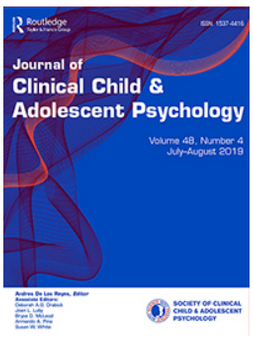
Testing syndromes of psychopathology in parent and youth ratings across societies

Achenbach, Thomas M.; Rescorla, Leslie A.; Guo, Jiesi; Althoff, Robert R.; Kan, Kees-Jan; Almqvist, Fredrik; Begovac, Ivan; Broberg, Anders G.; Chahed, Myriam; da Rocha, Marina Monzani

Published in:
Journal of Clinical Child and Adolescent Psychology

DOI:
10.1080/15374416.2017.1405352

Publication date:
2019



Routledge
Taylor & Francis Group

Journal of Clinical Child & Adolescent Psychology

ISSN: 1537-4416 (Print) 1537-4424 (Online) Journal homepage: https://www.tandfonline.com/loi/hcap20

Testing Syndromes of Psychopathology in Parent and Youth Ratings Across Societies

Masha Y. Ivanova, Thomas M. Achenbach, Leslie A. Rescorla, Jiesi Guo, Robert R. Althoff, Kees-Jan Kan, Fredrik Almqvist, Ivan Begovac, Anders G. Broberg, Myriam Chahed, Marina Monzani da Rocha, Anca Dobrean, Manfred Döepfner, Nese Erol, Eric Fombonne, Antonio Castro Fonseca, Maria Forns, Alessandra Frigerio, Hans Grietens, Nohelia Hewitt-Ramirez, Fernando Juarez, Ilona Kajokienė, Yasuko Kanbayashi, Young-Ah Kim, Bo Larsson, Patrick Leung, Xianchen Liu, Alfio Maggiolini, Asghar Minaei, Paulo A.S. Moreira, Kyung Ja Oh, Djaouida Petot, Cecilia Pisa, Rolando Pomalima, Alexandra Roussos, Vlasta Rudan, Michael Sawyer, Mimoza Shahini, Edwiges Ferreira de Mattos Silvares, Zeynep Simsek, Hans-Christoph Steinhausen, Lajos Szirovicza, Jose Valverde, Laura Viola, Sheila Weintraub, Christa Winkler Metzke, Tomasz Wolanczyk, Bernardine Woo, Eugene Yuqing Zhang, Nelly Zilber, Rita Žukauskienė & Frank C. Verhulst

To cite this article: Masha Y. Ivanova, Thomas M. Achenbach, Leslie A. Rescorla, Jiesi Guo, Robert R. Althoff, Kees-Jan Kan, Fredrik Almqvist, Ivan Begovac, Anders G. Broberg, Myriam Chahed, Marina Monzani da Rocha, Anca Dobrean, Manfred Döepfner, Nese Erol, Eric Fombonne, Antonio Castro Fonseca, Maria Forns, Alessandra Frigerio, Hans Grietens, Nohelia Hewitt-Ramirez, Fernando Juarez, Ilona Kajokienė, Yasuko Kanbayashi, Young-Ah Kim, Bo Larsson, Patrick Leung, Xianchen Liu, Alfio Maggiolini, Asghar Minaei, Paulo A.S. Moreira, Kyung Ja Oh, Djaouida Petot, Cecilia Pisa, Rolando Pomalima, Alexandra Roussos, Vlasta Rudan, Michael Sawyer, Mimoza Shahini, Edwiges Ferreira de Mattos Silvares, Zeynep Simsek, Hans-Christoph Steinhausen, Lajos Szirovicza, Jose Valverde, Laura Viola, Sheila Weintraub, Christa Winkler Metzke, Tomasz Wolanczyk, Bernardine Woo, Eugene Yuqing Zhang, Nelly Zilber, Rita Žukauskienė & Frank C. Verhulst (2019) Testing Syndromes of Psychopathology in Parent and Youth Ratings Across Societies, Journal of Clinical Child & Adolescent Psychology, 48:4, 596-609, DOI: 10.1080/15374416.2017.1405352

To link to this article: https://doi.org/10.1080/15374416.2017.1405352

Journal of Clinical Child & Adolescent Psychology, 48(4), 596–609, 2019

ISSN: 1537-4416 print/1537-4424 online
DOI: https://doi.org/10.1080/15374416.2017.1405352

Testing Syndromes of Psychopathology in Parent and Youth Ratings Across Societies

Masha Y. Ivanova[1], Thomas M. Achenbach[1], Leslie A. Rescorla[2], Jiesi Guo[3], Robert R. Althoff[1], Kees-Jan Kan[4], Fredrik Almqvist[5], Ivan Begovac[6], Anders G. Broberg[7], Myriam Chahed[8], Marina Monzani da Rocha[9], Anca Dobrean[10], Manfred Döepfner[11], Nese Erol[12], Eric Fombonne[13], Antonio Castro Fonseca[14], Maria Forns[15], Alessandra Frigerio[16], Hans Grietens[17], Nohelia Hewitt-Ramirez[18], Fernando Juarez[19], Ilona Kajokienė[20], Yasuko Kanbayashi[21], Young-Ah Kim[22], Bo Larsson[23], Patrick Leung[24], Xianchen Liu[25], Alfio Maggiolini[26], Asghar Minaei[27], Paulo A.S. Moreira[28], Kyung Ja Oh[29], Djaouida Petot[30], Cecilia Pisa[31], Rolando Pomalima[32], Alexandra Roussos[33], Vlasta Rudan[6], Michael Sawyer[34], Mimoza Shahini[35], Edwiges Ferreira de Mattos Silvares[36], Zeynep Simsek[37], Hans-Christoph Steinhausen[38], Lajos Szirovicza[39], Jose Valverde[32], Laura Viola[40], Sheila Weintraub[5], Christa Winkler Metzke[38], Tomasz Wolanczyk[41], Bernardine Woo[42], Eugene Yuqing Zhang[43], Nelly Zilber[44], Rita Žukauskienė[20], and Frank C. Verhulst[45]

[1]*Department of Psychiatry, University of Vermont*
[2]*Department of Psychology, Bryn Mawr College*
[3]*Institute for Positive Psychology and Education, Australian Catholic University*
[4]*College of Child Development and Education, University Amsterdam*
[5]*Department of Child Psychiatry, University of Helsinki*
[6]*Department of Psychological Medicine, University of Zagreb*
[7]*Department of Psychology, University of Gothenburg*
[8]*Department of Psychology, Université Paris Ouest Nanterre La Défense*
[9]*Center for Biological and Health Sciences, Mackenzie Presbyterian University*
[10]*Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University*
[11]*Department of Child and Adolescent Psychiatry, University of Cologne*
[12]*Department of Child and Adolescent Mental Health, Ankara University*
[13]*Department of Psychiatry, Oregon Health and Science University*
[14]*Faculdade de Psicologia e de Ciências da Educação, Universidade de Coimbra*
[15]*Department of Personality, Evaluation, and Psychological Treatment, University of Barcelona*
[16]*Child Psychopathology Unit, Scientific Institute, IRCCS Eugenio Medea*
[17]*Centre for Special Needs Education & Youth Care, University of Groningen*
[18]*Programa de Psicología, Universidad de San Buenaventura*
[19]*Escuela de Administración, Universidad del Rosario*
[20]*Institute of Psychology, Mykolas Romeris University*
[21]*Faculty of Letters, Chuo University*
[22]*Huno Inc.*
[23]*Department of Neuroscience, the Norwegian University of Science and Technology*
[24]*Department of Psychology, the Chinese University of Hong Kong*
[25]*Center for Studies of Psychological Application, South China Normal University*
[26]*Department of Psychology, Università degli Studi di Milano – Bicocca*
[27]*Department of Educational and Psychological Measurement, Allameh Tabataba'i University*
[28]*Instituto de Psicologia e Ciências da Educação, Universidade Lusíada Norte (Porto)*
[29]*Department of Psychology, Yonsei University*

Correspondence should be addressed to Masha Y. Ivanova, Department of Psychiatry, University of Vermont, 1 South Prospect Street, Burlington, VT 05401. E-mail: Masha.Ivanova@uvm.edu

[30]Department of Psychology, Université Paris Ouest Nanterre La Défense
[31]Minotauro Istituto di Analisi dei Codici Affettivi
[32]Instituto Nacional de Salud Mental Honorio Delgado Hideyo Noguchi
[33]Attiki Child Psychiatric Hospital
[34]School of Medicine, University of Adelaide & Research and Evaluation Unit, Women's and Children's Health Network
[35]Department of Child and Adolescent Psychiatry, University Clinical Center of Kosovo
[36]Department of Clinical Psychology, Universidade de São Paulo
[37]Department of Public Health, Harran University
[38]Department of Child and Adolescent Psychiatry, University of Zurich
[39]Institute for Anthropological Research
[40]Department of Child Psychiatry, Hospital de Niños, Sociedad Española
[41]Department of Child Psychiatry, Medical University of Warsaw
[42]Department of Child and Adolescent Psychiatry, Institute of Mental Health Singapore
[43]Institute of Psychology, Chinese Academy of Sciences
[44]Falk Institute for Mental Health Studies, Jerusalem & CRFJ (French Research Center in Jerusalem)
[45]Department of Child and Adolescent Psychiatry, Erasmus University Medical Center-Sophia's Children's Hospital

As societies become increasingly diverse, mental health professionals need instruments for assessing emotional, behavioral, and social problems in terms of constructs that are supported within and across societies. Building on decades of research findings, multisample alignment confirmatory factor analyses tested an empirically based 8-syndrome model on parent ratings across 30 societies and youth self-ratings across 19 societies. The Child Behavior Checklist for Ages 6–18 and Youth Self-Report for Ages 11–18 were used to measure syndromes descriptively designated as *Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior*, and *Aggressive Behavior*. For both parent ratings (N = 61,703) and self-ratings (N = 29,486), results supported aggregation of problem items into 8 first-order syndromes for all societies (configural invariance), plus the invariance of item loadings (metric invariance) across the majority of societies. Supported across many societies in both parent and self-ratings, the 8 syndromes offer a parsimonious phenotypic taxonomy with clearly operationalized assessment criteria. Mental health professionals in many societies can use the 8 syndromes to assess children and youths for clinical, training, and scientific purposes.

As societies become progressively more diverse, mental health professionals increasingly need to serve children of diverse backgrounds (we use "children" to include youths). They therefore need assessment instruments that are generalizable to culturally diverse populations. Such instruments should measure constructs of psychopathology that are supported within and across societies (we use "societies" to include countries, plus other geopolitically demarcated populations that are not countries, such as Hong Kong.)

Standardized systems for assessing a broad range of children's emotional, behavioral, and social problems—such as the Achenbach System of Empirically Based Assessment (Achenbach, 2009)—are widely used by mental health professionals around the world. These systems appeal to practitioners of different disciplines and levels of training working in many settings because they are easy to administer, score, and interpret and are useful with different populations. However, because these systems have been developed mainly in Anglophone societies, it is essential to test their generalizability to other societies.

Before using an assessment instrument developed in one society in another society, it is necessary to test whether it measures the same constructs in the two societies. For example, does an assessment instrument for anxiety developed in society X also measure anxiety in society Y? It is also necessary to test whether the instrument measures constructs in the same way in the two societies. For example, does a particular score on the instrument reflect the same severity of anxiety in society Y as it does in society X? Failure to conduct these tests may lead to inaccurate assessment results and misguided treatment planning in the new society.

The importance of testing the generalizability of assessment instruments across different contexts is increasingly recognized not only in multicultural research but also in multi-informant assessment. A growing literature suggests that it is important to

establish that an assessment instrument performs similarly across informants (e.g., parent and child) before comparing its scale scores based on different informants' reports (De Los Reyes & Ohannessian, 2016; Gross, Fleming, Mason, & Haggerty, 2017; Janssens et al., 2015).

Over the past decade, the confirmatory factor analysis (CFA) framework of measurement invariance (MI; Millsap, 2011) has been used to test the generalizability of assessment instruments across societies. MI rests on the idea that an observed item score on a measure of psychopathology is a function of several knowable influences, such as the assessed person's standing on the underlying "true" factor of psychopathology, as well as systematic influences (e.g., societally appropriate norms, procedural differences in data collection) and unsystematic influences (i.e., measurement error) that are not associated with the underlying factor of psychopathology.

MI posits testable hypotheses about how sets of items measure hypothesized latent factors of psychopathology across multiple societies (Gregorich, 2006). These hypotheses pertain to different components of factor models, such as factor loadings (indices of item-factor associations) and intercepts/thresholds (indices of systematic influences) on item ratings that are unrelated to the underlying factor. *Configural* MI is the hypothesis that particular items load on the same factors across societies. *Metric* MI is the hypothesis that items have similar loadings (item-factor associations) across societies. Finally, *scalar* MI is the hypothesis that item intercepts (or thresholds for categorical data) are equivalent, that is, that systematic influences on item ratings that are not associated with the underlying factors are the same across societies. Starting from configural invariance, MI hypotheses are hierarchical, with each level building on the prior level.

MI Testing of the Child Behavior Checklist for Ages 6–18 and Youth Self-Report

Translated into many languages, the Child Behavior Checklist for Ages 6–18 (CBCL/6–18) and Youth Self-Report (YSR; Achenbach & Rescorla, 2001) obtain parent and self-reports of children's emotional, behavioral, and social problems and competencies. The CBCL/6–18 and YSR are parallel in format and content, enabling users to document cross-informant variations in ratings of children's functioning (e.g., De Los Reyes, Thomas, Goodman, & Kundey, 2013; De Los Reyes et al., 2015; De Los Reyes & Kazdin, 2005). Norms are available for a representative U.S. national household sample and for many other societies (Achenbach & Rescorla, 2015). Factor analyses of the CBCL/6–18 and YSR problem items have yielded eight syndromes descriptively designated as *Anxious/Depressed, Withdrawn/Depressed, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior*, and *Aggressive Behavior* (Achenbach & Rescorla, 2001).

Ivanova et al. (2007a, 2007b) have tested the generalizability of the eight-syndrome model by performing CFAs on CBCL/6–18 ratings of 58,051 6- to 18-year-olds from 30 societies in Asia; Africa; Australia; the Caribbean; Eastern, Western, Southern, and Northern Europe; the Middle East; and North America. Fit indices indicated good model fit for all 30 societies. Ivanova et al. also tested the eight-syndrome model using YSR ratings by 30,243 participants 11 to 18 years of age in 23 societies from the same world regions. Fit indices indicated good model fit for all societies, for both genders, and for younger versus older adolescents. More recently, CFAs of CBCL/6–18 data from 11 additional societies and YSR data from 10 additional societies yielded good fit for all societies, which now included several from South America (Rescorla et al., 2012).

The eight-syndrome model has thus been supported by CFAs of CBCL/6–18 ratings of 69,866 children in 42 societies and of YSR ratings of 38,070 youths in 34 societies (Rescorla et al., 2012). Ivanova and colleagues also demonstrated the generalizability of the eight-syndrome model to the parallel Teacher's Report Form for Ages 6–18 in CFAs of teachers' ratings of 36,317 participants 6 to 15 years of age in 36 societies in Asia; Australia; the Caribbean; Eastern, Western, and Northern Europe; and the Middle East (Ivanova et al., 2007c; Rescorla et al., 2012). In summary, the eight-syndrome model has been supported in parent, self, and teacher ratings of children's emotional, behavioral, and social problems in separate CFAs of population samples from very diverse societies.

Alignment CFA Approach to MI

Multigroup CFA may be used to statistically compare different components of a factor model in multiple groups in order to test higher levels of MI beyond configural invariance. However, the computational demands of multigroup CFA preclude its use for testing factor models as complex as the eight-syndrome model across large numbers of societies. *Alignment CFA* was developed as a less computationally demanding approach to testing MI for complex models, such as those used in our multicultural research (Asparouhov & Muthén, 2014; Muthén & Asparouhov, 2014).

Alignment CFA estimates the entire factor model in the context of multisample CFA without requiring scalar invariance by allowing modest parameter noninvariance (approximate invariance). Equally important, because alignment models can absorb some noninvariance when model misspecifications are trivial, alignment CFA estimates all model parameters simultaneously, rather than terminating the estimation process and requiring post hoc sequential model modification. Moreover, Marsh et al. (2017) found that alignment CFA outperformed the traditional MI approach by yielding more accurate parameter estimates in the scalar model, including estimates of latent factor

TABLE 1
Descriptive Information for the Samples

Society	*Reference*	N	*Age Range*	*M Age (SD)*	*% Male*	*Sampling Frame*	N *(%) Invariant Loadings*	N *(%) Invariant Thresholds*
1. Australia	CBCL: Sawyer et al. (2000)[b]	3078	6–16	11.1 (3.2)	50	National household	85 (89)	49 (51)
	YSR: Sawyer et al. (2000)[b]	1114	11–16	14.5 (1.1)	49	National household	83 (94)	54 (61)
2. Brazil	CBCL: Rocha et al. (2013)[b]	1228	6–11	8.1 (1.3)	51	Regional convenience	88 (92)	50 (52)
	YSR: Rocha (2012)[b]	2728	11–16	13.6 (1.6)	47	Regional school-based	81 (92)	55 (63)
3. China	CBCL: Liu et al. (1999)[b]	4858	6–15	10.5 (2.9)	51	Regional household	75 (78)	40 (42)
	YSR: Wang, Zhang, and Leung (2005)[b]	967	11–16	14.7 (1.4)	51	Regional school-based	78 (89)	60 (68)
4. Colombia	CBCL: Hewitt et al. (2012)[a,b]	1491	6–16	10.6 (2.6)	53	Regional school-based	84 (88)	47 (49)
5. Croatia	CBCL: Rudan, Begovac, Szirovicza, Filipovi, and Skocic (2005)[b]	2373	6–16	11.1 (2.8)	48	National school-based	89 (93)	58 (60)
6. Finland	CBCL: Weintraub (2004)[a,b]	2093	6–16	10.3 (2.4)	49	Regional school-based	87 (91)	48 (50)
7. France	CBCL: Fombonne and Vermeersch (1997)[b]	2133	6–16	11.2 (3.2)	52	National electric/gas company	85 (89)	51 (53)
8. Germany	CBCL: Döpfner et al. (1997)[b]	2184	6–16	11.6 (3.0)	50	National household	88 (92)	65 (68)
	YSR: Döpfner et al. (1997)[b]	1497	11–16	13.3 (1.7)	51	National household	77 (88)	64 (73)
9. Greece	CBCL: Roussos et al. (1999)[b]	1113	6–11	8.7 (1.6)	50	National school-based	92 (96)	50 (52)
	YSR: Roussos et al. (2001)[b]	1222	11–16	14.7 (1.3)	49	National school-based	82 (93)	55 (63)
10. Hong Kong	CBCL: Leung et al. (2006)[b]	2033	6–16	11.3 (3.0)	49	Regional school-based	82 (85)	52 (54)
	YSR: Leung et al. (2006)[b]	1331	11–16	14.1 (1.3)	52	Regional school-based	81 (92)	50 (57)
11. Iran	CBCL: Minaei (2005)[b]	1205	6–16	11.4 (3.1)	54	Regional school-based	85 (89)	56 (58)
12. Israel	CBCL: Zilber, Auerbach, and Lerner (1994)[b]	1117	6–16	11.0 (3.1)	51	Regional household	91 (95)	75 (78)
13. Italy	CBCL: Frigerio et al. (2004)[b]	1063	6–16	11.0 (3.1)	48	Regional school-based	93 (97)	63 (66)
	YSR: Pisa and Maggiolini (2011)[a,b]	1224	11–16	15.1 (.8)	44	Regional school-based	75 (85)	51 (58)
14. Japan	CBCL: Itani et al. (2001)[b]	4645	6–16	10.8 (2.8)	48	Regional school-based	76 (79)	50 (52)
	YSR: Kuramoto et al. (2002)[b]	2542	11–16	13.1 (1.3)	48	Regional school-based	70 (80)	36 (41)
15. Korea (South)	CBCL: Oh, Lee, Hong, and Ha (1997a)[b]	3081	6–16	12.0 (3.1)	50	National school-based	82 (85)	48 (50)
	YSR: Oh, Hong, and Lee (1997b)[b]	2750	11–16	14.4 (1.2)	40	National school-based	65 (73)	45 (51)
16. Kosovo	CBCL: Shahini et al. (2015a)[b]	1332	6–16	11.0 (2.7)	51	National school-based	82 (85)	58 (60)
	YSR: Shahini, Rescorla, Wancata, and Ahmeti (2015b)[b]	1143	11–16	13.7 (1.7)	51	National school-based	70 (80)	45 (51)
17. Lithuania	CBCL: Zukauskiene, Ignataviciene, and Daukantaite (2003)[b]	2920	6–16	11.3 (3.3)	49	National school-based	84 (88)	40 (42)
	YSR: Zukauskiene and Kajokiene (2004)[b]	2022	11–16	13.8 (1.7)	49	National school-based	81 (92)	48 (55)
18. Netherlands	CBCL: Verhulst, Van Der Ende, and Koot (1996)[b]	1715	6–16	10.9 (3.2)	51	National household	83 (86)	64 (67)
19. Peru	CBCL: Pomalima and Valverde (2009)[a,b]	2296	6–16	10.7 (3.2)	49	Regional household	85 (89)	46 (48)
	YSR: Pomalima and Valverde (2009)[a,b]	1180	11–16	13.4 (1.7)	50	Regional household	82 (93)	57 (65)
20. Poland	CBCL: Wolanczyk (2003)[b]	2479	6–16	12.3 (2.7)	49	National school-based	88 (92)	43 (45)
	YSR: Wolanczyk (2003)[b]	1660	11–16	13.7 (1.7)	49	National school-based	83 (94)	61 (69)
21. Portugal	CBCL: Fonseca et al. (1995)[b]	1120	6–11	9.2 (1.6)	50	Regional school-based	89 (93)	54 (56)
	YSR: Moreira and Oliveira (2012)[a,b]	946	12–16	14.6 (1.3)	48	Regional school-based	78 (89)	61 (69)
22. Romania	CBCL: Domuta (2004)[b]	990	6–11	7.3 (1.2)	50	Regional school-based	88 (92)	55 (57)
23. Russia	CBCL: Hellinckx, Grietens, and De Munter (2000)[b]	1998	12–16	14.0 (1.4)	50	National household	80 (83)	39 (41)
24. Singapore	CBCL: Woo et al. (2007)[b]	1605	6–11	9.4 (1.8)	48	National school-based	92 (96)	57 (59)
25. Spain	YSR: Abad, Forns, and Gomez (2002)[b]	1337	11–16	13.7 (1.6)	51	Regional school-based	76 (86)	56 (64)

(*Continued*)

TABLE 1
(Continued)

Society	Reference	N	Age Range	M Age (SD)	% Male	Sampling Frame	N (%) Invariant Loadings	N (%) Invariant Thresholds
26. Sweden	CBCL: Larsson and Frisk (1999)[b]	1354	6–16	11.2 (3.0)	48	Regional school-based	93 (97)	68 (71)
	YSR: Broberg et al. (2001)[b]	1928	11–16	14.3 (.96)	51	Regional school-based	72 (82)	55 (63)
27. Switzerland (German Speaking)	CBCL: Steinhausen, Metzke, Meier, and Kannenberg (1997)[b]	2060	6–16	11.5 (2.9)	51	Regional school-based	83 (86)	59 (61)
	YSR: Steinhausen et al. (1997)[b]	1132	11–16	13.8 (1.6)	51	Regional school-based	87 (100)[c]	63 (72)
28. Tunisia	CBCL: Chahed (2010)[b]	1187	6–16	10.8 (3.1)	41	Regional school-based	90 (94)	50 (52)
29. Turkey	CBCL: Erol and Simsek (1997)[b]	3790	6–16	11.0 (3.2)	52	National household	75 (78)	52 (54)
	YSR: Erol and Simsek (1997)[b]	1760	11–16	13.7 (1.7)	51	National household	79 (90)	61 (69)
30. United States	CBCL: Achenbach and Rescorla (2001)[b]	1788	6–16	11.3 (3.1)	53	National household	87 (91)	56 (58)
	YSR: Achenbach and Rescorla (2001)[b]	1003	11–16	13.3 (1.7)	53	National household	88 (100)	68 (77)
31. Uruguay	CBCL: Viola, Garrido, and Rescorla (2011)[b]	1374	6–11	8.7 (1.7)	49	National school-based	91 (95)	55 (57)

Note. CBCL = Child Behavior Checklist; YSR = Youth Self-Report.
[a]Unpublished data.
[b]Random sample.

means. Alignment thus offers an innovative way to test complex models across many societies and to calibrate assessment instruments. This line of research could provide clinicians with psychometrically sound instruments for assessing children from diverse backgrounds.

Purpose of This Study

The alignment approach makes it possible to use multisample CFA to test the degree to which complex factor models of emotional, behavioral, and social problems fit data obtained from many societies. Our purpose was to test the multisociety MI of the CBCL/6–18 and YSR syndrome structures across many societies by applying alignment CFA. To our knowledge, this was the first multisample test of the configural, metric, and scalar MI of empirically based syndromes of child and youth psychopathology across dozens of societies. We hypothesized that the CBCL and YSR would demonstrate configural invariance across all tested societies and that more than half of their items would demonstrate approximate metric invariance across at least 80% of societies. Given the stringency of scalar invariance, the complexity of our tested model, and the large number of tested societies, we did not make predictions pertaining to scalar invariance.

METHOD

Assessment Instruments

Data were obtained with the parent-report CBCL/6–18 and self-report YSR, which assess a broad spectrum of emotional, behavioral, and social problems and competencies. Informants rate each problem item as 0 (*not true [as far as you know]*), 1 (*somewhat or sometimes true*), or 2 (*very true or often true*), based on the preceding 6 months. We analyzed the 96 CBCL/6–18 and 88 YSR problem items that were analyzed by Ivanova et al. (2007a, 2007b) and by Rescorla et al. (2012). Except for the Australian and U.S. samples that used the original English forms, parents and youths responded to translations of the CBCL/6–18 and YSR in their own language. The translations use simple wording corresponding to the fifth-grade reading level of the English language CBCL/6–18 and YSR, as verified via back-translations.

Samples

As Table 1 shows, parents rated 61,703 children from 6 to 16 years of age from 30 societies on the CBCL/6–18, and 29,486 youths from 11 to 16 years of age from 19 societies rated themselves on the YSR. Of the 31 societies that were represented, 15 were non-European. Monte Carlo studies of the alignment CFA that we used indicated that parameter bias is minimized and replication coverage approximates 95% when sample sizes approach 1,000/per group (Muthén &

Asparouhov, 2014). Consequently, we selected samples with total sample sizes ranging from around 1,000 (i.e., 990 Romanian CBCLs and 946 Portuguese YSRs) to the largest available (i.e., 4,878 Chinese CBCLs and 2,750 Korean YSRs; Rescorla et al., 2012). All samples were general population samples recruited with random sampling procedures. They were recruited independently by indigenous researchers who followed their institutional requirements for obtaining informed consent. Ivanova et al. (2007a, 2007b) and Rescorla et al. (2012) provided details of the samples.

Tested Models

We tested the first-order eight-factor model tested by Ivanova et al. (2007a, 2007b) and Rescorla et al. (2012). Table 2 lists the items loading on the factors, which were all modeled as intercorrelated first-order factors. Measurement errors were not correlated, and item cross-loadings were not allowed.

Alignment CFA of the Eight-Factor Model

The eight-factor model was tested via alignment CFA implemented in Mplus (Muthén & Muthén, 1998–2016). To avoid low-frequency cells, item scores were dichotomized as 0 versus 1 or 2, and tetrachoric correlations were computed between the dichotomized item ratings. Because of the ordered categorical nature of CBCL/6–18 and YSR item ratings, we used the robust Bayesian estimator rather than the Maximum Likelihood estimator, which is appropriate for normally distributed data. The Mplus (version 7.11 and higher; Muthén & Muthén, 1998–2016) alignment procedure generates invariance results for each parameter based on the largest set of the same parameters across groups that are not statistically different from each other.

RESULTS

The alignment CFA converged without warnings for the CBCL/6–18 and YSR, indicating that the same eight factors were identified across societies (i.e., a common configural model). Table 1 presents results of invariance-testing for item loadings and thresholds organized by society, whereas Table 2 presents the same results organized by item.

CBCL/6–18

Aligned Loadings

The number of invariant CBCL/6–18 loadings was high. By society, the mean number of invariant loadings across the 96 items was 85.7 loadings per society (*SD* = 5.0; 25th percentile = 83, 50th percentile = 86, 75th percentile = 89). The number of invariant loadings ranged from 75 loadings (China and Turkey) to 93 loadings (Sweden and Italy).

By item, the mean number of invariant loadings across the 30 societies was 26.8 (89.3%) loadings per item (*SD* = 2.6; 25th percentile = 25, 50th percentile = 27, 75th percentile = 29). The number of invariant loadings ranged from 17 loadings (*88.Sulks*) to 30 loadings (*21.Destroys others' things; 22.Disobedient at home; 40.Hears things that aren't there; 57.Attacks people; 59.Plays with sex parts in public; 67. Runs away; 70.Sees things that aren't there; 72.Sets fires; 73.Sexual problems; 91.Talks or thinks suicide; 101.Truant*)

Full metric invariance (i.e., invariance of loadings across all societies) was found for 11 items, whereas approximate metric invariance (which we defined as invariance for 80% or more of societies) was found for 74 additional items. For the remaining 11 items, metric invariance was found for 57% or more of societies.

Aligned Thresholds

The proportion of invariant CBCL/6–18 thresholds was smaller than the proportion of invariant CBCL/6–18 loadings. By society, the mean number of invariant thresholds across the 96 items was 53.3 thresholds per society (*SD* = 8.4; 25th percentile = 48, 50th percentile = 52, 75th percentile = 58). They ranged from 39 thresholds (Russia) to 75 thresholds (Israel).

By item, the mean number of invariant thresholds across the 30 societies was 16.6 (55.3%) thresholds per item (*SD* = 4.5; 25th percentile = 13, 50th percentile = 16, 75th percentile = 20). The mean ranged from eight thresholds (*27.Jealous; 104.Unusually loud*) to 28 thresholds (*82. Steals outside of home*).

Eight items (*1.Acts young; 25.Doesn't get along with other kids; 40.Hears things that aren't there; 56g.Vomiting; 59.Plays with sex parts in public; 70.Sees things that aren't there; 82. Steals outside of home; 84.Strange behavior*) demonstrated high invariance of thresholds (i.e., invariant thresholds across 80% to 93% of societies) in the context of high to full invariance of loadings (i.e., invariant loadings across 87% to 100% of societies). These eight items thus approached scalar invariance (i.e., invariance of both loadings and thresholds) across the tested societies. Fifty-six additional items demonstrated threshold invariance for 53% or more of societies and loading invariance for 73% or more of societies.

Across the eight CBCL/6–18 syndromes, the percentage of invariant loadings ranged from 86% (*Anxious/Depressed* and *Aggressive Behavior*) to 93% (*Thought Problems*) with a mean of 90% (*SD* = 3%), and the percentage of invariant thresholds ranged from 46% (*Aggressive Behavior*) to 62% (*Thought Problems*) with a mean of 56% (*SD* = 6%). A one-way analysis of variance (ANOVA) was conducted to compare proportions of invariant loadings and thresholds across

TABLE 2
Invariance Results for Aligned Loadings and Thresholds for CBCL and YSR Items

Syndrome Item	N*(%) of Societies With Invariant Aligned CBCL Loadings*/N *(%) of Societies With Invariant Aligned CBCL Thresholds*	R^2 *Aligned CBCL Loadings*/R^2 *Aligned CBCL Thresholds*	N*(%) of Societies With Invariant Aligned YSR Loadings*/N*(%) of Societies With Invariant Aligned YSR Thresholds*	R^2 *Aligned YSR Loadings*/R^2 *Aligned YSR Thresholds*
Anxious/Depressed				
14. Cries	23 (77)/13 (43)	.22/.68	15 (79)/13 (68)	.00/.71
29. Fears	25 (83)/12 (40)	.40/.56	18 (95)/11 (58)	.14/.11
30. Fears School	27 (90)/16 (53)	.22/.00	16 (84)/7 (37)	.36/.08
31. Fears Doing Bad	24 (80)/14 (47)	.39/.42	15 (79)/14 (74)	.42/.58
32. Must Be Perfect	25 (83)/11 (37)	.54/.40	15 (79)/8 (42)	.19/.20
33. Feels Unloved	27 (90)/22 (73)	.03/.74	16 (84)/16 (84)	.04/.53
35. Feels Worthless	25 (83)/13 (43)	.44/.00	18 (95)/17 (89)	.57/.71
45. Nervous/Tense	27 (90)/10 (33)	.71/.66	14 (74)/11 (58)	.49/.22
50. Fearful/Anxious	26 (87)/12 (40)	.50/.65	14 (74)/12 (63)	.16/.47
52. Feels Too Guilty	29 (97)/18 (60)	.63/.66	16 (84)/13 (68)	.27/.56
71. Self-Conscious	22 (73)/14 (47)	.35/.43	16 (84)/14 (74)	.04/.22
91. Talks/Thinks Suicide[a]	30 (100)/22 (73)	.17/.37	13 (68)/13 (68)	.38/.40
112. Worries	24 (80)/13 (43)	.41/.47	12 (63)/11 (58)	.05/.48
Withdrawn/Depressed				
42. Rather Be Alone	28 (93)/23 (77)	.40/.25	18 (95)/15 (79)	.55/.22
65. Won't Talk	29 (97)/15 (50)	.32/.62	18 (95)/9 (47)	.12/.31
69. Secretive	26 (87)/21 (70)	.39/.67	17 (89)/16 (84)	.53/.44
75. Shy/Timid	27 (90)/13 (43)	.32/.19	18 (95)/10 (53)	.37/.43
102. Lacks Energy	29 (97)/16 (53)	.44/.70	16 (84)/14 (74)	.40/.63
103. Sad	27 (90)/15 (50)	.24/.39	17 (89)/14 (74)	.00/.42
111. Withdrawn	27 (90)/23 (77)	.12/.57	15 (79)/10 (53)	.16/.02
Somatic Complaints				
47. Nightmares	24 (80)/17 (57)	.00/.08	17 (89)/7 (37)	.04/.00
49. Constipated	27 (90)/15 (50)	.00/.27	not on YSR	
51. Dizzy	29 (97)/18 (60)	.00/.21	16 (84)/9 (47)	.32/.12
54. Overtired[b]	29 (97)/15 (50)	.00/.52	19 (100)/10 (53)	.47/.46
56a. Aches/Pains[b]	29 (97)/20 (67)	.57/.78	19 (100)/15 (79)	.63/.78
56b. Headaches	29 (97)/18 (60)	.80/.82	18 (95)/14 (74)	.57/.80
56c. Nausea	26 (87)/15 (50)	.00/.48	16 (84)/13 (68)	.21/.68
56d. Eye problems	26 (87)/14 (47)	.39/.42	17 (89)/11 (58)	.37/.39
56e. Skin problems	29 (97)/11 (37)	.69/.48	18 (95)/9 (47)	.53/.35
56f. Stomachaches	27 (90)/21 (70)	.70/.75	17 (89)/10 (53)	.33/.48
56g. Vomiting	28 (93)/24 (80)	.40/.69	18 (95)/13 (68)	.14/.48
Social Problems				
11. Too Dependent[b]	27 (90)/10 (33)	.34/.57	19 (100)/9 (47)	.45/.34
12. Lonely	27 (90)/16 (53)	.66/.53	17 (89)/14 (74)	.00/.00
25. Doesn't Get Along[b]	26 (87)/26 (87)	.60/.75	19 (100)/17 (89)	.73/.29
27. Jealous	29 (97)/8 (27)	.40/.49	18 (95)/8 (42)	.44/.17

34. Others Out to Get Him/Her	27 (90)/20 (67)	.59/.47	18 (95)/8 (42)	.41/.36
36. Accident-Prone	22 (73)/19 (63)	.29/.04	14 (74)/9 (47)	.13/.40
38. Teased[b]	28 (93)/18 (60)	.00/.58	19 (100)/10 (53)	.54/.68
48. Not Liked	24 (80)/19 (63)	.65/.64	18 (95)/16 (84)	.47/.70
62. Clumsy	25 (83)/16 (53)	.14/.04	17 (89)/11 (58)	.11/.34
64. Prefers Younger Children	23 (77)/16 (53)	.60/.54	16 (84)/9 (47)	.28/.13
79. Speech Problem	29 (97)/23 (77)	.25/.40	18 (95)/8 (42)	.32/.15
Thought Problems				
9. Can't Get Mind Off Thoughts	28 (93)/15 (50)	.31/.34	16 (84)/11 (58)	.47/.64
18. Harms Self	29 (97)/22 (73)	.00/.28	16 (84)/14 (74)	.00/.15
40. Hears Things[a,b]	30 (100)/27 (90)	.09/.31	19 (100)/10 (53)	.53/.52
46. Twitches	29 (97)/16 (53)	.47/.49	17 (89)/11 (58)	.50/.57
58. Picks Skin	24 (80)/9 (30)	.40/.17	18 (95)/11 (58)	.44/.37
59. Plays With Sex Parts In Public[a]	30 (100)/25 (83)	.37/.00	not on YSR	
60. Plays With Sex Parts Too Much	28 (93)/18 (60)	.49/.34	not on YSR	
66. Compulsions	29 (97)/16 (53)	.37/.62	17 (89)/11 (58)	.49/.71
70. Sees things[a]	30 (100)/24 (80)	.31/.62	18 (95)/13 (68)	.10/.67
76. Sleeps Less Than Other Kids	29 (97)/16 (53)	.37/.04	18 (95)/12 (63)	.20/.61
83. Stores Up Things	24 (80)/17 (57)	.12/.45	17 (89)/8 (42)	.29/.23
84. Strange Behavior	28 (93)/25 (83)	.00/.57	18 (95)/14 (74)	.09/.59
85. Strange Ideas	28 (93)/20 (67)	.00/.35	17 (89)/14 (74)	.02/.67
92. Sleepwalking	25 (83)/15 (50)	.41/.11	not on YSR	
100. Sleep Problems	27 (90)/14 (47)	.52/.36	17 (89)/8 (42)	.10/.33
Attention Problems				
1. Acts Young	27 (90)/25 (83)	.62/.72	15 (79)/17 (89)	.14/.51
8. Can't Concentrate	26 (87)/20 (67)	.17/.61	18 (95)/15 (79)	.61/.76
10. Can't Sit Still	27 (90)/18 (60)	.48/.60	18 (95)/10 (53)	.63/.49
13. Confused	27 (90)/13 (43)	.29/.68	18 (95)/13 (68)	.00/.29
17. Daydreams	23 (77)/12 (40)	.12/.18	15 (79)/10 (53)	.46/.44
41. Impulsive[b]	28 (93)/16 (53)	.44/.38	19 (100)/16 (84)	.82/.72
61. Poor Schoolwork	26 (87)/17 (57)	.42/.63	17 (89)/11 (58)	.23/.51
80. Stares Blankly	29 (97)/16 (53)	.32/.58	not on YSR	
Rule-Breaking Behavior				
26. Lacks Guilt	23 (77)/16 (53)	.35/.31	15 (79)/19 (100)	.02/.53
39. Bad Friends	24 (80)/20 (67)	.20/.32	18 (95)/12 (63)	.00/.29
43. Lies/Cheats	28 (93)/15 (50)	.50/.44	15 (79)/9 (47)	.00/.00
63. Prefers Older Kids	24 (80)/13 (43)	.46/.49	16 (84)/12 (63)	.34/.40
67. Runs Away[a]	30 (100)/21 (70)	.34/.00	18 (95)/15 (79)	.29/.00
72. Sets Fires[a]	30 (100)/14 (47)	.42/.46	17 (89)/10 (53)	.30/.11
73. Sex Problems[a]	30 (100)/22 (73)	.20/.40	not on YSR	
81. Steals/Home	28 (93)/22 (73)	.30/.44	18 (95)/11 (58)	.23/.00

(*Continued*)

TABLE 2
(Continued)

Syndrome Item	N*(%) of Societies With Invariant Aligned CBCL Loadings*/N *(%) of Societies With Invariant Aligned CBCL Thresholds*	R^2 *Aligned CBCL Loadings*/R^2 *Aligned CBCL Thresholds*	N*(%) of Societies With Invariant Aligned YSR Loadings*/N*(%) of Societies With Invariant Aligned YSR Thresholds*	R^2 *Aligned YSR Loadings*/R^2 *Aligned YSR Thresholds*
82. Steals/Outside Home	29 (97)/28 (93)	.22/.58	17 (89)/10 (53)	.20/.00
90. Swearing	24 (80)/14 (47)	.00/.24	16 (84)/12 (63)	.04/.22
96. Thinks of Sex Too Much	29 (97)/19 (63)	.24/.56	17 (89)/10 (53)	.00/.46
101. Truant[a]	30 (100)/16 (53)	.29/.22	18 (95)/14 (74)	.53/.50
106. Vandalism	28 (93)/19 (63)	.06/.09	not on YSR	
Aggressive Behavior				
3. Argues	22 (73)/11 (37)	.41/.00	15 (79)/7 (37)	.25/.14
16. Mean	27 (90)/12 (40)	.20/.25	15 (79)/13 (68)	.27/.33
19. Demands Attention	24 (80)/12 (40)	.28/.32	17 (89)/13 (68)	.45/.37
20. Destroys Own Things	28 (93)/14 (47)	.48/.55	18 (95)/14 (74)	.41/.50
21. Destroys Others' Things[a]	30 (100)/22 (73)	.54/.62	17 (89)/14 (74)	.00/.31
22. Disobedient/Home[a]	30 (100)/17 (57)	.61/.53	18 (95)/10 (53)	.35/.45
23. Disobedient/School	29 (97)/12 (40)	.49/.54	16 (84)/11 (58)	.00/.62
37. Fights[b]	29 (97)/14 (47)	.36/.29	19 (100)/0 (0)	.27/.02
57. Attacks[a]	30 (100)/16 (53)	.44/.18	16 (84)/14 (74)	.00/.28
68. Screams	27 (90)/16 (53)	.40/.65	16 (84)/8 (42)	.40/.34
86. Stubborn/Sullen	28 (93)/13 (43)	.35/.14	13 (68)/10 (53)	.33/.37
87. Mood Changes	23 (77)/19 (63)	.24/.55	15 (79)/16 (84)	.46/.71
88. Sulks	17 (57)/10 (33)	.16/.39	not on YSR	
89. Suspicious	20 (67)/11 (37)	.29/.43	12 (63)/13 (68)	.33/.38
94. Teases	21 (70)/9 (30)	.35/.38	17 (89)/13 (68)	.22/.68
95. Temper[b]	29 (97)/16 (53)	.77/.61	19 (100)/9 (47)	.62/.25
97. Threatens	29 (97)/17 (57)	.36/.46	14 (74)/13 (68)	.00/.17
104. Loud	24 (80)/8 (27)	.26/.52	18 (95)/14 (74)	.00/.36

Note: R^2 indicates degree of invariance, with values ranging from 0 to 1. CBCL = Child Behavior Checklist; YSR = Youth Self-Report.

[a]Item met criterion for full metric invariance (invariance of factor loading across all tested societies) for the CBCL.

[b]Item met criterion for full metric invariance for the YSR.

CBCL/6–18 syndromes. ANOVA results indicated an overall difference in the proportion of invariant CBCL/6–18 thresholds, $F(7, 88) = 2.5$, $p < .05$, but Bonferroni post hoc group comparison tests indicated no significant group differences.

YSR

Aligned Loadings

The number of invariant loadings was also high for the YSR. Considered by society, the mean number of invariant loadings across the 88 items was 78.4 loadings per society ($SD = 6.1$; 25th percentile = 75, 50th percentile = 79, 75th percentile = 82). The number of invariant loadings ranged from 65 (Korea) to 88 (United States).

By item, the mean number of invariant loadings per item across the 19 societies was 16.7 (88%) ($SD = 1.6$; 25th percentile = 16, 50th percentile = 17, 75th percentile = 18). The number ranged from 12 loadings (*89.Suspicious; 112. Worries*) to 19 loadings (*11.Too dependent; 25.Doesn't get along with other kids; 37.Gets in fights; 38.Gets teased; 40.Hears things; 41.Impulsive; 54.Overtired; 56a. Aches, pains; 95.Hot temper*).

Full metric invariance was indicated for nine items and approximate metric invariance (as indicated by invariance for 80% or more of societies) for 61 additional items. For 19 other items, metric invariance was indicated for 63% or more of societies.

Aligned Thresholds

Like the pattern found for the CBCL/6–18, the proportion of invariant YSR thresholds was smaller than the proportion of invariant YSR loadings. By society, the mean number of invariant thresholds across the 88 items was 55.0 thresholds per society ($SD = 7.8$; 25th percentile = 50, 50th percentile = 55, 75th percentile = 61). The number of invariant thresholds ranged from 36 (Japan) to 68 thresholds (United States).

By item, the mean number of invariant thresholds across the 19 societies was 11.7 (62%) thresholds per item ($SD = 3.0$; 25th percentile = 10, 50th percentile = 11, 75th percentile = 14). The number of invariant thresholds ranged from zero (*37.Gets in fights*) to 19 (*26.Lacks guilt*).

Nine items (*87.Mood changes; 33.Feels unloved; 69. Secretive; 48.Not liked; 41.Impulsive; 1.Acts young*; 35. *Feels worthless; 25.Doesn't get along; 26.Lacks guilt*) demonstrated high to full invariance of thresholds (i.e., invariant thresholds across 80% to 100% of societies) in the context of high to full invariance of loadings (i.e., invariant loadings across 79% to 100% of societies). These nine items thus approached scalar invariance across the tested societies. Sixty additional items demonstrated threshold invariance for 53% or more of societies and loading invariance for 63% or more of societies.

The percentage of invariant loadings ranged from 80% (*Anxious/Depressed*) to 92% (*Somatic Complaints* and *Social Problems*) with a mean of 89% ($SD = 4\%$), and the percentage of invariant thresholds ranged from 57% (*Social Problems*) to 69% (*Attention Problems*) with a mean of 62% ($SD = 4\%$). One-way ANOVAs were conducted to compare proportions of invariant loadings and thresholds across YSR syndromes. ANOVA results indicated significant syndrome differences in the proportion of invariant loadings, $F(7, 80) = 3.4$, $p < .01$. Bonferroni post hoc group comparison tests indicated that the proportion of invariant loadings for the *Anxious/Depressed* syndrome was significantly smaller than for the *Somatic Complaints*, *Social Problems*, and *Thought Problems* syndromes.

The R^2 Alignment Fit Statistic

Table 2 presents R^2 values for aligned item loadings and thresholds for the CBCL/6–18 and YSR. The R^2 alignment fit statistic indicates the degree of item invariance for an item parameter. Ranging from 0 to 1, the R^2 indicates how much variation in the parameter is due to the variation in the underlying factor mean and the factor's variance rather than to error variance. Higher R^2 values indicate higher invariance, with $R^2 = 1$ meaning that all variance in the parameter is due to the variation of its underlying factor mean and the factor's variance. For the CBCL/6–18, the R^2 for loadings ranged from .00 to .80 ($M = .34$, $SD = .20$) and the R^2 for thresholds ranged from .00 to .82 ($M = .43$, $SD = .21$). For the YSR, the R^2 for loadings ranged from .00 to .82 ($M = .29$, $SD = .21$) and the R^2 for thresholds ranged from .00 to .80 ($M = .39$, $SD = .22$).

The R^2 thus indicated variability in the invariance of item loadings and thresholds for both the CBCL/6–18 and YSR. Together with proportions of invariant thresholds and loadings, the R^2 also provided information about sources of invariance across societies. For example, YSR item *12.Lonely* demonstrated relatively high proportions of invariant loadings and thresholds (89% and 74%, respectively). Its R^2 of 0.00 for both loadings and intercepts suggested that for societies that were noninvariant, the variability of these parameters was due to influences other than factor means and factor variances, that is, influences that were included in measurement error.

Aligned Loadings and Thresholds by Society Characteristics

To test whether item invariance was associated with cultural characteristics, we first classified societies according to the culture clusters derived in the Global Leadership and Organizational Behavior Effectiveness study (GLOBE; House, Hanges, Javidan, Dorfman, & Gupta, 2004). We then compared the number of invariant loadings and thresholds between clusters using one-way ANOVAs. Societies were assigned to the GLOBE clusters as follows: *Anglo*: Australia, United States; *Middle East*: Turkey, Tunisia; *Confucian*: Hong Kong, Korea, China, Japan, Singapore; *Eastern Europe*: Croatia, Greece, Kosovo, Lithuania, Poland, Romania, Russia; *Germanic Europe*: Germany,

Netherlands, Switzerland (German-speaking region); *Latin America*: Brazil, Colombia, Peru, Uruguay; *Latin Europe*: France, Israel, Italy, Portugal, Spain; *Nordic Europe*: Finland, Sweden; *Southern Asia*: Iran. We found no significant differences in numbers of invariant loadings and thresholds between cluster groups for the CBCL/6–18 or YSR.

DISCUSSION

Multisample CFAs of parents' ratings of 61,703 children 6 to 16 years of age from 30 societies and self-ratings by 29,486 children 11 to 16 years of age from 19 societies supported the eight-syndrome model that was previously supported by separate CFAs of each sample (Ivanova et al., 2007a, 2007b; Rescorla et al., 2012). Our multisample CFAs of large samples of non-normally distributed item ratings with robust estimators were made possible by the new alignment method developed by Asparouhov and Muthén (2014). We applied the alignment method to the 30 largest CBCL/6–18 and 19 largest YSR samples used by Ivanova et al. (2007a, 2007b) and Rescorla et al. (2012), as each of these samples approached or exceeded the total sample size of 1,000 that a Monte Carlo simulation study indicated as the minimum needed for a fair test of the method (Muthén & Asparouhov, 2014).

For most of the 96 CBCL/6–18 and 88 YSR items that we tested, item loadings were equivalent across most of the compared societies, supporting high levels of metric invariance. Results indicated full metric invariance for 11 CBCL/6–18 and nine YSR items. Approximate metric invariance, which we defined as invariance across 80% or more of societies, was found for an additional 74 CBCL/6–18 and 61 YSR items. For all remaining items, metric invariance was indicated for more than half the societies, 57% or more for the CBCL/6–18 and 63% or more for the YSR.

A substantial, though smaller, proportion of items demonstrated some degree of scalar invariance, that is, invariance of both item loadings and thresholds. Eight CBCL/6–18 and nine YSR items approached scalar invariance by demonstrating approximate to full invariance of thresholds in the context of approximate to full invariance of loadings. For 56 additional CBCL/6–18 items and 60 additional YSR items, threshold invariance was found for 53% or more of the tested societies. For these items, loading invariance was also found for 73% or more and 63% or more of societies, respectively.

Proportions of invariant loadings and thresholds were similar across syndromes, although the YSR *Anxious/Depressed* syndrome had fewer invariant loadings than the YSR *Somatic Complaints, Social Problems*, and *Thought Problems* syndromes. The item-factor associations for the YSR *Anxious/Depressed* syndrome were more variable across societies than the item-factor associations for the YSR *Somatic Complaints, Social Problems*, and *Thought Problems* syndromes. Because items composing the YSR *Anxious/Depressed* syndrome assess harder-to-define internal problems than problems composing the YSR *Somatic Complaints, Social Problems*, and *Thought Problems* syndromes, ratings of these problems may be affected more by socioculturally determined interpretations of internal experiences than responses to other problems. The YSR *Anxious/Depressed* syndrome may thus be more influenced by sociocultural factors in the measurement of psychopathology than other syndromes.

There is growing consensus that complete scalar invariance is unattainable for complex models tested across many groups (Van De Schoot, Schmidt, De Beuckelaer, Lek, & Zondervan-Zwijnenburg, 2015). Our results and results of other large-scale multicultural studies of MI (e.g., Nagengast & Marsh, 2014; Zercher, Schmidt, Cieciuch, & Davidov, 2015) support this position. In our study, full scalar invariance would mean equivalence of 8,640 parameters for the CBCL/6–18 (288 parameters across 30 societies) and 5,073 for the YSR (267 parameters across 19 societies). Besides the sheer statistical improbability of such equivalence, it is especially improbable in multicultural research, where the vicissitudes of translation may prevent items from being identically construed by all respondents in every society.

China and Turkey had the fewest invariant CBCL/6–18 item loadings (75 loadings, or 78%), whereas Italy and Sweden had the most (93 loadings, or 97%). Korea had the fewest invariant YSR item loadings (65 loadings, or 73%), whereas the United States had the most (88 loadings, or 100%). Even for societies with the fewest invariant loadings, most loadings were invariant. Consequently, we do not recommend procedures for adjusting the CBCL/6–18 and YSR in these societies. Furthermore, we did not find any associations between patterns of invariance and sociocultural differences, such as those characterizing societies grouped according to the culture clusters (e.g., Confucian) identified in the GLOBE study (House et al., 2004).

The alignment CFA enabled us to advance beyond testing syndrome structures in individual societies to formally testing measurement invariance in multisociety analyses. This is a significant advance in large-scale multicultural MI testing of empirically based syndromes. Although alignment CFA does not directly test model fit, the multisociety convergence of the eight-factor model supported the configural invariance of the CBCL/6–18 syndromes across 30 societies and of the YSR syndromes across 19 societies. Moreover, the results indicated full or approximate metric invariance (invariance of item loadings) for 89% of the CBCL/6–18 items and 79% of the YSR items. This suggests that most items of the CBCL/6–18 and YSR are equivalent or approach equivalence in their correlations with the

corresponding underlying constructs of psychopathology across the tested societies.

The eight empirically derived syndromes thus demonstrated configural and, generally, metric validity across societies that differ markedly in language, culture, political and social systems, ethnic composition, educational systems, and childrearing practices. The consistency of our findings between parent- and self-reports further supported the validity of the eight syndromes as phenotypic structures. Given the multitude of cultural, linguistic, and methodological factors that could contribute to differences between societies and between informants, our findings argue for the multicultural robustness of the tested syndromes.

Taken together, our findings support the eight syndromes as robust phenotypic structures of psychopathology across the tested societies. Our findings also suggest that most CBCL/6–18 and YSR items are similarly representative of the corresponding syndromes across societies and that many CBCL/6–18 and YSR items are quite comparable across all or most tested societies.

Limitations of the Findings

Nevertheless, our results should be interpreted in light of certain limitations. To optimize the performance of the alignment approach, we needed large samples, which limited the number of qualifying societies. Because the alignment approach cannot yet accommodate hierarchical structural relations in the context of robust estimation, we could not test hierarchical relations among the eight syndromes. Finally, the alignment approach needs further testing. For example, in estimating model parameters, the alignment approach assumes that there are few large, noninvariant parameters and many approximately invariant parameters. The performance of the alignment approach needs to be tested with this assumption and its alternative (many medium-sized noninvariant parameters) under different data conditions.

An important goal of MI testing is to determine where societies stand on an underlying factor of psychopathology by comparing latent factor means across societies. If cross-society differences in latent factor means are found, norms can be used to take account of these differences. However, because scalar invariance is often unattainable, large-scale multisociety studies rarely advance to this stage of MI testing.

A fundamental innovation of alignment CFA is that it estimates latent factor means without requiring full scalar invariance. However, we cannot yet recommend accounting for differences in alignment-based latent factor means when calculating CBCL/6–18 and YSR scale scores for two reasons. First, because the alignment approach is so new, more work needs to be done to understand its performance under different conditions. Second, although all our samples were large, they utilized different random sampling procedures. Therefore, estimates of latent factor means based on these samples may not be fully generalizable within each society. Our findings do indicate that alignment offers a promising methodology to achieve intricate cross-societal calibration of assessment instruments. Our findings also indicate the power of factor-analytic methods to identify robust phenotypic structures of psychopathology.

Clinical Significance of the Findings

Although CFAs conducted separately on data from each society have previously supported aggregations of CBCL/6–18 and YSR items into eight syndromes (configural invariance), our findings are the first to support the syndromes via simultaneous multisociety analyses. Because multisociety analyses are much more comprehensive than single-society analyses, the present study markedly strengthens the evidence for the multicultural robustness of the syndromes. The findings should increase clinicians' confidence in using the syndromes as meaningful clinical constructs for assessing children of diverse backgrounds. For most items, the results also supported metric invariance, that is, comparable loadings of items on their respective factors across societies. Overall, our study provides evidence that the CBCL/6–18 and YSR syndromes offer clinicians psychometrically robust ways to conceptualize problems reported for children of diverse backgrounds.

ACKNOWLEDGMENTS

Drs. Ivanova, Achenbach, and Rescorla receive support from the nonprofit Research Center for Children, Youth, and Families, which publishes the Child Behavior Checklist for Ages 6-18 and the Youth Self-Report.

REFERENCES

Abad, J., Forns, M., & Gomez, J. (2002). Emotional and behavioral problems as measured by the YSR: Gender and age differences in Spanish adolescents. *European Journal of Psychological Assessment*, *18*(2), 149–157. doi:10.1027//1015-5759.18.2.157

Achenbach, T. M. (2009). *Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications*. Burlington: University of Vermont, Research Center of Children, Youth & Families.

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

Achenbach, T. M., & Rescorla, L. A. (2015). *Multicultural guide for the ASEBA forms and profiles for ages 1½-59*. Burlington: University of Vermont: Research Center for Children, Youth, and Families.

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495–508. doi:10.1080/10705511.2014.919210

Broberg, A. G., Ekeroth, K., Gustafsson, P. A., Hansson, K., Hagglof, B., Ivarsson, T., & Larsson, B. (2001). Self-reported competencies and problems among Swedish adolescents: A normative study of the YSR. *European Child and Adolescent Psychiatry*, *10*(3), 186–193. doi:10.1007/s007870170025

Chahed, M. (2010). Contributions of the Child Behavior Checklist and Youth Self-Report to the study of behavioral and emotional problems among Tunisian children and adolescents (Unpublished doctoral dissertation). Universite de Paris Ouest, Nanterre, France.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*(4), 858–900. doi:10.1037/a0038498

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, *131*(4), 483–509. doi:10.1037/0033-2909.131.4.483

De Los Reyes, A., & Ohannessian, C. M. (2016). Introduction to the Special Issue: Discrepancies in adolescent-parent perceptions of the family and adolescent adjustment. *Journal of Youth and Adolescence*, *45*, 1957–1972. doi:10.1007/s10964-016-0533-z

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology*, *9*, 123–149. doi:10.1146/annurev-clinpsy-050212-185617

Domuta, A. (2004). *Evaluarea sindromului ADHD la varsta prescolara ADHD assessment in preschoolers* (Unpublished doctoral dissertation). Babes-Bolyai University, Cluj Napoca, Romania.

Döpfner, M., Plück, J., Berner, W., Fegert, J. M., Huss, M., Lenz, J., … Lehmkuhl, G. (1997). Psychic disturbances of children and adolescents in Germany: Results of representative study: Method, age, gender, and rater effects. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, *25*(4), 218–233.

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

Fombonne, E., & Vermeersch, S. (1997). Children of the GAZEL cohort: 1-Prevalence and patterns of service utilization for emotional and behavioural symptoms. *Revue d'Epidemiologie et de Sante Publique*, *45*(1), 29–40.

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

Frigerio, A., Cattaneo, C., Cataldo, M., Schiatti, A., Molteni, M., & Battaglia, M. (2004). Behavioral and emotional problems among Italian children and adolescents aged 4 to 18 years as reported by parents and teachers. *European Journal of Psychological Assessment*, *20*(2), 124–133. doi:10.1027/1015-5759.20.2.124

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the Confirmatory Factor Analysis framework. *Medical Care*, *44*, S78–S94. doi:10.1080/00220979809601406

Gross, T. J., Fleming, C. B., Mason, W. A., & Haggerty, K. P. (2017). Alabama Parenting Questionnaire – 9: Longitudinal measurement invariance across parents and youth during the transition to high school. *Assessment*, *24*(5), 646–659. doi:10.1177/1073191115620839

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

Hewitt, N., Jaimes, S., Vera, L. A., & Villa, M. C. (2012). [Child Behavior Checklist]. Unpublished raw data.

House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies* (3rd ed.). Thousand Oaks, CA: Sage Publications.

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

Ivanova, M. Y., Achenbach, T. M., Dumenci, L., Rescorla, L. A., Almqvist, F., Weintraub, S., … Verhulst, F. C. (2007). Testing the 8-syndrome structure of the Child Behavior Checklist in 30 societies. *Journal of Clinical Child and Adolescent Psychology*, *36*(3), 405–417. doi:10.1080/15374410701444363

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bilenberg, N., … Verhulst, F. C. (2007a). The generalizability of the Youth Self-Report syndrome structure in 23 societies. *Journal of Consulting and Clinical Psychology*, *75*(5), 729–738. doi:10.1037/0022-006X.75.5.729

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Dumenci, L., Almqvist, F., Bathiche, M., … Verhulst, F. C. (2007b). Testing the generalizability of the Teacher's Report Form syndromes in 20 societies. *School Psychology Review*, *36*, 468–483.

Janssens, A., Goossens, L., Van Den Noortgate, W., Colpin, H., Verschueren, K., & Van Leeuwen, K. (2015). Parents' and adolescents' perspectives on parenting: Evaluating conceptual, structure, measurement invariance and criterion validity. *Assessment*, *22*(4), 473–489. doi:10.1177/1073191114550477

Kuramoto, H., Kanbayashi, Y., Nakata, Y., Fukui, T., Mukai, T., & Negishi, Y. (2002). Standardization of the Japanese version of the Youth Self Report (YSR). *Japanese Journal of Child and Adolescent Psychiatry*, *43*, 17–32.

Larsson, B., & Frisk, M. (1999). Social competence and emotional/behaviour problems in 6–16 year-old Swedish school children. *European Child and Adolescent Psychiatry*, *8*(1), 24–33. doi:10.1007/s007870050080

Leung, P. W. L., Kwong, S. L., Tang, C. P., Ho, T. P., Hung, S. F., Lee, C. C., … Liu, W. S. (2006). Test-retest reliability and criterion validity of the Chinese version of CBCL, TRF, and YSR. *Journal of Child Psychology and Psychiatry*, *47*(9), 970–973. doi:10.1111/j.1469-7610.2005.01570.x

Liu, X., Kurita, H., Guo, C., Miyake, Y., Ze, J., & Cao, H. (1999). Prevalence and risk factors of behavioral and emotional problems among Chinese children aged 6 through 11 years. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(6), 708–715. doi:10.1097/00004583-199906000-00018

Marsh, H. W., Guo, J., Nagengast, B., Parker, P. D., Asparouhov, T., Muthén, B., & Dicke, T. (2017). What to do when scalar invariance fails: The extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychological Methods*. Epub ahead of print. doi:10.1037/met0000113

Millsap, R. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

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

Moreira, P. A. S., & Oliveira, J. T. (2012). [The Psychology of Positive Development Research Center database: Youth Self-Report]. Unpublished raw data.

Muthén, B., & Asparouhov, T. (2014). IRT studies of many groups: The alignment method. *Frontiers in Psychology*, *5*, 1–7. doi:10.3389/fpsyg.2014.00978

Muthén, L. K., & Muthén, B. O. (1998-2014). Mplus. In *The comprehensive modelling program for applied researchers: User's guide* (Vol. 7).

Muthén, L. K. & Muthén, B. O. (1998–2016). Mplus [Statistical Software]. Los Angeles, CA: Muthén & Muthén.

Nagengast, B., & Marsh, H. W. (2014). Motivation and engagement in science around the globe: Testing measurement invariance with multigroup SEMs across 57 countries using PISA 2006. In L. Rutkowski, M. Von Davier, & D. Rutkowski (Eds.), *A Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 317–344). New York, NY: Taylor & Francis.

Oh, K. J., Hong, K. E., & Lee, H. R. (1997b). *Korean-Youth Self Report (K-YSR)*. Seoul, Korea: Jungang Aptitude Research Center.

Oh, K. J., Lee, H., Hong, K.-E., & Ha, E. H. (1997a). *Korean-child behavior checklist*. Seoul, Korea: Jungang Aptitude Research Center.

Pisa, C., & Maggiolini, A . (2011). [Youth Self-Report]. Unpublished raw data.

Pomalima, R., & Valverde, J. (2009). [Child Behavior Checklist and Youth Self-Report]. Unpublished raw data.

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*(12), 1273–1283. doi:10.1016/j.jaac.2012.09.012

Rocha, M. M. (2012). *Evidencias de validade do "Inventario de Autoavaliac ao para Adolescentes" (YSR/2001) para a popilacao Brasileira* (Unpublished doctoral dissertation). Universidade de Sao Paulo, Sao Paulo, Brazil.

Rocha, M. M., Rescorla, L. A., Emerich, D. R., Silvares, E. F., Borsa, J. C., Araujo, L. S., … Assis, S. G. (2013). Behavioural/emotional problems in Brazilian children: Findings from parents' reports on the Child Behavior Checklist. *Epidemiology and Psychiatric Sciences*, *22*(4), 329–338. doi:10.1017/S2045796012000637

Roussos, A., Karantanos, G., Richardson, C., Hartmann, C., Karajiannis, D., Kyprianos, S., … Zoubou, V. (1999). Achenbach's Child Behavioral Checklist and Teacher's Report Form in a normative sample of Greek children 6-12 years old. *European Child and Adolescent Psychiatry*, *8*(3), 165–172. doi:10.1007/s007870050125

Roussos, A. C., Francis, K., Zoubou, V., Kiprianos, S., Prokopiou, A., & Richardson, C. (2001). The standardization of Achenbach's Youth Self-Report in Greece in a national sample of high school students. *European Child and Adolescent Psychiatry*, *10*(1), 47–53. doi:10.1007/s007870170046

Rudan, V., Begovac, I., Szirovicza, L., Filipovi, O., & Skocic, M. (2005). The Child Behavior Checklist, Teacher Report Form and Youth Self-Report problem scales in a normative sample of Croatian children and adolescents aged 7-18. *Collegium Antropologicum*, *29*(1), 17–26.

Sawyer, M. G., Kosky, R. J., Graetz, B. W., Arney, F., Zubrick, S. R., & Baghurst, P. (2000). The National Survey of Mental Health and Wellbeing: The child and adolescent component. *Australian and New Zealand Journal of Psychiatry*, *34*(2), 214–220. doi:10.1046/j.1440-1614.2000.00729.x

Shahini, M., Rescorla, L. A., Ahmeti, A. P., Begovac, I., Dobrean, A., Marcovic, J., … Zukauskiene, R. (2015a). Parent-reported behavioural and emotional problems in Albanian Kosovar children. *Epidemiology and Psychiatric Sciences*, *24*(3), 266–273. doi:10.1017/S204579601400016X

Shahini, M., Rescorla, L., Wancata, J., & Ahmeti, A. (2015b). Mental health problems in Kosovar adolescents: Results from a national mental health survey. *Neuropsychiatrie*, *29*(3), 125–132. doi:10.1007/s40211-015-0155-9

Steinhausen, H. C., Metzke, C. W., Meier, M., & Kannenberg, R. (1997). Behavioral and emotional problems reported by parents for ages 6 to 17 in a Swiss epidemiological study. *European Child and Adolescent Psychiatry*, *6*(3), 136–141. doi:10.1007/BF00538985

Van De Schoot, R., Schmidt, P., De Beuckelaer, A. D., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology*, *6*, 1064–1068. doi:10.3389/fpsyg.2015.011064

Verhulst, F. C., Van Der Ende, J., & Koot, H. M. (1996). *Handleiding voor de CBCL/4-18*. Rotterdam, the Netherlands: Erasmus University, Department of Child and Adolescent Psychiatry.

Viola, L., Garrido, G., & Rescorla, L. A. (2011). Testing multicultural robustness of the Child Behavior Checklist in a national epidemiological sample in Uruguay. *Journal of Abnormal Child Psychology*, *39*(6), 897–908. doi:10.1007/s10802-011-9500-z

Wang, J., Zhang, Y. Q., & Leung, P. W. L. (2005). The application of CBCL, YSR and TRF for 12-18 years of age students in Beijing. *Chinese Mental Health Journal*, *19*(9), 621–625.

Weintraub, S. (2004). [Child Behavior Checklist]. Unpublished raw data.

Wolanczyk, T. (2003). *Zaburzenia emocjonalne i behawioralne u dzieci i mlodziezy w Polsce*. Warszaw, Poland: Akademia Medyczna w Warszawie.

Woo, B. S., Ng, T. P., Fung, D. S., Chan, Y. H., Lee, Y. P., Koh, J. B., & Cai, Y. (2007). Emotional and behavioural problems in Singaporean children based on parent, teacher, and child reports. *Singapore Medical Journal*, *48*(12), 1100–1106.

Zercher, F., Schmidt, P., Cieciuch, J., & Davidov, E. (2015). The comparability of the universalism value over time and across countries in the European Social Survey: Exact vs. approximate measurement invariance. *Measurement Invariance*, *6*, 1–11. doi:10.3389/fpsyg.2015.00733

Zilber, N., Auerbach, J., & Lerner, Y. (1994). Israeli norms for the Achenbach Child Behavior Checklist: Comparison of clinically-referred and non-referred children. *Israeli Journal of Psychiatry and Related Sciences*, *31*(1), 5–12.

Zukauskiene, R., Ignataviciene, K., & Daukantaite, D. (2003). Subscale scores of the Lithuanian version of CBCL: Preliminary data on the emotional and behavioural problems in childhood and adolescence. *European Child and Adolescent Psychiatry*, *12*(3), 136–143. doi:10.1007/s00787-003-0321-7

Zukauskiene, R., & Kajokiene, I. (2004, October). *Child Behavior Checklist profile types in a random sample and in a child psychiatric clinic population in Lithuania*. Paper presented at the Workshop on Child and Adolescent Assessment and Research Methodology, Vilnius, Lithuania.