemy
National Institute of Mental Health

Marina M. da Rocha
Mackenzie Presbyterian University

Jeroen Decoster and Johnny R. J. Fontaine
Ghent University

Yasuko Funabiki
Kyoto University Hospital

Halldór S. Guðmundsson
University of Iceland

Valerie S. Harder
University of Vermont

Young Ah Kim
Huno Consulting, Seoul, Republic of Korea

Patrick Leung
The Chinese University of Hong Kong

David M. Ndetei
Africa Mental Health Foundation, Nairobi, Kenya

Jelena Srdanović Maraš
Clinic for Psychiatry, Clinical Center of Vojvodina, Novi Sad, Serbia

Jasminka Marković
Clinic for Psychiatry, Clinical Center of Vojvodina, Novi Sad, Serbia

Kyung Ja Oh
Yonsei University

Virginia C. Samaniego
Pontificia Universidad Católica Argentina

Sandra Sebre
University of Latvia

Edwiges Silvares
University of Sao Paolo

Roma Simulioniene
Klaipeda University

Elvisa Sokoli
University of Tirana

Natalia Vazquez
Pontificia Universidad Católica Argentina

Ewa Zasepa
The Maria Grzegorzewska Academy of Special Education

This study tested for similarities and differences across societies in self-ratings of problems, personal strengths, and aspects of adaptive functioning on the Adult Self-Report (ASR) for nonclinical samples of adults ages 18 to 59 in 17 societies (N = 10,197). Results indicated considerable consistency across societies regarding mean ratings on the ASR problem items. Most effect sizes (ESs) for societal differences in problem scales were small (2–5%). Hierarchical linear modeling (HLM) analyses indicated that culture clusters and society accounted for small percentages of variance in Internalizing, Externalizing, and Total Problems scores, with most of the variation accounted for by individual differences within societies. In contrast to the small effects of society on problem scores, for the ASR Personal Strengths scale the societal ES was 34% and culture cluster accounted for 12% of the variance. Worse reported relations with spouse/partner were associated with higher problem scores. Overall, findings indicated considerable similarity but also some important differences in self-reported problems and adaptive functioning across 17 societies.

Keywords: adult psychopathology, ASR, self-reported problems, cross-cultural, international comparisons

Mental disorders comprise about 14% of the global health burden worldwide, are linked to many other health problems, and are among the most costly disorders to treat (Tomlinson et al., 2009). Tomlinson et al. noted that global implementation of evidence-based mental health practices is hampered by the fact that most mental health research has been done in Western countries. This argues for research on mental health assessment in more countries than have been studied to date. Research that compares findings from the same instrument in different societies exemplifies what Pike (1967) called the *etic* approach. Etic research is often contrasted with *emic* examination of constructs specific to particular societies.

Etic epidemiological mental health research compares the prevalence, distribution, and correlates of mental health problems in different populations. Such research can reveal similarities and differences among societies in (a) the prevalence of categorically defined disorders (e.g., depression diagnosed according to criteria such as those of the American Psychiatric As-

This article was published Online First February 15, 2016.

Leslie A. Rescorla, Department of Psychology, Bryn Mawr College; Thomas M. Achenbach, Masha Y. Ivanova, Lori V. Turner, and Robert R. Althoff, Department of Psychiatry, University of Vermont; Hervör Alma Árnadóttir, Faculty of Social Work, University of Iceland; Alma Au, Department of Applied Social Sciences, Hong Kong Polytechnic University; Monica Bellina, Department of Child Psychiatry, Eugenio Medea Scientific Institute; J. C. Caldas, Department of Social Sciences and Behavior, Instituto Superior de Ciencias da Saude-Norte; Yi-Chuen Chen, Department of Psychology, National Chung Cheng University; Ladislav Csemy, National Institute of Mental Health; Marina M. da Rocha, Developmental Disorders Program, Centre for Health and Biological Sciences, Mackenzie Presbyterian University; Jeroen Decoster and Johnny R. J. Fontaine, Department of Personnel Management, Work, and Organizational Psychology, Ghent University; Yasuko Funabiki, Department of Psychiatry, Kyoto University Hospital; Halldor S. Guðmundsson, Faculty of Social Work, University of Iceland; Valerie S. Harder, Department of Pediatrics, University of Vermont; Young Ah Kim, Huno Consulting, Seoul, Republic of Korea; Patrick Leung, Department of Psychology, The Chinese University of Hong Kong; David M. Ndetei, Africa Mental Health Foundation, Nairobi, Kenya; Jelena Srdanović Maraš, Clinic for Psychiatry, Clinical Center of Vojvodina, Novi Sad, Serbia; Jasminka Marković, Clinic for Psychiatry, Clinical Center of Vojvodina, Novi Sad, Serbia; Kyung Ja Oh, Department of Psychology, Yonsei University; Virginia C. Samaniego, Department of Psychology, Pontificia Universidad Católica Argentina; Sandra Sebre, Department of Psychology, University of Latvia; Edwiges Silvares, Institute of Psychology, University of Sao Paolo; Roma Simulioniene, Department of Psychology, Klaipeda University; Elvisa Sokoli, Department of Psychology, University of Tirana; Natalia Vazquez, Department of Psychology, Pontificia Universidad Católica Argentina; Ewa Zasepa, The Maria Grzegorzewska Academy of Special Education.

The ASR is published by the nonprofit Research Center for Children, Youth, and Families from which Drs. Achenbach, Althoff, Ivanova, and Rescorla receive remuneration.

Correspondence concerning this article should be addressed to Leslie A. Rescorla, Department of Psychology, Bryn Mawr College, 101 North Merion Avenue, Bryn Mawr, PA 19010. E-mail: lrescorl@brynmawr.edu

sociation's (APA) *Diagnostic and Statistical Manual, Fifth Edition* (*DSM–5*; APA, 2013), and/or (b) scores on quantitative scales derived from rating instruments (e.g., scales measuring depression).

The World Health Organization (WHO) conducted an etic multinational comparison of adult psychiatric disorders in 14 societies (WHO, 2004) using interviews based on the *DSM–IV* (APA, 1994). Four general classes of disorders were studied in each society (anxiety disorders, mood disorders, impulse-control disorders, and substance use disorders), although the number of different diagnoses assessed within each of these broad categories varied across societies. Response rates ranged from 46% (France) to 88% (Colombia). Similarities in findings across the societies included that anxiety disorders were generally the most common diagnoses and that severity of disorders was significantly associated with impairment. Despite these similarities, prevalence estimates for meeting the criteria for at least one diagnosis ranged from 4% in Shanghai to 26% in the United States (U.S.).

Although diagnostic interviews may be useful for large multinational mental health projects, their cost may be prohibitive for indigenous investigators conducting epidemiological research in their own societies. It is more practical for indigenous investigators in different societies to use low-cost, standardized assessment instruments to assess mental health, at least as a first step in multistep screening. However, for such instruments to be valid for international applications, they need to be tested in different societies. Many theorists have proposed methods for assessing the international equivalence of assessment instruments, as summarized briefly below.

Geisinger (1994) argued that "substantial evidence of the comparability" of a translated/adapted instrument and the original instrument is needed as a basis for cross-national assessment. Poortinga (1989) discussed various procedures to test for invariance of instruments across cultures, including comparisons of correlations among scales, of factor structure, and of item difficulty order. Poortinga also noted that when instruments are not invariant across cultures, the chances of finding cross-cultural differences increases. Van de Vijver and Poortinga (1997) stressed the importance of discriminating between spurious cultural differences based on measurement artifacts (denoted as "bias") and valid group differences (denoted as "impact"). Additionally, van de Vijver and Poortinga (1997) argued that testing equivalence of an instrument across cultures should include looking for both method bias (e.g., cultural differences in response sets/styles, or response formats favoring some cultures but not others) and item bias/differential item functioning (e.g., relations between item scores and total scores varying across cultures).

Butcher and Han (1996) proposed several methods for testing instrument equivalence across different societies, including examining whether items perform similarly, using confirmatory factor analysis (CFA), and comparing levels and correlates of scale scores. Leung and Bond (1989) described several ways to test associations between variables *X* and *Y* in samples composed of different cultural groups, including taking the mean for *X* and *Y* in each of 10 cultures and correlating them versus computing one correlation between all the *X* and *Y* values within each culture. Byrne and Campbell (1999) demonstrated that even when data from different countries fit a prescribed factor structure, item factor loadings, skew, kurtosis, and response frequencies may differ across countries. Schwartz and Boehnke (2004) reported that they used several different methods to test the equivalence of the same theoretical model of 10 basic human values across countries, including multidimensional scaling, similarity structure analysis, and CFA. In sum, various methods have been used to establish equivalence of instruments across societies.

The ASR (Achenbach & Rescorla, 2003) is a rating form standardized in the U.S. for assessing behavioral, emotional, social, and thought problems; substance use; personal strengths; and adaptive functioning. It is designed for adults ages 18–59 and takes about 15–20 min to complete. The ASR is a revision and extension of the Young Adult Self-Report (Achenbach, 1997) for ages 18 to 30. The ASR contains 120 problem items written at a fifth-grade reading level that respondents rate as 0 (*not true*), 1 (*somewhat or sometimes true*), and 2 (*very true or often true*) based on the preceding 6 months. Factor analyses of the ASR's 120 problem items using U.S. data yielded eight syndromes: Anxious/Depressed, Withdrawn, and Somatic

Complaints (all loading on a broad-band, second-order factor designated as Internalizing); Rule-Breaking Behavior, Aggressive Behavior, and Intrusive (all loading on a broad-band, second-order factor designated as Externalizing); and Thought Problems and Attention Problems (not loading differentially on either second-order factor) (Achenbach & Rescorla, 2003). Subsets of the 120 problem items are also used to score six *DSM*-oriented scales (Depressive Problems, Anxiety Problems, Somatic Problems, Avoidant Personality Problems, Attention Deficit Hyperactivity Problems, and Antisocial Personality Problems), which comprise ASR items identified by experts from 17 societies as being very consistent with diagnostic categories of the *DSM–5* (APA, 2013). Subsets of problem items also comprise scales denoted as Obsessive-Compulsive Problems and Sluggish Cognitive Tempo, based on research by others. Note that ratings of all 120 problem items are summed to yield the broad-band Total Problems score, while subsets of items are summed to yield scores on the other problem scales specified above. Interspersed among the 120 problem items are 11 Personal Strengths items (e.g., "I make good use of my opportunities"). These are rated on the same 0–1–2 scale as the 120 problem items.

In addition to the 120 problem items and 11 Personal Strengths items, items comprising five Adaptive Functioning scales (Friends, Spouse/Partner, Family, Job, and Education) are included in the ASR. The Spouse/Partner, Job, and Education scales include some items with the same 0–1–2 Likert scale as the problem and Personal Strengths items. All five Adaptive Functioning scales have other items requesting specific information (e.g., whether or not the respondent has had a spouse/partner or a job in the preceding 6 months). The ASR also includes three substance use items, but results for these items are not presented in this article.

In a multinational study of the ASR, Ivanova et al. (2015) used separate CFAs for 18- to 59-year-olds in 29 societies differing widely in social, political, and economic systems (N = 17,152) to test the fit of ASR problem item ratings to the eight-syndrome model derived from factor analyses of U.S. data. Ivanova et al. used the term "society" rather than "nation" or "country" because some of the populations from which they obtained samples do not constitute nations, such as Hong Kong and Flanders. Ivanova did not use CFAs to test the three broad-band scales, the *DSM*-oriented scales, or the Adaptive Functioning scales. In all 29 samples, the eight-syndrome model converged, the primary fit index (root mean square error of approximation) indicated good model fit, and the secondary indices (comparative fit index and Tucker-Lewis index) indicated acceptable to good fit. Only 0.06% of the 8,301 tested parameters fell outside the admissible parameter space. Item loadings were robust across societies. Ivanova et al. thus concluded that the results supported the eight-syndrome model in all 29 samples.

Ivanova et al.'s (2015) CFA study provides an important first step in testing the ASR's international use. However, the ASR is a broad instrument with a hierarchical structure (items, narrow-band scales, and broad-band scales), and it taps many different aspects of adult functioning (i.e., problems, substance use, strengths, and adaptive functioning). These features of the ASR enable testing of similarities and differences across various societies in multiple areas of functioning and using multiple methods.

To address issues not addressed by Ivanova et al. and to further advance international mental health research based on the ASR, we analyzed data from 17 of the 29 societies analyzed by Ivanova et al. (2015). We excluded 12 demographically limited samples (such as samples that included only university students). Consistent with Ivanova et al., we used the term "society" rather than "nation" to encompass samples from populations that are not nations. We used data from the 17 non-U.S. societies to address the following issues. First, we tested internal consistencies of the ASR's scales. Second, we determined the prevalence of specific problems such as "I have trouble planning for the future" and "I think about killing myself" by analyzing means of the 0–1–2 ratings for the 120 problem items in each society. Third, we tested the effects of society, age, and gender on the eight-syndrome scales Ivanova tested and on the three broad-band scales, the *DSM*-oriented scales, the Personal Strengths scale, and the Adaptive Functioning scales. Moreover, we used HLM to test the effects of culture clusters and per capita income on problem scores.

Purpose of the Study

Our purpose was to identify similarities as well as differences across societies in adults' self-reports of behavioral, emotional, social, and thought problems, personal strengths, and adaptive functioning. We used data from 17 non-U.S. societies differing widely in economic, political, ethnic, religious, and cultural characteristics (N = 10,197) to answer the following questions: (a) What are the internal consistencies of ASR scales across societies? (b) How similar are the 17 societies with respect to mean ratings on the ASR's 120 problem items? (c) What are the effects of society, gender, and age on ASR scale scores? (d) How much variance is accounted for by society, culture cluster, and individual differences within societies when multilevel models are tested with two different typologies of culture cluster? We hypothesized (a) that the pattern of internal consistencies of ASR scales would be similar to those reported for the U.S., (b) that there would be strong consistency across societies with respect to prevalence of specific problems based on means of the 0–1–2 ratings, (c) that societal effects on ASR scores would be significant but relatively modest for most problem scales and Adaptive Functioning scales but large for Personal Strengths, (d) that age and gender effects would be small but consistent across societies, and (e) that culture cluster and society would account for small portions of the variance in problem scores when analyzed using HLM, with most of the variation due to individual differences within societies.

Method

Samples

The U.S. normative data were collected in 1999, and the ASR was published in 2003. Over the ensuing decade, investigators in many societies requested permission to translate the ASR and use the translated form for data collection in their own societies. The data for each society were collected by indigenous investigators from 2010 to 2012 and then sent to the lead authors for analysis. In each of the 17 societies, conventions for obtaining informed consent required by the investigator's research institution were followed.

As noted above, we used data from 17 of the 29 societies analyzed by Ivanova et al. (2015). Following the recommendation of Nunnally and Bernstein (1994), we required a minimum N of 300 per society, with Ns ranging from 302 (Latvia) to 1,548 (Flanders). As shown in Table 1, rigorous random sampling methods were used in some societies, resulting in representative population samples. However, in other societies, various methods of convenience sampling were used, resulting in samples of unknown representativeness. Although our primary analyses involved the 17 non-U.S. societies (N = 10,197), we also conducted subsidiary analyses for 18 societies (N = 12,217) by adding data from the U.S. to those from the other 17 societies, so that all 18 societies could be compared statistically.

Instrument

Foreign language versions of the ASR (Achenbach & Rescorla, 2015) were developed by indigenous mental health researchers who first did translations and then obtained independent back-translations. The ASR's 120 problem items tap a wide variety of emotional, behavioral, social, and thought problems, such as "I worry about my family"; "I am stubborn, sullen, or irritable"; "I argue a lot"; and "I have thoughts that other people would think are strange." The 11 Personal Strengths items (e.g., "I make good use of my opportunities," "I work up to my ability," "I am pretty honest," "I meet my responsibilities to my family," "I try to be fair to others," and "I am a happy person") are rated on the same 0–1–2 scale as the problem items, with high ratings indicating positive characteristics.

On all five Adaptive Functioning scales, higher scores indicate better functioning. On the Friends scale, respondents report on number of friends, frequency of contacts with friends, getting along with friends, and visits by friends and family. On the Spouse/Partner scale, there are eight items rated 0, 1, or 2 dealing with how well respondents get along with their partner, share responsibilities, enjoy similar activities, are satisfied with their partner, and like their partner's friends and family. The Family scale contains items that respondents rate on a Likert scale (*worse than average*, *variable or average*, *better than average*) about how well they get

Table 1
Reference, N, Percent Male, Mean Age (SD), and Sampling Procedure in 17 Societies (N = 10,197), Plus the U.S.[a]

Society	Reference	*N*	Male (%)	Response rate (%)	Mean age (*SD*)	Sampling procedure
Albania	Sokoli (2013)	750	50	76	37.3 (12.8)	Random sampling of enumeration areas (EAs) of 80–100 families, based on census districts. Representative sample stratified by age, gender, region and urban-rural areas within each EA; interviewed at home.
Argentina[b]	Samaniego and Vázquez (2012)	679	48	unknown	35.7 (12.0)	Student research assistants recruited acquaintances; forms completed via interview; convenience sample stratified by level of educational attainment to be representative of the greater Buenos Aires area.
Brazil	Silvares and da Rocha (2012)	813	41	85	34.5 (11.7)	Recruited in public places (e.g., shops, stations, waiting rooms); form completed then or returned by mail; convenience sample stratified by region, age, gender, and socioeconomic status, representative of the metropolitan population.
Czech Republic	Csemy (2012)	588	51	90	37.8 (12.4)	Recruited from randomly selected households within randomly selected electoral districts; participants interviewed at home; stratified by region, age, gender, and education to be representative of the Czech population.
Flanders (Belgium)	Decoster and Fontaine (2012)	1,548	50	unknown	38.6 (12.2)	Student assistants recruited acquaintances via the Internet; convenience sample stratified by region, gender, age, and education, to be representative of Flanders, the Dutch-speaking region of Belgium.
Hong Kong	Au and Leung (2012)	324	39	95	29.4 (12.7)	Student research assistants recruited acquaintances; ASRs completed in person; convenience sample stratified by age and gender to be representative of the Hong Kong population.

Table 1 (*continued*)

Society	Reference	*N*	Male (%)	Response rate (%)	Mean age (*SD*)	Sampling procedure
Iceland	Guðmundsson and Árnadóttir (2012)	353	45	48	37.5 (12.0)	Recruited by stratified random sampling using national register; ASRs obtained by mail or website; representative sample of Icelandic population.
Italy	Bellina (2012)	519	46	35	38.0 (12.4)	Recruited a stratified random sample drawn from electoral rolls; ASRs mailed or dropped at the home; representative sample of Lecco province.
Japan	Funabiki (2012)	1,000	47	unknown	38.2 (10.7)	Professional firm recruited participants from all regions of Japan via the Internet; Japanese national sample stratified by age and gender.
Kenya	Harder and Ndetei (2012)	427	40	46	38.9 (85)	Adult relatives of regional sample of school-aged children, with children's names randomly drawn from class rosters.
Korea[b]	Kim, Kim, Lee, Kim, and Oh (2014)	1,000	51	unknown	37.9 (9.8)	Recruited in different regions by a survey firm; respondents interviewed at home; representative national sample, randomly drawn from the national registry, stratified by age, gender, and education.
Latvia	Sebre (2012)	302	43	unknown	33.9 (12.7)	Student research assistants recruited acquaintances according to stratification criteria; ASRs completed in person; convenience sample stratified by age, gender, educational attainment, and region to be representative of Latvia.
Lithuania[b]	Simulionienė, Brazdeikienė, Rugevičius, Gedutienė, and Žakaitienė (2010)	573	48	96	35.3 (11.1)	Recruited by stratified random sampling using national statistics/census information; participants interviewed at home; representative national sample randomly drawn from the Lithuanian national registry, with stratification by gender, age, and educational attainment.

(*table continues*)

Table 1 (*continued*)

Society	Reference	*N*	Male (%)	Response rate (%)	Mean age (*SD*)	Sampling procedure
Poland	Zasepa (2012)	310	37	71	36.7 (11.9)	Recruited by stratified random sampling using national statistics/census information; participants interviewed at home; representative sample stratified by age, gender, residence, and educational attainment to be representative of the Polish population.
Portugal	Caldas (2012)	397	49	unknown	35.4 (12.0)	Student research assistants recruited acquaintances; ASRs sent and returned by mail or in person; convenience sample stratified by age and gender to be representative of the Portuguese population.
Serbia	Markovic (2012)	314	42	72	35.7 (10.6)	Participants recruited by age-stratified random sampling using national statistics/ population registry information; ASRs sent and returned by mail; representative sample of the Novi Sad metropolitan area.
Taiwan	Chen (2012)	300	50	99	37.0 (11.9)	Participants self-referred based on notices posted on bulletin boards, Internet message boards, or by word of mouth; ASRs sent and returned by mail; convenience sample stratified by region, gender, and age to be representative of the Taiwan population.
U.S.[b]	Achenbach and Rescorla (2003)	2,020	41	94	39.1 (12.0)	Recruited by stratified random sampling via households in 40 states using national statistics/census information; participants interviewed at home; representative sample stratified by age, gender, and urban-suburban-rural residence to be representative of the U.S. population.

Note. ASR = Adult Self-Report.
[a] The U.S. sample was not included in the primary analyses reported here, which were based on 17 non-U.S. societies However, for some subsidiary analyses, we included the U.S. data and report findings for 18 societies (N = 12,217). [b] Indicates a published source for the data file. All other data files represent unpublished sources.

along with various family members (e.g., siblings, parents, children, etc.). On the Job scale, there are eight items (rated 0, 1, or 2) dealing with how well respondents get along with coworkers and bosses, do their work, find their work satisfying or stressful, worry about work, or do things that may cause them to lose their job. On the Education scale, there are five items (rated 0, 1, or 2) dealing with how well respondents get along with fellow students, achieve in their studies, finish their work, feel satisfied with their educational situation, and do things that may cause them to fail. Respondents who were not living with a spouse/partner, did not have a job, or were not enrolled in an education program during the preceding 6 months omitted those sections.

Based on U.S. data, Achenbach and Rescorla (2003) reported alphas of .89 to .97 for the ASR broad-band Internalizing, Externalizing, and Total Problems scales, .51 to .88 for the syndromes, .68 to .84 for the *DSM*-oriented scales, and .51 to .68 for the Adaptive Functioning scales. As reported by Achenbach and Rescorla (2003), the ASR's 1-week test-retest correlations (*r*s) were .89 to .94 for the broad-band scales, .78 to .91 for the syndromes, .77 to .86 for the *DSM*-oriented scales, and .71 to .85 for the Adaptive Functioning scales. ASR items and scales significantly discriminated between demographically similar clinically referred and nonreferred samples of adults, with referral status accounting for 50% of the variance in mean Adaptive Functioning and 13% of the variance in Total Problems.

Data Analysis

Forms that lacked ratings for $>$ 8 problem items were excluded from all analyses (0% of total forms for Argentina, Czech Republic, Japan, Korea, Lithuania, Poland, Serbia, Albania, and Taiwan to 6.4% for Iceland). ASR problem scale scores were positively skewed in every sample, because many people in nonclinical samples report relatively few problems. However, general linear models are very robust with respect to deviations from normality, especially with very stringent criteria for significance and large samples having similar skew (Kirk, 1995). Accordingly, we analyzed untransformed raw scores for all analyses, including analyses of variance (ANOVAs) on scale scores and correlations on 0–1–2 item ratings. Because of the high statistical power, we set alpha at $p \leq .001$ for all analyses.

Because the 17 societies did not have equal sample sizes for all four age/gender groups (i.e., each gender aged 18–35 and 36–59), we calculated the societal means for each scale by averaging means for the four age/gender groups. We then averaged the 17 societal means for each scale to obtain the omnicultural mean (i.e., average of the 17 society means; Ellis & Kimmel, 1992) and its *SD*. However, ANOVAs were conducted with each participant in each society contributing equally to the results.

In our first set of analyses, we computed Cronbach's alphas to test for internal consistencies of the ASR's problem and Adaptive Functioning scales in each society (except for the Family scale, which does not have a uniform set of items for all cases). Next, we calculated mean item ratings for each of the 120 problem items across the full sample for each society (i.e., not separately for age/gender groups). This yielded 17 rank orderings of endorsement for 120 items (one per society), which served as our measure of rank ordering of problem frequency/severity in each society. We then computed correlations between the mean problem item ratings from each society and the mean item ratings from each other society. These *r*s were in effect *Q* correlations between the mean of the 0–1–2 ratings for each item in each pair of societies. (Note that the mean item ratings were distributed continuously from a possible minimum of 0 to a possible maximum of 2). Next, we converted each society's 16 bisociety *r*s to Fisher's *z*s, averaged these 16 *z*s for each society to obtain 17 mean bisociety *z*s, and then averaged these 17 *z*s to yield the omnicultural mean *z*. We reconverted these *z*s back to *r*s for purposes of reporting results, so that the values would be interpretable on the *r* scale (-1 to $+1$). As an additional test of concordance, we computed Kendall's *W* for mean item ratings after assigning ranks to all items in each society based on the mean item ratings for that society.

In our third set of analyses, we tested ASR problem scale scores, Personal Strength scale scores, and Adaptive Functioning scale scores using society, gender, and age (18–35, 36–59) as factors in ANOVAs for each scale. Because we used $p \leq .001$ for all analyses, we report ESs rather than *F* and *p* values. ESs for ANOVAs

were measured by η^2 and characterized using Cohen's (1988) criteria (*small* = .01 to .059, *medium* = .06 to .139, *large* ≥ .14).

In our fourth set of analyses, we used multilevel modeling to test effects of individual differences, societies, and culture clusters on several scales using the Global Leadership and Organizational Behavior Effectiveness (GLOBE) definitions of culture clusters (House, Hanges, Javidan, Dorfman, & Gupta, 2004). Additionally, we clustered the societies by per capita income and then performed the same HLM analyses.

Results

Internal Consistency of Scales in Different Societies

Alpha coefficients are displayed in Table 2. For each of the 17 societies, alphas for Total Problems were ≥ .94, while alphas for Internalizing and Externalizing were ≥ .87 and ≥ .83, respectively. For eight scales, the lowest alpha was > .60 and for nine scales the lowest alpha was > .70. However, as shown in Table 2, alphas in some societies were < .60 for seven scales (Thought Problems, Rule-Breaking Behavior, *DSM*-Anxiety Problems, Obsessive-Compulsive Problems, Sluggish Cognitive Tempo, Jobs, and Education). Consequently, we do not report further findings for these scales.

Mean Item Ratings

All bisociety *r*s for the mean problem item ratings were significant, with the range being from .43 (Kenya with Iceland and Japan) to .89 (Argentina with Poland, Hong Kong with Taiwan). Because the lowest mean *r* was .56 (Kenya), 17 mean bisociety *r*s were large according to Cohen's (1988) criteria. The omnicultural mean *r* across all 17 societies was .77. This indicates that the societies were fairly similar with respect to mean item ratings. When bisociety *r*s were calculated across 18 societies, with the U.S. included, the omnicultural mean *r* was also .77. Results of an additional test of concordance yielded a Kendall's *W* of .78 for ranked mean item ratings, very close to the omnicultural mean *r* of .77.

The 17 mean item ratings for each of the 120 problem items were averaged to yield an omnicultural mean item rating. The 120 omnicultural mean item ratings ranged from .03 to 1.17 (*M* = .36, *SD* = .21). Two items had a mean rating of ≥1.0, 26 had a mean rating from .50 to .99, 52 had a mean rating from .25 to .49, 26 had a mean rating from .10 to .24, and 14 had a mean rating of < .10. The problem items with the highest omnicultural mean 0–1–2-ratings averaged across all 17 societies are shown in Table 3. These items included many problems from the Internalizing scale (e.g., worrying, being nervous or tense, lacking self-confidence, feeling overwhelmed by responsibilities), some from the Attention Problems scale (e.g., problems concentrating, daydreaming, having trouble planning for the future), and two from the Aggressive Behavior scale (arguing a lot and being stubborn, sullen, or irritable). The problem items with the lowest mean ratings are also shown in Table 3. These items included such problems as self-injury, seeing or hearing things that aren't there, using drugs, stealing, repeating acts over and over, and attacking people.

Problem Scale Scores

For each scale, Table 4 displays the smallest and the largest societal mean, the omnicultural mean (and its *SD*), and the range of the within-society *SD*s. For example, the societal means for Total Problems ranged from 34.1 (*SD* = 23.3) to 52.6 (*SD* = 25.5), the omnicultural mean was 42.7 (*SD* = 6.1), and the range of the within-society *SDs* was 20.6 to 31.6. Results were very similar when we included the U.S., with an omnicultural mean of 42.5 (*SD* = 6.0). Nine societies had mean Total Problems scores within one *SD* of the omnicultural mean (10 including the U.S.), whereas three societies had lower mean Total Problems scores (Iceland, Japan, and Taiwan) and five societies had higher mean Total Problems scores (Albania, Brazil, Kenya, Latvia, and Lithuania). To test whether the societies with high Internalizing scores also tended to have high Externalizing scores, we calculated across the 17 societies the correlation between Internalizing mean scores and Externalizing mean scores, which contain no shared items. Although the resulting *r* of .74 suggested that societies had a general tendency toward lower or higher scores, this correlation should be interpreted cautiously because the problem scores were positively skewed.

Table 2
Internal Consistency Alpha Coefficients for ASR Scales in 17 Societies

ASR Scale (number of items)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Broad-band scales																	
Total Problems (120)	.91	.87	.90	.89	.92	.93	.93	.90	.94	.90	.92	.90	.91	.93	.88	.91	.91
Internalizing (39)	.93	.87	.84	.90	.85	.91	.86	.84	.91	.87	.90	.85	.89	.89	.83	.90	.88
Externalizing (35)	.96	.94	.94	.95	.95	.97	.96	.94	.97	.96	.96	.94	.95	.95	.96	.96	.96
Syndromes																	
Anxious/Depressed (18)	.86	.80	.85	.85	.90	.90	.91	.87	.92	.81	.88	.86	.87	.90	.83	.86	.90
Withdrawn (9)	.72	.73	.73	.70	.77	.75	.73	.75	.85	.67	.79	.76	.77	.80	.67	.74	.73
Somatic Complaints (12)	.80	.76	.77	.73	.75	.82	.80	.73	.82	.84	.82	.74	.80	.82	.77	.82	.72
Thought Problems (10)	.51	.54	.64	.57	.60	.70	.74	.47	.69	.69	.53	.65	.64	.51	.71	.61	.52
Attention Problems (15)	.82	.78	.80	.79	.83	.83	.84	.78	.87	.78	.82	.80	.78	.78	.79	.84	.84
Rule-Breaking Behavior (14)	.80	.72	.66	.81	.69	.79	.70	.57	.78	.76	.74	.68	.72	.76	.59	.75	.64
Aggressive Behavior (15)	.89	.82	.78	.83	.80	.83	.82	.80	.86	.75	.86	.81	.82	.83	.80	.85	.84
Intrusive (6)	.74	.66	.69	.75	.68	.80	.67	.63	.75	.62	.69	.71	.69	.70	.61	.73	.71
DSM-oriented scales																	
Depressive Problems (14)	.83	.73	.76	.76	.82	.83	.84	.78	.87	.79	.82	.78	.78	.82	.75	.77	.82
Anxiety Problems (6)	.58	.62	.61	.64	.75	.75	.78	.71	.77	.60	.70	.63	.66	.69	.59	.65	.77
*S*omatic Problems (9)	.73	.73	.74	.69	.68	.78	.79	.67	.79	.78	.81	.69	.74	.78	.72	.77	.68
Avoidant Personality Problems (7)	.65	.67	.71	.71	.74	.74	.79	.76	.85	.63	.78	.70	.74	.79	.67	.70	.75
Attention Deficit Hyperactivity (13)	.81	.76	.77	.79	.80	.82	.83	.75	.84	.73	.81	.77	.72	.75	.78	.80	.83
Antisocial Personality Problems (20)	.86	.74	.63	.82	.69	.80	.72	.66	.83	.79	.80	.74	.78	.76	.79	.79	.70
Other Problem Scales																	
Obsessive-Compulsive Problems (8)	.54	.58	.63	.56	.71	.76	.76	.64	.74	.66	.66	.63	.57	.62	.61	.63	.71
Sluggish Cognitive Tempo (5)	.61	.58	.63	.55	.69	.73	.72	.62	.74	.59	.73	.62	.55	.62	.61	.66	.75
Strengths/Adaptive Functioning scales[a]																	
Personal Strengths (11)	.75	.65	.67	.66	.61	.78	.67	.60	.78	.78	.76	.65	.69	.74	.63	.67	.74
Friends (4)	.61	.75	.61	.68	.65	.71	.63	.64	.74		.67	.71	.72	.77	.70	.63	.69
Spouse/Partner (8)	.82	.75	.79	.73	.72	.78	.74	.73	.79		.78	.65	.75	.75	.65	.76	.67
Job (9)	.49	.48	.50	.56	.57	.66	.57	.50	.72		.66	.63	.56	.66	.35	.63	.58
Education (5)	.53	.59	.51	.47	.67	.58	.67	.66	.50		.70	.58	.63	.52	.58	.73	.73

Note. ASR = Adult Self-Report.1 = Albania; 2 = Argentina; 3 = Brazil; 4 = Czech Republic; 5 = Flanders; 6 = Hong Kong; 7 = Iceland; 8 = Italy; 9 = Japan; 10 = Kenya; 11 = Korea; 12 = Latvia; 13 = Lithuania; 14 = Poland; 15 = Portugal; 16 = Serbia; 17 = Taiwan.
[a] Alphas for the Adaptive Functioning scales based on 16 societies (Kenya omitted Adaptive Functioning items); no alpha is computed for the Family scale because the same items are not rated by all individuals.

Table 3
ASR Items With the 21 Highest and 21 Lowest Omnicultural Mean Ratings Across 17 Societies

Items with highest omnicultural mean ratings[a]	Mean (*SD*)	Items with lowest omnicultural mean ratings	Mean (*SD*)
72. I worry about my family	1.17 (.44)	85. I have thoughts that other people would think are strange	.16 (.10)
22. I worry about my future	1.17 (.31)	56d. Problems with eyes	.16 (.10)
112. I worry a lot	.82 (.24)	84. I do things that other people think are strange	.14 (.09)
32. I feel that I have to be perfect	.75 (.28)	122. I have trouble keeping a job	.14 (.07)
69. I am secretive or keep things to myself	.72 (.24)	110. I wish I were of the opposite sex	.14 (.08)
1. I am too forgetful	.72 (.10)	66. I repeat certain acts over and over	.13 (.08)
99. I dislike staying in one place for very long	.70 (.21)	101. I stay away from my job even when not sick/not on vacation	.11 (.07)
45. I am nervous or tense	.68 (.18)	79. I have a speech problem	.09 (.04)
3. I argue a lot	.68 (.20)	56g. Vomiting, throwing up	.09 (.05)
44. I feel overwhelmed by my responsibilities	.63 (.20)	37. I get in many fights	.09 (.07)
17. I day dream a lot	.62 (.25)	20. I damage or destroy my things	.08 (.05)
47. I lack self-confidence	.62 (.13)	92. I do things that may cause me trouble with the law	.07 (.04)
24. I don't eat as well as I should	.61 (.28)	91. I think about killing myself	.07 (.04)
8. I have trouble concentrating/paying attention for long	.60 (.11)	97. I threaten to hurt people	.06 (.05)
86. I am stubborn, sullen, or irritable	.59 (.15)	18. I deliberately try to hurt or kill myself	.06 (.04)
93. I talk too much	.57 (.14)	70. I see things that other people think aren't there	.06 (.06)
78. I have trouble making decisions	.57 (.10)	40. I hear sounds or voices that other people think aren't there	.05 (.04)
42. I would rather be alone than with others	.57 (.14)	21. I damage or destroy things belonging to others	.05 (/03)
118. I am too impatient	.56 (.11)	6. I use drugs for nonmedical purposes	.04 (.03)
53. I have trouble planning for the future	.56 (.11)	57. I physically attack people	.04 (.03)
71. I am self-conscious or easily embarrassed	.56 (.13)	82. I steal	.03 (.02)

Note. ASR = Adult Self-Report. The omnicultural mean rating for each item was obtained by averaging the 17 mean item ratings obtained from the full sample in each society.
[a] Items are listed in descending order of mean scores.

As shown in Table 5, 13 of the 14 problem scale societal ESs (η^2) were small (ranging from 2% to 5%), and one was medium (6%). Gender effects were significant for 12 scales, with the largest ES being 3% (Somatic Complaints). Women scored higher than men on 9 scales, whereas men scored higher on Externalizing, Intrusive, and *DSM*-Antisocial Personality Problems. Younger adults (ages 18–35) scored significantly higher than older adults (ages 36–59) on 11 scales (ESs $\leq$2%), whereas older adults scored higher than younger adults on one scale (*DSM*-Somatic Problems; ES $<$1%). Of the many interactions, about half were not significant, while ESs for the rest were $<$1% (not displayed in Table 5), indicating strong similarity in age and gender patterns across societies. Very similar results were obtained when the U.S. was included, as shown in Table 5.

HLM Analyses of Culture Clusters and per Capita Income

GLOBE culture clusters. Georgas and Berry (1995) developed a taxonomy of societies based on six parameters (e.g., ecology, education, economics). As their taxonomy does not reflect the current geopolitical or economic status of many of the societies included in our

Table 4
Range of Society Means, Omnicultural Means and Their SDs, and Ranges of 17 Society SDs (N = 10,197)

ASR Scale	Minimum mean	Maximum mean	Omnicultural mean	Omnicultural mean *SD*	Within-society *SD* range
Broad-band scales					
Total Problems	34.1	52.6	42.7	6.1	20.6–31.6
Internalizing	11.4	18.0	14.6	2.1	8.2–12.0
Externalizing	8.0	14.3	10.5	1.8	5.9–10.1
Syndromes					
Anxious/Depressed	6.5	10.3	8.0	1.1	4.6–6.6
Withdrawn	2.2	4.8	3.3	.8	2.3–3.5
Somatic Complaints	2.6	5.2	3.3	.7	2.6–4.8
Attention Problems	5.6	8.3	6.6	.7	3.9–5.0
Aggressive Behavior	4.0	7.8	5.3	1.0	3.6–5.4
Intrusive	1.7	3.9	2.4	.6	1.8–2.5
DSM-oriented scales					
Depressive Problems	3.9	6.1	4.9	.6	3.4–407
Somatic Problems	1.4	4.0	2.2	.6	2.0–3–7
Avoidant Personality Problems	2.3	3.9	3.0	.4	2.1–3.1
Attention Deficit Hyperactivity Problems	4.4	6.9	5.7	.7	3.4–4.6
Antisocial Personality Problems	2.9	6.1	4.2	1.0	2.5–5.0
Strengths/Adaptive Functioning scales					
Personal Strengths	9.1	17.6	15.5	2.1	2.6–4.1
Friends	6.3	9.0	8.2	.7	1.9.–3.0
Spouse/Partner	3.0	5.6	4.3	.8	2.2–3.3
Family	1.3	1.8	1.5	.1	.31–.52

Note. Omnicultural *M* = mean of the 17 society means. Friends, Spouse/Partner, and Family means are for only 16 societies because Kenya omitted Adaptive Functioning items.

study, the GLOBE culture clusters (House et al., 2004) offered a more contemporary taxonomy with which we could test the degree to which cultural influences transcending specific societies might have affected our results. The GLOBE project included >200 scholars from 69 countries. Building on work by previous theorists, such as Hofstede (1980) and Triandis (1995), House et al. formulated nine GLOBE dimensions, including 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 countries, they were able to classify countries into 10 culture clusters based on these dimensions.

We used HLM to determine the percentage of variance accounted for by individual differences, society, and GLOBE culture cluster. Building on work by Stankov (2011), we used HLM to parse the variance of Total Problems, Internalizing, and Externalizing into components reflecting individual differences (Level 1), society (Level 2), and culture cluster (Level 3). Only those GLOBE clusters for which we had data from more than one society were used, which excluded the U.S. sample (as the only Anglo culture), the Icelandic sample (Nordic), the Flanders sample (Germanic European), and the Kenya sample (African). The remaining societies represented the following four culture clusters: Eastern European (Czech Republic, Latvia, Lithuania, Poland, Serbia, Albania); Latin American (Argentina, Brazil); Latin European (Italy, Portugal); and Confucian Asian (Hong Kong, Japan, Korea, Taiwan). We report findings only with $p < .001$ as significant.

In the three separate HLMs for Total Problems, Internalizing, and Externalizing, the addition of society accounted for 3% to 5% of the total variance, all significant but small effects. Addition of the third level of culture cluster accounted for 0.02% to 0.3% of the total variance, which was not significant. Individual differences within societies accounted for 95% to 97% of the total variance, all $p < .001$. An

Table 5
Significant Effect Sizes (η^2) for Society, Gender, and Age on ASR Scale Scores in 17 Societies/18 Societies

Scale	Society	Gender	Age
Broad-band scales			
Total Problems	5%/4%	<1%/<1%[a]	1%/1%[c]
Internalizing	3%/4%	2%/2%[a]	<1%/<1%[c]
Externalizing	4%/4%	<1%/2%[b]	2%/2%[c]
Syndromes			
Anxious/Depressed	4%/4%	2%/2%[a]	<1%/<1%[c]
Withdrawn/Depressed	5%/5%	*ns*/*ns*	*ns*/*ns*
Somatic Complaints	2%/2%	3%/3%[a]	*ns*/*ns*
Attention Problems	2%/2%	<1%/<1%[a]	2%/1%[c]
Aggressive Behavior	5%/4%	<1%/<1%[a]	1%/1%[c]
Intrusive	6%/5%	<1%/<1%[b]	2%/2%[c]
DSM-oriented scales			
Depressive Problems	2%/2%	2%/2%[a]	<1%/<1%[c]
Somatic Problems	3%/3%	2%/2%[a]	<1%/<1%[d]
Avoidant Personality Problems	2%/2%	<1%/<1%[a]	<1%/<1%[c]
Attention Deficit Hyperactivity Problems	3%/3%	*ns*/*ns*	2%/1%[c]
Antisocial Personality Problems	4%/4%	1%/1%[b]	1%/1%[c]
Strengths/Adaptive Functioning scales			
Personal Strengths	34%/31%	*ns*/*ns*	*ns*/*ns*
Friends	8%	<1%[a]	3%[c]
Spouse/Partner	5%	<1%[b]	<1%[c]
Family	7%	<1%[a]	*ns*/*ns*

Note. ASR = Adult Self-Report; 17 societies (N = 10,197); 18 societies include U.S. (N = 12,217); Friends, Spouse/Partner, and Family ESs for 16 societies only (not Kenya); *ns* = not significant at $p < .001$. Interactions were not significant or had an ES of < 1%.
[a] Females > males. [b] Males > females. [c] Younger adults > older adults. [d] Older adults > younger adults.

additional HLM was performed for the Personal Strengths scale, which had the largest ES for society in our ANOVAs. Compared with the three broad-band scales, larger percentages of variance were accounted for by society (14%) and by culture cluster (12%), with concomitantly less variance accounted for by individual differences (74%).

Per capita income clusters. We used the same HLM methodology to test the effects of per capita income clusters on ASR scores. Specifically, we clustered our 17 societies by the World Bank's (2015) estimates of purchasing power parity per capita (PPP); because the World Bank did not report PPP for Latvia or Argentina, International Monetary Fund estimates for these two societies were used. We grouped societies into PPP categorical levels of 10–20K (Brazil, Serbia, Albania), 20–35K (Argentina, Czech Republic, Italy, Korea, Latvia, Lithuania, Poland, Portugal), 35–50K (Flanders, Iceland, Japan, Taiwan), and over 50K (Hong Kong, U.S.A.). (K = 1,000 in "international" dollars.) PPP cluster was used for the Level 3 variable. In this model, PPP accounted for ≤1% of the variance and society accounted for 4–5% of the variance, indicating that most of the variation in Total Problems, Internalizing, and Externalizing (94–95%) was due to individual differences within societies (all significant effects). Society accounted for 27% of the variance for Personal Strengths, PPP cluster was not significant (<1%, $p > .500$), and 73% of the variance was accounted for by individual differences.

Personal Strengths Scale

The societal ES of 34% for Personal Strengths was much larger than those found for the problem scales. Also, unlike for most problem scales, age and gender effects were not significant. The four Asian societies (Japan, Korea, Hong Kong, and Taiwan) had the lowest mean Personal Strength scores, with the means for Japan and Korea being > 1 *SD* (2.1) below the omnicultural mean of 15.5: Japan's M = 9.1 (SD = 4.1) and Korea's M = 13.0 (SD = 4.0).

Societies with the highest mean Personal Strengths scores included Brazil, Albania, Portugal, and Flanders.

Educational Level

The ASR asks respondents to indicate their highest level of education. Educational levels were coded differently for different societies reflecting differences in educational systems, but most had at least five levels (e.g., primary, secondary, some postsecondary, college/university, some postcollege). Correlations between educational level and Total Problems score computed separately within each society yielded nonsignificant *r*s for eight societies (−.07 to .09 for Flanders, Hong Kong, Italy, Japan, Korea, Poland, Portugal, and Taiwan). In the other nine societies, *r*s were significant at $p < .01$ or $p < .001$, negative, and very small (from −.10 in Albania to −.17 in Latvia and Lithuania). These results indicate minimal associations between self-reported problem scores and education.

Adaptive Functioning Scales

Friends scale. Most participants in the 16 non-U.S. samples with Adaptive Functioning scores (not Kenya) completed the Friends scale (N = 9,665). The societal ES was 8%, with Japan having the lowest mean score and Portugal and Iceland having the highest mean scores. Women obtained slightly higher scores than men (ES < 1%), and younger adults obtained higher scores than older adults (ES = 3%). All interactions were < 1% or not significant.

Spouse/Partner scale. Only 6,093 participants completed this scale, indicating that 40% of the sample had not lived with a spouse/partner in the past 6 months. The societal ES was 5%, with Hong Kong and Korea having the lowest means and Italy and Portugal having the highest. Gender and age ESs were both < 1%, with males and younger adults having slightly higher scores than females and older adults.

Family scale. This scale was completed by 9,651 participants. The societal ES was 7%, with Poland and Latvia having the lowest means and Iceland and Albania having the highest means. Women scored slightly higher than men (ES < 1%), but the age effect was not significant.

Associations with problem scores. The Friends scale had significant but modest *r*s with Internalizing (−.17) and Total Problems (−.07). The Spouse/Partner scale was more strongly correlated with the three broad-band problem scales, with *r*s of −.39 (Internalizing), −.33 (Externalizing), and −.38 (Total Problems). The *r*s between the Family scale and the three broad-band problem scales were all significant but small (−.15, −.17, −.17).

Discussion

Alphas obtained in this study were very similar to those reported by Achenbach and Rescorla (2003). As in the U.S. sample, the largest alphas were for Total Problems, Internalizing, and Externalizing, and the smallest were for Thought Problems. All alphas were > .70 for nine scales and > .60 for eight other scales.

Results indicated strong similarities across 17 non-U.S. societies regarding rank ordering of problem items by mean item ratings (r = .77). This large r approximates the omnicultural mean r of .70 for adolescent self-ratings in 34 societies reported by Rescorla et al. (2012) based on data obtained for the Youth Self-Report (YSR; Achenbach & Rescorla, 2001). Serious problems such as self-injury, seeing or hearing things that aren't there, using drugs, stealing, destroying things, getting in fights, trouble with the law, and attacking people were rarely endorsed. The items with the highest omnicultural mean ratings (e.g., worrying, being nervous or tense, lacking self-confidence, difficulties concentrating, feeling overwhelmed by responsibilities, having trouble making decisions, arguing, and being irritable), when considered individually, are not necessarily symptoms of serious psychopathology. However, individuals reporting multiple problems of this sort may need professional help.

Mean Total Problems scores for the 17 societies (18 including the U.S.) ranged from 34.1 to 52.6, thereby spanning less than 8% of the possible range from 0 to 240. Although nine quite different societies had mean Total Problems scores within one *SD* (6.1) of the omnicultural mean of 42.7, three societies had mean Total Problems scores > 1 *SD* below the omnicultural mean while five had mean Total Problems scores > 1 *SD* above the omnicultural mean.

Our HLM analyses indicated that GLOBE culture clusters and society accounted for much smaller percentages of variance in Internalizing, Externalizing, and Total Problems scores than were accounted for by individual differences within societies. Our results are consistent with Stankov's (2011) findings for self-ratings of neuroticism by students in 45 societies, namely, that individual differences accounted for 95.3% of the variance in self-ratings, whereas societal differences accounted for only 2.0% and GLOBE culture cluster differences accounted for only 2.7%. Similar results were obtained from our HLM analyses using economic status as the clustering factor.

No obvious "cultural dichotomy" such as Eastern/Western, collectivistic/individualistic, or developing/developed appears to explain why some societies tended to have lower versus higher mean problem scores. For example among Confucian Asian societies, Japan and Taiwan had relatively low problem scores but Korea and Hong Kong did not. Similarly, among former "East Bloc" societies, Lithuania, Latvia, and Albania had relatively high problem scores, but Poland, the Czech Republic, and Serbia did not. It is possible that response sets (e.g., a tendency toward social desirability responding) or response styles (e.g., avoiding extreme ratings) may have contributed to differences in overall score levels, but we did not have any measures of these factors. Consistent with our previous multicultural comparisons (Rescorla et al., 2007, 2012), ASR scores varied widely within each society (see Table 4). Pinpointing the reasons for the large individual differences in scores within societies was beyond the scope of our study, but socioeconomic status, ethnic, religious, education, political, cultural, familial, and genetic differences might be relevant.

Age and gender effects were found on most problem scales, but societal interactions with age and gender were either not significant or very small, indicating strong similarity across societies in age and gender patterns. Consistent with our YSR findings for adolescents (Rescorla et al., 2007), women obtained higher scores than men on most problem scales, but men obtained higher scores on Externalizing, Intrusive, and *DSM*-Antisocial Personality Problems. Higher scores were obtained by 18–35-year-olds than by 36–59-year-olds on most problem scales.

Whereas most of the societal ESs on problem scales were small (2% to 5%), the societal ES on the Personal Strengths scale was 34%, mainly because within-society variance was very small. Rescorla et al. (2007) found the same pattern for the YSR, where the societal ES was 27% for the Positive Qualities scale. HLM results for Personal Strengths using the GLOBE clusters indicated larger effects for society and culture cluster (14% and 12%) than were found for the three broad-band problem scales. In the HLM analysis of economic clusters, the societal effect was much larger than the economic cluster effect (27% vs. $< 1\%$). The lowest mean scores on Personal Strengths were found in the four Confucian Asian societies, with Japan and Korea having the lowest scores. Overall, these findings suggest that self-rated positive qualities may be more "culturally embedded" than are most self-rated problems. Furthermore, the consistency of these findings for both adolescents and adults argues for emic investigation of self-reports of Personal Strengths in Confucian societies, especially Japan and Korea.

Limitations

A limitation of our study is that some of our samples were of unknown representativeness because they were obtained using convenience rather than random sampling methods. Other limitations include low response rates plus low internal consistencies for some ASR scales in some societies. Furthermore, reports of problems and adaptive functioning by collaterals might yield different results than self-reports, a possibility that we are currently testing. Because ours was an etic study, the same instrument was used in every society. Inclusion of other items might yield different results. It should also be noted that our societal differences in problem scores may, to some degree, reflect societal differences in response styles or differential willingness to admit problems. Finally, our ANOVA results for syndrome scores must be considered in light of the fact that, although Ivanova et al.'s (2015) within-society CFAs supported the eight-syndrome model in the 17 societies whose data we analyzed, they could not test for various kinds

of invariance across all societies in a single multigroup CFA.

Conclusions

Despite these limitations, our study has numerous strengths. Sample sizes were large and the 17 societies (18 including the U.S.) differed in economic, political, religious, and ethnic characteristics. Because data for all the societies were combined in a single data set, we could do direct statistical tests of similarities and differences across societies. Strong consistency across societies was found in correlations between mean item ratings and in age and gender patterns, but significant differences between societies were found in scale scores, with a particularly large ES for the Personal Strengths scale. Although we had no measures of response style (e.g., social desirability), it is possible that unmeasured response style differences affected societal differences in scores.

Research and Clinical Applications

Our study indicates that epidemiological data on mental health problems can be obtained for large samples at relatively low cost when indigenous investigators use a standardized assessment instrument that does not require professional time for administration or scoring. Data were easily obtained in a diverse set of societies (e.g., Albania, Brazil, Iceland, Kenya, Latvia, and Serbia), including some with turbulent recent histories. Because the same assessment instrument was used in all societies (following a rigorous translation and back-translation process), and the data comprised quantitative ratings, they could be easily merged to enable international comparisons. ASR items appeared to be interpreted quite similarly in the 17 different societies, as reflected in the omnicultural mean *r* of .77 for mean item ratings. The list of items with the highest and lowest mean ratings should be useful for researchers who survey mental health problems in other societies.

Our findings also have important implications for clinical practice. Because the ASR assesses problems, substance use, personal strengths, and adaptive functioning, it is an efficient way to tap multiple issues related to mental health. Although societal ESs were modest for most problem and adaptive functioning scales, the ANOVA ES was 34% for Personal Strengths. This finding and the 14% of variance accounted for by society and 12% by culture cluster in our GLOBE HLM analysis indicate that cultural factors may have stronger effects on adults' self-reported positive qualities than on their reports of problems, an important consideration for clinicians treating clients from different cultural backgrounds.

References

Achenbach, T. M. (1997). *Manual for the Young-Adult Self-Report and Young Adult Behavior Checklist*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

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

Achenbach, T. M., & Rescorla, L. A. (2003). *Manual for the ASEBA adult forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Achenbach, T. M., & Rescorla, L. A. (2015). *Multicultural supplement to the Manual for the ASEBA adult forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Au, A., & Leung, P. (2012). [Hong Kong ASR data]. Unpublished raw data.

Bellina, M. (2012). [Italian ASR data]. Unpublished raw data.

Butcher, J. N., & Han, K. (1996). Methods of establishing cross-cultural equivalence. In J. N. Butcher (Ed.), *International adaptations of the MMPI-2: Research and clinical applications* (pp. 44–63). Minneapolis, MN: University of Minnesota Press.

Byrne, B. M., & Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology, 30,* 555–574. http://dx.doi.org/10.1177/0022022199030005001

Caldas, J. C. (2012). [Portuguese ASR data]. Unpublished raw data.

Chen, Y. C. (2012). [Taiwanese ASR data]. Unpublished raw data.

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

Csemy, L. (2012). [Czech ASR data]. Unpublished raw data.

Decoster, J., & Fontaine, J. R. J. (2012). [Flemish ASR data]. Unpublished raw data.

Ellis, B. B., & Kimmel, D. D. (1992). Identification of unique cultural response patterns by means of item response theory. *Journal of Applied Psychology, 77,* 177–184. http://dx.doi.org/10.1037/0021-9010.77.2.177

Funabiki, Y. (2012). [Japanese ASR data]. Unpublished raw data.

Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6,* 304–312. http://dx.doi.org/10.1037/1040-3590.6.4.304

Georgas, J., & Berry, J. W. (1995). An ecocultural taxonomy for cross-cultural psychology. *Cross-Cultural Research: The Journal of Comparative Social Science, 29,* 121–157. http://dx.doi.org/10.1177/106939719502900202

Guðmundsson, H. S., & Árnadóttir, H. A. (2012). [Icelandic ASR data]. Unpublished raw data.

Harder, V. S., & Ndetei, D. M. (2012). [Kenyan ASR data]. Unpublished raw data.

Hofstede, G. (1980). *Culture's consequences*. Beverly Hills, CA: Sage Publications.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE Study of 62 societies*. Thousand Oaks, CA: Sage.

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Turner, L. V., Ahmeti-Pronaj, A., Au, A., . . . Zasepa, E. (2015). Syndromes of self-reported psychopathology for ages 18–59 in 29 societies. *Journal of Psychopathology and Behavioral Assessment, 37,* 171–183. http://dx.doi.org/10.1007/s10862-014-9448-8

Kim, M. Y., Kim, Y. A., Lee, J., Kim, H. J., & Oh, K. J. (2014). Development of Korean version of the Adult Self Report. *The Korean Journal of Clinical Psychology, 33,* 615–632.

Kirk, R. R. (1995). *Experimental design procedures for the behavioral sciences*. Pacific Grove, CA: Brookside, Cole.

Leung, K., & Bond, M. H. (1989). On the empirical identification of dimensions for cross-cultural comparisons. *Journal of Cross-Cultural Psychology, 20,* 133–151. http://dx.doi.org/10.1177/0022022189202002

Markovic, J. (2012). [Serbian ASR data]. Unpublished raw data.

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

Pike, K. L. (1967). *Language in relation to a unified theory of the structure of human behavior* (2nd ed.). The Hague, Netherlands: Mouton De Gruyter. http://dx.doi.org/10.1037/14786-000

Poortinga, Y. H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology, 24,* 737–756. http://dx.doi.org/10.1080/00207598908247842

Rescorla, L., Achenbach, T. M., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., . . . Verhulst, F. (2007). Epidemiological comparisons of problems and positive qualities reported by adolescents in 24 countries. *Journal of Consulting and Clinical Psychology, 75,* 351–358. http://dx.doi.org/10.1037/0022-006X.75.2.351

Rescorla, L., Ivanova, M. Y., Achenbach, T. M., Begovac, I., Chahed, M., Drugli, M. B., . . . Zhang, E. Y. (2012). International epidemiology of child and adolescent psychopathology II. Integration and applications of dimensional findings from 44 societies. *Journal of the American Academy of Child & Adolescent Psychiatry, 51,* 1273–1283.e8.

Samaniego, V. C., & Vázquez, N. (2012, July). *Adult psychopathology: Is there any agreement between self-reports and reports by other informants?* Paper presented at the 30th International Congress of Psychology, Cape Town, South Africa.

Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality, 38,* 230–255. http://dx.doi.org/10.1016/S0092-6566(03)00069-2

Sebre, S. (2012). [Latvian ASR data]. Unpublished raw data.

Silvares, E., & da Rocha, M. M. (2012). [Brazilian ASR data]. Unpublished raw data.

Simulionienė, R., Brazdeikienė, L., Rugevičius, M., Gedutienė, R., & Žakaitienė, A. (2010). The psychometric properties of the Lithuanian version of ASEBA adult forms. *Psichologija, 42,* 23–43.

Sokoli, E. (2013). [Albanian ASR data]. Unpublished raw data.

Stankov, L. (2011). Individual, country and societal cluster differences on measures of personality, attitudes, values, and social norms. *Learning and Individual Differences, 21,* 55–66. http://dx.doi.org/10.1016/j.lindif.2010.09.002

Tomlinson, M., Rudan, I., Saxena, S., Swartz, L., Tsai, A. C., & Patel, V. (2009). Setting priorities for global mental health research. *Bulletin of World Health Organization, 87,* 438–446. http://dx.doi.org/10.2471/BLT.08.054353

Triandis, H. C. (1995). *Individualism & collectivism*. Boulder, CO: Westview press.

van de Vijver, F. J., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment, 13,* 29–37. http://dx.doi.org/10.1027/1015-5759.13.1.29

World Bank. (2015). *Purchasing power parities and the real size of world economies*. Washington, DC: World Bank.

World Health Organization World Mental Health Survey Consortium. (2004). Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization world mental health surveys. *Journal of the American Medical Association, 291,* 2581–2590. http://dx.doi.org/10.1001/jama.291.21.2581

Zasepa, E. (2012). [Polish ASR data]. Unpublished raw data.

Received April 26, 2014
Revision received October 9, 2015
Accepted October 13, 2015 ■