



International Journal of Clinical and Health Psychology

www.elsevier.es/ijchp



ORIGINAL ARTICLE

Syndromes of collateral-reported psychopathology for ages 18-59 in 18 Societies



Masha Y. Ivanova^{a,*}, Thomas M. Achenbach^a, Leslie A. Rescorla^b, Lori V. Turner^a, Hervör Alma Árnadóttir^c, Alma Au^d, J. Carlos Caldas^e, Nebia Chaalal^f, Yi Chuen Chen^g, Marina M. da Rocha^h, Jeroen Decosterⁱ, Johnny R.J. Fontaineⁱ, Yasuko Funabiki^j, Halldór S. Guðmundsson^c, Young Ah Kim^k, Patrick Leung^l, Jianghong Liu^m, Sergey Malykhⁿ, Jasminka Marković^o, Kyung Ja Oh^p, Jean-Michel Petot^f, Virginia C. Samaniego^q, Edwiges Ferreira de Mattos Silveiras^h, Roma Šimulionienė^r, Valentina Šobot^o, Elvira Sokoli^s, Guiju Sun^t, Joel B. Talcott^u, Natalia Vázquez^q, Ewa Zasepa^v

^a University of Vermont, USA

^b Bryn Mawr College, USA

^c University of Iceland, Iceland

^d Hong Kong Polytechnic University, China

^e Instituto Superior de Ciências da Saúde - Norte, Portugal

^f Université de Paris Ouest, France

^g National Chung Cheng University, Taiwan

^h University Paulista (Unip), Brazil

ⁱ Ghent University, Belgium

^j Kyoto University Hospital, Japan

^k Huno Inc., Korea

^l The Chinese University of Hong Kong, China

^m University of Pennsylvania, USA

ⁿ Psychological Institute of the Russian Academy of Education, Russia

^o University of Novi Sad, Serbia

^p Yonsei University, South Korea

^q Pontificia Universidad Católica Argentina, Argentina

^r Klaipeda University, Lithuania

^s University of Tirana, Albania

^t Southeast University, China

^u Aston University, United Kingdom

^v The Maria Grzegorzewska Academy of Special Education, Poland

Received 12 April 2014; accepted 17 July 2014

Available online 29 August 2014

* Corresponding author: 1 South Prospect Street, Burlington, VT 05401, USA.
E-mail address: masha.ivanova@uvm.edu (M.Y. Ivanova).

KEYWORDS

Collateral reports;
Multicultural;
International;
Adult Behavior
Checklist;
Descriptive survey
study

PALABRAS CLAVE

Informes de
allegados;
Multicultural;
Internacional;
Adult Behavior
Checklist;
Estudio descriptivo
mediante encuestas

Abstract The purpose was to advance research and clinical methodology for assessing psychopathology by testing the international generalizability of an 8-syndrome model derived from collateral ratings of adult behavioral, emotional, social, and thought problems. Collateral informants rated 8,582 18-59-year-old residents of 18 societies on the Adult Behavior Checklist (ABCL). Confirmatory factor analyses tested the fit of the 8-syndrome model to ratings from each society. The primary model fit index (Root Mean Square Error of Approximation) showed good model fit for all societies, while secondary indices (Tucker Lewis Index, Comparative Fit Index) showed acceptable to good fit for 17 societies. Factor loadings were robust across societies and items. Of the 5,007 estimated parameters, 4 (0.08%) were outside the admissible parameter space, but 95% confidence intervals included the admissible space, indicating that the 4 deviant parameters could be due to sampling fluctuations. The findings are consistent with previous evidence for the generalizability of the 8-syndrome model in self-ratings from 29 societies, and support the 8-syndrome model for operationalizing phenotypes of adult psychopathology from multi-informant ratings in diverse societies.

© 2014 Asociación Española de Psicología Conductual. Published by Elsevier España, S.L.U. All rights reserved.

Síndromes psicopatológicos informados por personas allegadas a adultos de 18 a 59 años de edad en 18 países

Resumen El propósito fue avanzar en la metodología clínica y de investigación de la evaluación psicopatológica mediante el examen de la generalización internacional de un modelo de 8 síndromes derivados de evaluaciones de personas allegadas a adultos, en sus problemas emocionales, sociales y de pensamiento. Informantes allegados a los adultos calificaron a 8.582 residentes de 18 países entre 18 y 59 años de edad con el Adult Behavior Checklist (ABCL). Un análisis factorial confirmatorio examinó el ajuste del modelo de 8 síndromes a las puntuaciones provenientes de cada país. El índice primario de ajuste del modelo (RMSEA) mostró un buen ajuste del modelo para todas las sociedades, mientras que índices secundarios (TLI, CFI) mostraron un ajuste de aceptable a bueno para 17 países. Las cargas factoriales fueron robustas a través de los países e ítems. Los hallazgos son consistentes con evidencia previa existente para la generalización del modelo de 8 síndromes en autoinformes de 29 sociedades. Además, los resultados respaldan el modelo de 8 síndromes para operacionalizar fenotipos de psicopatología del adulto provenientes de evaluaciones de múltiples informantes en diversas sociedades.

© 2014 Asociación Española de Psicología Conductual. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Globalization is a theme of our times. World communities are becoming increasingly diverse as traditional boundaries between societies are changing. Mental health professionals increasingly need clinical constructs and methods for operationalizing such constructs that can be used with people of many different backgrounds.

One approach is to define constructs on the basis of experts' knowledge and judgments, as exemplified by the [American Psychiatric Association's \(2013\) Diagnostic and Statistical Manual \(DSM\)](#) and the [World Health Organization's \(2010\) International Classification of Disease \(ICD\)](#). A second approach derives constructs from statistical analyses of problems reported for samples of individuals. This approach can be applied to various kinds of data, including experts' judgments of individuals' problems. An early example is [Wittenborn's \(1951\)](#) factor analysis of

psychiatrists' and nurses' ratings of U.S. inpatients on 55 symptoms, which yielded nine syndromes resembling diagnostic constructs of that era.

The statistical approach can also be applied to self-reported problems. As an example, [Goekoop et al. \(1992\)](#) derived five factors from exploratory factor analysis (EFA) of 192 Dutch outpatients' interview reports regarding 65 items of the Comprehensive Psychopathological Rating Scale ([Asberg, Montgomery, Perris, Schalling, & Sedvall, 1978](#)). In a study of 91 UK Tourette syndrome patients, EFA of 11 self-rated personality and psychopathology measures yielded obsessiveness and anxiety/depression factors ([Eapen, Fox-Hiley, Banerjee, & Robertson, 2004](#)).

The statistical approach has been used to derive syndrome constructs from a broad spectrum of problems assessed by the Adult Self-Report (ASR), which is designed

for clinical, research, and epidemiological assessment of 18-59-year-olds (Achenbach & Rescorla, 2003). Literate respondents can rate its 120 problem items in about 15 minutes, or it can be administered by lay interviewers. The items describe particular kinds of behavioral, emotional, social, and thought problems, which are rated 0 = *not true*, 1 = *somewhat or sometimes true*, or 2 = *very true or often true*, based on the preceding 6 months. Every item had to meet at least one of the following criteria: (a) it discriminated significantly between individuals referred for mental health or substance use services versus demographically similar individuals not referred for services; (b) it was identified by international experts as being very consistent with ≥ 1 DSM diagnostic category; (c) it loaded significantly on at least one of eight syndromes derived by EFAs and confirmatory factor analyses (CFAs) of self-ratings by 2,968 U.S. referred and nonreferred 18-59-year-olds (Achenbach & Rescorla, 2003). The syndromes are designated as Anxious/Depressed, Withdrawn, Somatic Complaints, Thought Problems, Attention Problems, Aggressive Behavior, Rule-Breaking Behavior, and Intrusive. The first three syndromes aggregate into a broad-band Internalizing scale, while the last three aggregate into an Externalizing scale, based on second-order factor analyses.

The syndrome scales scored from the ASR have been found to correlate significantly with scales of the Symptom Checklist-90-Revised (Derogatis, 1994) in a U.S. clinical sample. Most of the syndrome scales have also been found to correlate significantly with MMPI scales in a Turkish non-clinical sample (Achenbach & Rescorla, 2003). Over 100 published studies have reported use of the ASR (Bérubé & Achenbach, 2014).

Beyond self-reports

Assessment of adult psychopathology tends to rely on data obtained mainly from the person being assessed. However, reports by other informants may not agree with self-reports. For example, Meyer et al. (2001) found a mean kappa of only .18 between diagnoses based solely on adults' self-reports versus diagnoses based on collateral reports. Meyer et al. also found that 70% of personality disorder diagnoses based on clinical interviews with adult patients were wrong when compared with diagnoses made from multiple sources of data. Furthermore, meta-analyses have yielded mean correlations between adult self- and informant-ratings of only .43 for internalizing problems (e.g., anxiety, depression) and .44 for externalizing problems (e.g., aggression, lying) (Achenbach, Krukowski, Dumenci, & Ivanova, 2005).

One method intended to overcome the limitations of self-report instruments is to use validity scales to correct for misleading self-reports. However, several reviews have concluded that validity scales do not improve the validity of results (e.g., Archer, Fontaine, & McCrae, 1998; Barthlow, Graham, Ben-Porath, Tellegen, & McNulty, 2002; McGrath, Mitchell, Kim, & Hough, 2010). In a cross-cultural study that found validity scales to reduce validity, the authors concluded that multimethod approaches are superior to validity scales, because "The best evidence on protocol validity, and the best alternative to the use of validity scales, comes from the comparison of self-report scores

with independent assessments," i.e., reports by informants who know the people being assessed (Piedmont, McCrae, Riemann, & Angleitner, 2000, p. 590). Although standardized multi-informant methods are often used for assessing child psychopathology (De Los Reyes, Thomas, Goodman, & Kundey, 2013), such methods may also improve assessment of adult psychopathology (De Los Reyes, Bunnell, & Beidel, 2013).

One way to advance standardized multi-informant assessment of adult psychopathology is to develop instruments for obtaining collateral reports that parallel self-report instruments. Accordingly, the Adult Behavior Checklist (ABCL) was developed for completion by collaterals to assess many of the same problems that are assessed by the ASR. The ABCL has counterparts of 115 of the 120 ASR problem items but omits items not likely to be ratable by informants (e.g., "heart pounding"). The ABCL also has three items that are not likely to be self-reported (e.g., "stares blankly"). Like the ASR items, ABCL items are rated 0-1-2, based on the preceding 6 months. EFAs and CFAs of ABCL ratings by collaterals of 1,660 U.S. referred and nonreferred 18-59-year-olds yielded counterparts of the eight syndromes found in ASR self-ratings (Achenbach & Rescorla, 2003). The collaterals included spouses/partners, family members, and friends. Correlations between ASR and ABCL scores were .43 for internalizing problems and .44 for externalizing problems, the same as the mean cross-informant correlations found in the meta-analyses of many instruments in many studies (Achenbach et al., 2005).

Testing syndromes in other societies

There are multiple reasons why we cannot assume that assessment instruments developed in one society would be applicable in other societies. As an example, genetic factors affecting behavioral covariation may differ across societies (Way & Lieberman, 2010). Additionally, different behaviors may be encouraged or discouraged in different societies, leading to different groupings of behaviors across societies (Weisz, Weiss, Suwanlert & Chaiyasit, 2006).

To evaluate the applicability of assessment instruments to societies other than where they were developed, they must be used to assess large samples from those societies. To test the degree to which ratings of problems in other societies fit particular syndrome models, CFAs should be applied to the ratings obtained in each society. CFAs of ASR self-ratings by 17,152 18-59-year-olds from 29 diverse societies have supported the eight-syndrome model in each society (Ivanova et al., in press). However, to our knowledge, no published studies report multi-society tests of syndromes of adult psychopathology derived statistically from collateral ratings.

Although focused on personality dimensions rather than psychopathology, EFAs of ratings of the NEO-PI-R (Costa & McCrae, 1992) by collaterals in multiple societies may be instructive. The NEO-PI-R is a 240-item questionnaire for assessing personality dimensions designated as Neuroticism, Extraversion, Openness-to-Experience, Agreeableness, and Conscientiousness. McCrae Terracciano and 78 Members of the Personality Profiles of Cultures Project (2005) performed EFAs on NEO-PI-R collateral ratings for 11,985

18- to 59-year-olds. Data were collected in 50 societies representing a wide range of world regions. Results supported the five NEO-PI-R personality factors in all tested societies, although support was weaker for the five African societies.

Purpose of the present study

Various studies indicate that reports by collaterals provide important information about adult psychopathology beyond what self-reports provide. The ABCL was developed to obtain informants' ratings of most of the problems that are self-rated on the ASR. Factor analyses of the ABCL and ASR for U.S. samples have yielded eight similar syndromes that enable researchers and clinicians to rigorously compare collateral versus self-reports. CFAs of ASR ratings from 29 societies have supported the eight-syndrome model in self-ratings but the multi-society generalizability of the syndromes scored from the ABCL needs to be tested. The purpose of this study was to test whether the ABCL eight-syndrome model would be supported in collateral ratings of adults in societies that differed in many ways from the U.S. The multi-society CFA findings for ASR self-ratings suggest that ABCL collateral ratings would also support the eight-syndrome model.

Method

Samples

The ABCL was completed by 8,582 informants from the 18 samples described in Table 1. Informants were instructed to complete the ABCL about an 18- to 59-year-old they knew well. The ABCLs were completed for 8,582 different adults, i.e., no adult was rated by > 1 informant. Each sample was collected by indigenous researchers, who followed procedures for protection of research participants in their respective institutions. Samples ranged from 282 (Poland) to 1,000 (Japan), with a mean of 477. The gender of assessment targets averaged 46% male. Following scoring conventions for the ABCL (Achenbach & Rescorla, 2003), forms with ≥ 8 missing items were excluded from analyses. Excluded forms comprised .0081 of the total sample of ABCLs (0% for Albania, Algeria, Argentina, Flanders, France, Japan, Korea, Lithuania, Portugal, Serbia, and Taiwan to 8.8% for Iceland).

Tested model

Each of the 93 items with significant factor loadings was assigned to only one factor. For each factor, the first item was specified as the metric item (i.e., the scale of the latent factor is set to the scale of that item). All latent factors were modeled as correlated first-order factors, and no hierarchical relations between factors were specified. Item 37. *Gets in many fights* was omitted for Taiwan, because it was not endorsed by any participant. Because endorsements would have required the investigator to report respondents to the police, four items assessing illegal behavior were not used in Japan (6. *Uses drugs (other than alcohol or nicotine) for nonmedical purposes*; 57. *Physically attacks people*; 82.

Steals; and 92. *Does things that may cause trouble with the law*).

Data analyses

The correlated eight-factor model was tested separately in each society, following the factor analytic procedures reported by Achenbach and Rescorla (2003). All 0-1-2 item ratings were transformed to 0 versus 1 or 2, and tetrachoric correlations were computed on these ratings. To account for the nonnormal distribution of the ratings, we used the robust WLSMV estimator (Muthén & Muthén, 1998-2012). It involves computing weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean- and variance-adjusted chi-square test statistic.

We selected the Root Mean Square Error of Approximation (RMSEA) as the primary model fit index because it was identified as the best performing index for the WLSMV (Yu & Muthén, 2002). Yu and Muthén (2002) found that RMSEA values of < .05-.06 indicated good model fit for ordered categorical variables in a simulation study. We also computed the Comparative Fit Index (CFI; Bentler, 1990) and Tucker Lewis Index (TLI; Tucker & Lewis, 1973), but considered their results to be secondary to the RMSEA. We did not use other model fit indices generated by Mplus for CFAs with the WLSMV (e.g., SRMR, WRMR), because Yu and Muthén (2002) found that they did not perform as reliably as the RMSEA, CFI, and TLI with binary or ordered categorical variables. Based on the results of their simulation study with three five-item factors, Hu and Bentler (1999) proposed that CFI and TLI values greater than .95 should be interpreted to indicate good model fit. However, Marsh, Hau, and Wen (2004) argued that Hu and Bentler's threshold was too stringent for complex models. Because our model was much more complex than Hu and Bentler's, we adopted the less stringent criteria of .80 to .90 to indicate acceptable model fit, and $\geq .90$ to indicate good model fit.

Results

The correlated eight-factor model converged for all samples. RMSEAs ranged from .016 (China and Iceland) to .029 (Albania and Lithuania), indicating good fit for all societies (Table 2). The RMSEAs equaled .021, .025, and .027 at the 25th, 50th, and 75th percentiles, respectively. CFIs ranged from .775 for Portugal to .955 for Iceland. TLI values were similar to CFI values within societies, ranging from .768 for Portugal to .954 for Iceland. CFIs and TLIs indicated acceptable to good model fit for all societies, except Portugal and Algeria.

For 10 societies, all 93 tested items had statistically significant loadings on their respective factors. As shown in Table 2, 17 item loadings were not statistically significant (one for Albania and Iceland, two for China, France, Lithuania, and Taiwan, three for Portugal, and four for Algeria). The 17 nonsignificant loadings comprised 1.01% of the 1,669 loadings we tested (92 for Taiwan and 89 for Japan, plus 93 for the remaining 16 societies).

Table 2 presents medians and ranges of factor loadings for each society. The median factor loading ranged from .53 (Algeria) to .79 (Japan), with an overall median of .70.

Table 1 Samples used for CFA.

Society	Reference	<i>N</i>	Targets' Age range	Targets' % Male	Sample
Albania	Sokoli (2012) ^a	750	18-59	50	Collaterals selected by ASR participants recruited in a representative national sample.
Algeria	Chaalal (2012) ^a	300	18-59	66	University students completed ABCLs for their acquaintances.
Argentina	Samaniego and Vázquez (2012)	679	18-59	48	Collaterals selected by ASR participants recruited in a regional sample stratified by level of educational attainment to be representative of the greater Buenos Aires area.
Brazil	Silvares et al. (2012) ^a	679	18-59	40	Collaterals selected with ASR participants recruited in a national sample stratified by region, age, gender, and socioeconomic status to be representative of the Brazilian metropolitan population.
China	Liu (2012) ^a	515	18-59	43	Collaterals selected by ASR participants recruited in regional samples drawn from regions of mainland China.
Flanders	Decoster and Fontaine (2012) ^a	737	18-59	50	Collaterals selected by ASR participants recruited in a regional sample stratified by region, gender, age, and educational attainment to be representative of Flanders.
France	Petot, Vrignaud, Zebdi and Camart (in press)	395	18-59	51	University students completed ABCLs for their acquaintances.
Hong Kong	Au and Leung (2012) ^a	330	18-59	40	Collaterals selected by ASR participants recruited in a community sample stratified by age and gender to be representative of the Hong Kong population.
Iceland	Guðmundsson and Árnadóttir (2012) ^a	299	18-59	46	A small portion of ABCLs was completed by collaterals selected by ASR participants recruited in a representative national sample randomly drawn from a national registry, with the rest being completed by ASR participants about their acquaintances.
Japan	Funabiki (2012) ^a	1,000	18-59	47	Collaterals selected by ASR participants recruited from multiple regions by a research company.
Korea	Oh and Kim (2011)	299	21-59	49	Collaterals selected by ASR participants recruited in a representative national sample randomly drawn from a national registry and stratified by age, gender, and educational attainment.
Lithuania	Šimulionienė, Brazdeikienė, Rugevičius, Gedutienė, & Žakaitienė (2010)	573	18-59	48	Collaterals selected by ASR participants recruited in a representative national sample randomly drawn from a national registry and stratified by age, gender, and educational attainment.
Poland	Zasępa (2012) ^a	282	18-59	39	Collaterals selected by ASR participants recruited in a national sample and stratified by age, gender, residence and educational attainment to be representative of the national population.

Table 1 (Continued)

Society	Reference	N	Targets' Age range	Targets' % Male	Sample
Portugal	Caldas (2012) ^a	397	18-59	49	Collaterals selected by ASR participants recruited in a regional sample and stratified by age and gender to be representative of the national population.
Russia	Malykh (2012) ^a	436	18-55	33	University students completed ABCLs for their acquaintances.
Serbia	Markovic (2012) ^a	312	18-59	43	Collaterals selected by ASR participants recruited in a randomly drawn sample of the Novi Sad metropolitan area, randomly drawn from a population registry and stratified by age.
Taiwan	Chen (2012) ^a	300	18-59	50	Collaterals selected by ASR participants recruited in a national sample stratified by region, gender, and age to be representative of the national population.
United Kingdom	Talcott et al. (2012) ^a	299	18-59	42	ASR participants recruited in a community sample completed ABCLs for their acquaintances.

^a Unpublished raw data.

Table 2 also presents medians and ranges for correlations between latent factors across the 18 societies. Median latent factor correlations ranged from .55 in the Argentine sample to .78 in the Albanian sample, with an overall median of .66.

Table 3 presents the means, medians, standard deviations, and ranges of the loadings for each item and for the

items comprising each syndrome across the 18 societies. For each item, the median factor loading across societies ranged from .49 (121. *Late for appointments*) to .87 (54. *Feels tired without good reason*), with an overall median of .71. For the items comprising each syndrome, median loadings ranged from .65 (Attention Problems) to .74 (Thought Problems).

Table 2 CFA results.

Society	N	RMSEA	CFI	TLI	Items with non-significant loadings ^a	Empirically under-identified items ^a	Median Factor Loadings (Range)	Median Factor Correlations (Range)
Albania	750	.029	.939	.938	76		.73 (.01-.95)	.78(.48-.96)
Algeria	300	.026	.795	.789	11, 19, 40, 52		.53(.11-.85)	.61(.15-.94)
Argentina	679	.027	.861	.857			.65(.28-.96)	.55(.20-.79)
Brazil	679	.026	.909	.906			.69(.34-.91)	.65(.16-.81)
China	515	.016	.926	.924	18, 82		.67(.07-.88)	.74(.52-.93)
Flanders	737	.024	.899	.896			.68(.16-.90)	.59(.28-.82)
France	395	.026	.876	.872	56d, 94		.68(.11-.93)	.58(.18-.77)
Hong Kong	330	.022	.917	.914			.72(.39-.99)	.74(.48-.90)
Iceland	299	.016	.955	.954	57	56c	.75(.35-1.001)	.70(.31-.88)
Japan	1000	.023	.927	.925		104	.79(.50-1.04)	.75(.57-.88)
Korea	299	.022	.937	.936			.74(.39-.91)	.67(.30-.88)
Lithuania	573	.029	.903	.900	6, 17		.67(.11-.95)	.70(.37-.85)
Poland	282	.026	.906	.903			.73(.28-.91)	.67(.17-.86)
Portugal	397	.027	.775	.768	7, 91, 92		.61(.01-.93)	.64(.11-.78)
Russia	436	.028	.871	.868			.66(.31-.95)	.64(.13-.82)
Serbia	312	.021	.932	.930		91	.73(.32-1.06)	.70(.36-.87)
Taiwan	300	.019	.893	.890	70, 92		.70(.14-.99)	.60(.29-.84)
United Kingdom	299	.021	.912	.909		40	.74(.26-1.07)	.61(.15-.77)

Note. RMSEA = Root Mean Square Error of Approximation, CFI = Comparative Fit Index, TLI = Tucker-Lewis Index.

^a The number is the item's number on the ABCL.

Table 3 Descriptive statistics for factor loadings across 18 societies by syndrome.

Syndromes and items	Mean loading	Median loading	SD	Range
<i>Anxious/Depressed</i>	.69	.72	.09	.57-.85
12. Lonely	.64	.66	.10	.39-.77
14. Cries a lot	.56	.59	.11	.29-.73
31. Fears doing bad	.57	.58	.16	.25-.86
33. Feels unloved	.77	.79	.08	.60-.88
34. Others out to get her	.77	.77	.08	.60-.95
35. Feels worthless	.77	.77	.08	.50-.87
45. Nervous, tense	.73	.76	.07	.61-.80
47. Lacks self-confidence	.70	.71	.07	.54-.80
50. Fearful, anxious	.69	.72	.12	.32-.87
52. Feels too guilty	.61	.65	.17	.11-.81
71. Self-conscious	.56	.57	.14	.31-.85
103. Unhappy, sad	.84	.85	.05	.73-.92
107. Can't succeed	.71	.72	.10	.52-.82
112. Worries	.57	.57	.14	.29-.76
<i>Withdrawn</i>	.69	.73	.10	.52-.77
25. Doesn't get along	.78	.77	.07	.66-.89
30. Poor relations with opposite sex	.66	.68	.10	.40-.81
42. Rather be alone	.52	.52	.07	.39-.61
48. Not liked	.74	.76	.14	.36-.95
60. Enjoys little	.74	.75	.12	.44-.91
65. Refuses to talk	.72	.73	.10	.52-.91
67. Trouble making friends	.76	.76	.06	.64-.89
69. Secretive	.50	.53	.17	.25-.74
111. Withdrawn	.68	.73	.15	.28-.85
<i>Somatic Complaints</i>	.68	.70	.13	.50-.87
51. Feels dizzy	.77	.76	.10	.56-.92
54. Tired without reason	.84	.87	.08	.66-.97
56a. Aches, pains	.67	.70	.13	.32-.80
56b. Headaches	.59	.57	.09	.39-.77
56c. Nausea, feels sick	.74	.76	.13	.41-1.001 ^a
56d. Eye problems	.49	.50	.17	.11-.79
56e. Skin problems	.50	.53	.13	.16-.69
56f. Stomachaches	.62	.64	.12	.30-.86
56g. Vomiting	.76	.77	.19	.26-1.00
<i>Thought Problems</i>	.73	.74	.07	.60-.84
9. Can't get mind off thoughts	.58	.60	.14	.32-.81
18. Harms self	.72	.78	.23	.10-.95
40. Hears sounds, voices	.69	.74	.24	.21-1.07 ^a
66. Repeats acts	.64	.67	.15	.34-.96
70. Sees things	.66	.75	.21	.23-.94
80. Stares blankly	.69	.70	.12	.38-.99
84. Strange behavior	.78	.76	.12	.50-.96
85. Strange ideas	.68	.71	.13	.43-.88
91. Suicidal thoughts	.79	.84	.19	.32-1.06 ^a
<i>Attention Problems</i>	.65	.65	.09	.49-.76
1. Forgetful	.50	.51	.05	.39-.57
8. Can't concentrate	.64	.64	.09	.40-.80
11. Too dependent	.63	.65	.13	.15-.73
13. Confused	.75	.76	.07	.56-.87
17. Daydreams	.52	.56	.16	.11-.77
53. Trouble planning	.69	.71	.07	.59-.83
59. Fails to finish	.73	.74	.08	.56-.87

Table 3 (Continued)

Syndromes and items	Mean loading	Median loading	SD	Range
61. Poor work performance	.75	.75	.09	.53-.92
64. Trouble setting priorities	.73	.75	.07	.58-.83
78. Trouble making decisions	.71	.72	.06	.62-.79
96. Lacks initiative	.65	.64	.07	.51-.79
101. Skips job	.63	.61	.15	.42-.94
102. Lacks energy	.68	.67	.10	.53-.87
105. Disorganized	.70	.71	.09	.57-.84
108. Loses things	.57	.57	.08	.41-.71
119. Not good at details	.58	.57	.11	.38-.79
121. Late for appointments	.48	.49	.09	.28-.62
<i>Aggressive Behavior</i>	<i>.70</i>	<i>.70</i>	<i>.06</i>	<i>.61-.80</i>
3. Argues	.61	.61	.09	.40-.78
5. Blames others	.68	.66	.09	.47-.82
16. Mean to others	.67	.69	.15	.48-.90
28. Gets along badly with family	.65	.66	.09	.44-.80
37. Gets in fights	.65	.61	.18	.22-1.00
55. Mood swings between elation and depression	.77	.75	.09	.64-.93
57. Attacks people	.67	.67	.12	.42-.88
68. Screams a lot	.66	.66	.09	.44-.80
81. Changeable behavior	.79	.80	.06	.62-.90
86. Stubborn, sullen, irritable	.72	.72	.09	.51-.82
87. Mood changes	.80	.79	.05	.69-.90
95. Hot temper	.71	.72	.08	.60-.88
97. Threatens people	.69	.71	.15	.46-.95
113. Sulks	.74	.75	.08	.49-.84
116. Easily upset	.76	.77	.11	.42-.91
118. Impatient	.68	.68	.11	.53-.91
<i>Rule-Breaking Behavior</i>	<i>.68</i>	<i>.71</i>	<i>.09</i>	<i>.51-.86</i>
6. Uses drugs	.48	.51	.13	.23-.65
23. Breaks rules	.66	.65	.10	.48-.84
26. Lacks guilt	.60	.63	.13	.35-.83
39. Bad friends	.66	.71	.16	.31-.93
41. Impulsive	.72	.72	.10	.47-.85
43. Lying, cheating	.71	.72	.09	.58-.83
76. Irresponsible	.79	.86	.21	.01-.93
82. Steals	.68	.72	.20	.07-.93
90. Gets drunk	.49	.52	.11	.26-.63
92. Trouble with the law	.64	.72	.24	.01-.85
114. Fails to pay debts	.69	.73	.17	.25-.93
117. Trouble managing money	.67	.68	.09	.47-.81
122. Trouble keeping jobs	.72	.71	.12	.54-.96
<i>Intrusive</i>	<i>.68</i>	<i>.69</i>	<i>.06</i>	<i>.59-.76</i>
7. Brags	.63	.67	.21	.06-.93
19. Demands attention	.69	.70	.18	.11-.89
74. Showing off, clowning	.65	.70	.17	.30-.89
93. Talks too much	.59	.59	.10	.41-.75
94. Teases a lot	.67	.68	.19	.12-.93
104. Loud	.75	.76	.12	.53-1.04 ^a

Note. Values in italics are descriptive statistics for median item loadings comprising the syndromes.

^a The 95% confidence intervals around out-of-range factor loadings included values that were in the admissible parameter space (0.00 - 1.00).

For four societies, one item had a negative residual, indicating that it was misspecified (item 56c. *Nausea, feels sick* for Iceland; item 104. *Is unusually loud* for Japan; item 91. *Talks about killing self* for Serbia; and item 40. *Hears sounds or voices that aren't there (describe)* for the UK). Because the 95% confidence intervals for all four parameters included the admissible parameter space, sampling fluctuations may explain their misspecification (Chen, Bollen, Paxton, Curran, & Kirby, 2001). These four out-of-range parameters comprised only 0.08% of the 5,007 parameters that we estimated (93 thresholds, item loadings, and residuals for 16 societies, 92 thresholds, item loadings, and residuals for Taiwan, and 89 thresholds, item loadings, and residuals for Japan).

Discussion/conclusions

The ABCL was developed to meet needs for multi-informant assessment of adult psychopathology by obtaining informant ratings of the same clinical constructs as rated by the people being assessed. Because findings in the society where an instrument originated cannot be assumed to generalize to other societies without empirical evidence, instruments and constructs should be empirically tested in multiple societies. The more diverse the societies in which an instrument's constructs are supported, the more confidence their generalizability warrants. We used CFAs to test the ABCL's eight-syndrome model in 18 societies from such very different regions as Africa, Asia, South America, and Eastern, Northern, and Western Europe.

We found that the eight-syndrome model scored from ABCL ratings by collaterals was supported by RMSEAs ranging from .016 for China and Iceland to .029 for Albania and Lithuania. The small size of the RMSEAs indicated very good fit for the eight-syndrome model in all 18 societies in which collaterals completed ABCLs. The secondary fit indices indicated acceptable to good fit in all societies except Algeria and Portugal. Of the 1,669 loadings that were tested, the median loading was .71, and only 17 (1.01%) were not significant. Of the 5,007 parameters that were estimated, 4 (0.08%) were out of range, but 95% confidence intervals included the admissible range for all 4. Our findings thus supported the same eight-syndrome model for collateral ratings as was previously supported for self-report ratings (Ivanova et al., in press).

Our findings indicate that the eight-syndrome model was supported for operationalizing phenotypes of mutually associated problems rated by collaterals in 18 societies that differed in ethnicity, language, religion, geographical region, and economic, mental health, and political systems. Combined with similar findings for the eight-syndrome model in ASR self-ratings from 17 of these societies (Ivanova et al., in press), our findings support the applicability of the syndromes scored from the parallel collateral- and self-rating instruments for 18-59-year-olds in many parts of the world. Additional items, instruments, and models are likely to be useful. However, the ASR and ABCL items and the eight syndromes scored from them provide a core data language for describing and quantifying psychopathology that can be widely used for research and clinical purposes in conjunction with whatever additional items, instruments, and/or models may be warranted in particular contexts and societies.

Implications

The findings support use of the ABCL to assess 18-59-year-olds in terms of the eight-syndrome model scored from collateral ratings in many societies. When the people being assessed rate themselves on the ASR, the parallel items and scales of the two instruments enable clinicians and researchers to identify specific similarities and differences between collateral- and self-reports. To evaluate the degree of deviance indicated by particular scores on each syndrome, research is needed to test differences between scores for different societies, males versus females in different age groups, and collateral- versus self-ratings, as is now being done (Rescorla, 2014; Rescorla et al., 2014). The findings provide a basis for constructing norms that take account of differences in distributions of scores for different societies, each gender/age group, and collateral- versus self-ratings. Computer software will then enable users to display syndrome scale scores in relation to norms appropriate for each society and gender/age group for collateral- and self-ratings.

Limitations

The present study tested the configural invariance of the ABCL syndrome structure in data from 18 societies. Configural invariance, which is the most fundamental component of measurement invariance, means that an assessment instrument measures the same construct(s) across different populations (e.g., Vandenberg & Lance, 2000). Because our tested model was complex and the WLSMV estimator (which we used to account for the ordered categorical nature of our data) is very computationally intensive, we were unable to successfully execute tests of other components of measurement invariance (i.e., metric, scalar, item residual, and factor), which all require multi-sample modelling.

Although the study reported here extends evidence-based collateral assessment of psychopathology across diverse societies, it is limited to what Pike (1967) called *etic* research, which applies the same standardized assessments to people in different societies. We did not hypothesize that the ABCL syndrome structure is universal nor that every ABCL item is equally applicable in all societies. Instead, the study's purpose was to test the degree to which the ABCL syndrome structure fit ratings of the ABCL items in 18 specific societies.

Detailed comparisons of scores in different societies (Rescorla, 2014; Rescorla et al., 2014) may reveal societal differences that warrant what Pike (1967) called *emic* research, which is customized for particular societies to explore society-specific phenomena. Emic research may also be used to test the validity of the ABCL items and syndromes according to society-specific criteria. The findings may argue for use of other items and syndrome models in particular societies.

References

- Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological*

- Bulletin*, 131, 361–382. <http://dx.doi.org/10.1037/0033-2909.131.3.361>
- 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.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)* (5th ed.). Washington, DC: Author.
- Archer, R. P., Fontaine, J., & McCrae, R. R. (1998). Effects of two MMPI-2 validity scales on basic scale relations to external criteria. *Journal of Personality Assessment*, 70, 87–102. <http://dx.doi.org/10.1207/s15327752jpa7001>
- Asberg, M., Montgomery, S. A., Perris, C., Schalling, D., & Sedvall, G. (1978). A comprehensive psychopathological rating scale. *Acta Psychiatrica Scandinavica*, 57(S271), 5–27. <http://dx.doi.org/10.1111/j.1600-0447.1978.tb02357.x>
- Barthlow, D. L., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., & McNulty, J. L. (2002). The appropriateness of the MMPI-2K correction. *Assessment*, 9, 219–229. <http://dx.doi.org/10.1177/1073191102009003001>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238–246. <http://dx.doi.org/10.1037/0033-2909.107.2.238>
- Bérubé, R. L., & Achenbach, T. M. (2014). *Bibliography of published studies using the Achenbach System of Empirically Based Assessment (ASEBA)*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods and Research*, 29, 468–508. <http://dx.doi.org/10.1177/0049124101029004003>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources Inc.
- De Los Reyes, A., Bunnell, B. E., & Beidel, D. C. (2013). Informant discrepancies in adult social anxiety disorder assessments: Links with contextual variations in observed behavior. *Journal of Abnormal Psychology*, 122, 376–386. <http://dx.doi.org/10.1037/a0031150>
- De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, 9, 123–149. <http://dx.doi.org/10.1146/annurev-clinpsy-050212-185617>
- Derogatis, L. R. (1994). *Symptom Checklist 90–R: Administration, scoring and procedures manual* (3rd ed.). Minneapolis, MN: National Computer Systems.
- Eapen, V., Fox-Hiley, P., Banerjee, S., & Robertson, M. (2004). Clinical features and associated psychopathology in a Tourette syndrome cohort. *Acta Neurologica Scandinavica*, 109, 255–260. <http://dx.doi.org/10.1046/j.1600-0404.2003.00228.x>
- Goekoop, J. G., Hoeksema, T., Knoppert-Van der Klein, E. A. M., Klinkhamer, R. A., Van Gaalen, H. A. E., Van Londen, L., De Weme, R., & Zwinderman, A. H. (1992). Multidimensional ordering of psychopathology: A factor-analytic study using the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica*, 86, 306–312. <http://dx.doi.org/10.1111/j.1600-0447.1992.tb03271.x>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Ivanova, M.Y., Achenbach, T.M., Rescorla, L.A., Turner, L.V., Ahmeti-Pronaj, A., Au, A., Avila Maese, C., Bellina, M., Caldas, J.C., Chen, Y.-C., Csemy, L., da Rocha, M.M., Decoster, J., Dobrean, A., Ezpeleta, L., Fontaine, J.R.J., Funabiki, Y., Gudmundsson, H.S., Harder, V.S., Leiner de la Cabada, M., Leung, P., Liu, J., Mahr, S., Malykh, S., Maras, J.S., Markovic, J., Ndetei, D.M., Oh, K.J., Petot, J.-M., Riad, G., Sakarya, D., Samaniego, V.C., Sebre, S., Shahini, M., Silvaes, E., Simulioniene, R., Sokoli, E., Talcott, J.B., Vazquez, N., & Zasepa, E. (in press). Syndromes of self-reported psychopathology for ages 18-59 in 29 societies. *Psychopathology and Behavioral Assessment*. [doi:10.1007/s10862-014-9448-8](https://doi.org/10.1007/s10862-014-9448-8)
- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cut-off values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320–341. http://dx.doi.org/10.1207/s15328007sem1103_2
- McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88, 547–561. <http://dx.doi.org/10.1037/0022-3514.88.3.547>
- McGrath, R. E., Mitchell, M., Kim, B. H., & Hough, L. (2010). Evidence for response bias as a source of error variance in applied assessment. *Psychological Bulletin*, 136, 450–470. <http://dx.doi.org/10.1037/a0019216>
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiscyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56, 128–165. <http://dx.doi.org/10.1037/0003-066X.56.2.128>
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.
- Oh, K. J., & Kim, Y. A. (2011). *Manual for the Korean Version of ASEBA Adult Forms and Profiles*. Seoul: Huno Press.
- Petot, J.M., Vrignaud, P., Zebdi, R., & Camart, N. (in press). Qualités psychométriques de la version française de l'inventaire de comportement pour adultes (Adult Behavior Checklist) d'Achenbach et Rescorla: Structure factorielle et normes provisoires. *Psychologie Française*.
- Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-report and observer ratings in volunteer samples. *Journal of Personality and Social Psychology*, 78, 582–593. <http://dx.doi.org/10.1037/0022-3514.78.3.582>
- Pike, K. L. (1967). *Language in relation to a unified theory of the structure of human behavior*. The Hague: Mouton.
- Rescorla, L.A., Achenbach, T.M., Ivanova, M.Y., Turner, L.V. Au, A., Caldas, J.C., Chen, Y.C., da Rocha, M.M., Decoster, J., Fontaine, J., Funabiki, Y., Gudmundsson, H., Leung, P., Maras, J.S., Markovic, J., Oh, K.J., Samaniego, V.C., Silvaes, E., Simulioniene, R., Sokoli, E., Vazquez, N., & Zasepa, E. 2014. Collateral reports of problems and adaptive functioning for adults in 14 societies. Manuscript in preparation.
- Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Turner, L. V., Au, A., Bellina, M., Caldas, J. C., Chen, Y.-C., Csemy, L., da Rocha, M. M., Decoster, J., Fontaine, J., Funabiki, Y., Gudmundsson, H., Leung, P., Maras, J. S., Markovic, J., Oh, K. J., Samaniego, V. C., Sebre, S., Silvaes, E., Simulioniene, R., Sokoli, E., Vazquez, N., & Zasepa, E. (2014). *Problems and adaptive functioning reported by adults in 17 societies*. *Manuscript under review*.
- Samaniego, V. C., & Vázquez, N. (2012). *Adult psychopathology: Is there any agreement between self-reports and reports by other informants? In Presented at the 30th International Congress of Psychology* Cape Town, South Africa.
- Šimulionienė, R., Brazdeikienė, L., Rugevičius, M., Gedutienė, R., & Žakaitienė, A. (2010). *Kai kurios lietuviško ASEBA suaugusiųjų klausimyno psichometrinės charakteristikos (The Psychometric Properties of the Lithuanian Version of ASEBA Adult Forms)*. *Psichologija*, 42, 23–43.
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10.

- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4–70. <http://dx.doi.org/10.1177/109442810031002>
- Way, B. M., & Lieberman, M. D. (2010). Is there a genetic contribution to cultural differences? Collectivism, individualism, and genetic markers of social sensitivity. *Social, Cognitive, and Affective Neuroscience*, 5, 203–211. <http://dx.doi.org/10.1093/scan/nsq059>
- Weisz, J. R., Weiss, B., Suwanlert, S., & Chaiyasit, W. (2006). Culture and youth psychopathology: Testing the syndromal sensitivity model in Thai and American adolescents. *Journal of Consulting and Clinical Psychology*, 74, 1098–1107. <http://dx.doi.org/10.1037/0022-006X.74.6.1098>
- Wittenborn, J. R. (1951). Symptom patterns in a group of mental hospital patients. *Journal of Consulting and Clinical Psychology*, 15, 290–302. <http://dx.doi.org/10.1037/h0054947>
- World Health Organization. (2010). *International Classification of Disease (ICD) (10th ed.)*. Geneva: Switzerland: Author.
- Yu, C. Y., & Muthén, B. O. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes (Technical Report)*. Los Angeles: UCLA: Graduate School of Education and Information Studies.