


RESEARCH ARTICLE

The generalizability of Older Adult Self-Report (OASR) syndromes of psychopathology across 20 societies

Masha Y. Ivanova¹  | Thomas M. Achenbach¹ | Leslie A. Rescorla² | Lori V. Turner¹ | Julie A. Dumas¹ | Vera Almeida^{3,4} | Meltem Anafarta-Sendag⁵  | Ieva Bite⁶ | Dorret I. Boomsma⁷ | J. Carlos Caldas^{3,8} | John W. Capps⁹ | Yi-Chuen Chen¹⁰ | Paola Colombo¹¹  | Margareth da Silva Oliveira¹² | Anca Dobrea¹³ | Nese Erol¹⁴ | Alessandra Frigerio¹¹ | Yasuko Funabiki¹⁵ | Reda Gedutienė¹⁶ | Halldór S. Guðmundsson¹⁷  | Min Quan Heo¹⁸  | Young Ah Kim¹⁹ | Tih-Shih Lee¹⁸ | Manuela Leite³ | Jianghong Liu²⁰ | Jasminka Markovic²¹ | Monika Misiec²² | Marcus Müller²³ | Kyung Ja Oh²⁴ | Verónica Portillo-Reyes²⁵ | Wolfgang Retz^{23,26} | Sandra B. Sebre⁶ | Shupeng Shi²⁰ | Sigurveig H. Sigurðardóttir¹⁷  | Roma Šimulionienė¹⁶ | Elvira Sokoli²⁷ | Tanja Tomasevic²⁸ | Jacqueline M. Vink²⁹ | Ewa Zasepa²²

¹Department of Psychiatry, Larner College of Medicine, University of Vermont, Burlington, VT, USA

²Department of Psychology, Bryn Mawr College, Philadelphia, PA, USA

³Instituto de Investigação e Formação Avançada em Ciências e Tecnologias da Saúde, CESPU, Gandra, Portugal

⁴UCIBIO/REQUIMTE, MedTec-Laboratório de Tecnologia Farmacêutica, Departamento de Ciências do Medicamento, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal

⁵Department of Psychology, Ufuk University, Ankara, Turkey

⁶Department of Psychology, University of Latvia, Riga, Latvia

⁷Department of Biological Psychology, Vrije University, Amsterdam, The Netherlands

⁸CINTESIS—Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

⁹Department of Psychology, The University of Texas at El Paso, El Paso, TX, USA

¹⁰Department of Psychology, National Chung Cheng University, Chiayi County, Taiwan

¹¹Child Psychopathology Unit, Scientific Institute Eugenio Medea, Bosisio Parini, Italy

¹²Faculty of Psychology, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brazil

¹³Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania

¹⁴Department of Child and Adolescent Mental Health, Ankara University, Ankara, Turkey

¹⁵Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

¹⁶Department of Psychology, Klaipėda University, Klaipėda, Lithuania

¹⁷Faculty of Social Work, University of Iceland, Reykjavik, Iceland

¹⁸Neuroscience and Behavioural Disorders Programme, Duke-National University of Singapore Medical School, Singapore

¹⁹Huno Inc., Seoul, Republic of Korea

²⁰Department of Family and Community Health, University of Pennsylvania School of Nursing, Philadelphia, PA, USA

²¹Private Practice, Novi Sad, Serbia

²²Department of Psychology, Maria Grzegorzewska University, Warsaw, Poland

²³Department of Psychiatry and Psychotherapy, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

²⁴Department of Psychology, Yonsei University, Seoul, Republic of Korea

²⁵Departamento de Psicología, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Mexico

²⁶Institute for Forensic Psychology and Psychiatry, Saarland University Homburg/Saar, Homburg, Germany

²⁷Department of Psychology, University of Tirana, Tirana, Albania

²⁸Faculty of Medicine, Institute of Public Health of Vojvodina, University of Novi Sad, Novi Sad, Serbia

²⁹Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

Correspondence

Masha Y. Ivanova, 1 S. Prospect Street,
Burlington, VT 05401.
Email: masha.ivanova@uvm.edu

Objectives: As the world population ages, psychiatrists will increasingly need instruments for measuring constructs of psychopathology that are generalizable to diverse elders. The study tested whether syndromes of co-occurring problems derived from self-ratings of psychopathology by US elders would fit self-ratings by elders in 19 other societies.

Methods/design: The Older Adult Self-Report (OASR) was completed by 12 826 adults who were 60 to 102 years old in 19 societies from North and South America, Asia, and Eastern, Northern, Southern, and Western Europe, plus the United States. Individual and multigroup confirmatory factor analyses (CFAs) tested the fit of the seven-syndrome OASR model, consisting of the Anxious/Depressed, Worries, Somatic Complaints, Functional Impairment, Memory/Cognition Problems, Thought Problems, and Irritable/Disinhibited syndromes.

Results: In individual CFAs, the primary model fit index showed good fit for all societies, while the secondary model fit indices showed acceptable to good fit. The items loaded strongly on their respective factors, with a median item loading of .63 across 20 societies, and 98.7% of the loadings were statistically significant. In multigroup CFAs, 98% of items demonstrated approximate or full metric invariance. Fifteen percent of items demonstrated approximate or full scalar invariance, and another 59% demonstrated scalar invariance across more than half of societies.

Conclusions: The findings supported the generalizability of OASR syndromes across societies. The seven syndromes offer empirically based clinical constructs that are relevant for elders of different backgrounds. They can be used to assess diverse elders and as a taxonomic framework to facilitate communication, services, research, and training in geriatric psychiatry.

KEYWORDS

alignment CFA, cross-cultural, elders, empirical syndromes, psychopathology

1 | INTRODUCTION

Over the next few decades, the number of older adults is expected to grow annually by 27.1 million, reaching 1.6 billion or 16.7% of the total world population by 2050.¹ Because older adults face significantly more physical and mental health problems than the rest of the population, this demographic shift will present formidable challenges to the world's health care systems. For example, combined with increasing globalization of world communities, such rapid aging of the global population will require increased capacity to assess the mental health needs of elders of diverse backgrounds. The growing population of immigrant elders may pose additional capacity challenges, as their mental health needs may be especially high.²

Mental health professionals will increasingly need assessment instruments for measuring constructs of psychopathology that are generalizable to diverse elders. However, as pointed out by Mindt et al,³ there is a paucity of assessment instruments that are appropriate for culturally and linguistically diverse older adults. Most assessment instruments have been developed in rather similar high-income societies and may not be generalizable to other societies. Before an assessment instrument developed in one society is used in another society, it is necessary to test whether it measures the same constructs in both societies. For example, does an assessment instrument for depression developed in one society also measure depression in the other society? It is also necessary to test whether the instrument measures constructs in the same way in both societies. For example, does a score of X on the instrument reflect the same severity of

depression in both societies? Failure to establish this equivalence between societies may lead to inaccurate assessment results and misguided treatment planning for patients in the new society.

The generalizability of constructs measured by the same instrument across societies is usually tested via Confirmatory Factor Analysis (CFA) using the framework of measurement invariance (MI).⁴ Measurement invariance posits that when an individual obtains a particular item score on a measure of psychopathology, that score is a function of several influences, such as the person's standing on the underlying "true" factor of psychopathology, as well as systematic (eg, social norms) and unsystematic (ie, error) influences that are not associated with the underlying factor of psychopathology. When a factor model representing the factor structure of an instrument is fit to data obtained with this instrument in different societies, the measurement invariance framework is translated into testable hypotheses about different components of the factor model. *Configural invariance* implies that certain items load on the same factors across societies. *Metric invariance* implies that items have similar loadings on their specified factors across societies (ie, that the item reflects the same level of the "true" latent factor of psychopathology across societies). Finally, *scalar invariance* implies that item intercepts (or thresholds for categorical data) are equivalent across societies (ie, that systematic influences on item ratings that are not associated with the underlying factors are the same across societies). These hypotheses are hierarchical, with each level depending on the prior level(s).

1.1 | Previous factor analytic studies of elder psychopathology

Standardized assessment instruments used in geriatric psychiatry include the Geriatric Depression Scale (GDS),⁵ the Hospital Anxiety and Depression Scale (HADS),⁶ the Beck Depression Inventory (BDI),⁷ and the Hamilton Depression Rating Scale (HAM-D).⁸ These instruments are excellent candidates for testing measurement invariance across societies because they are relatively short and easy to administer to diverse informants and in diverse settings. While we are not aware of published tests of the measurement invariance of these instruments across societies using CFA, several international studies have used exploratory factor analysis (EFA) to test the factor structures of the GDS and HADS.

Developed in the United States, the GDS is a 30-item self-rating instrument that was created specifically to assess depression in the elderly. It does not include physical symptoms of depression (ie, aches and pains), as they may lack specificity in this population. It uses a simplified response format to minimize the cognitive burden on elder respondents. EFAs of the GDS performed on self-ratings by US elders yielded five factors designated as *sad mood*, *positive mood*, *lack of energy*, *agitation*, and *social withdrawal*.⁹

Two international studies found a multifactor structure for the GDS. Kim et al¹⁰ administered the GDS to 782 elders participating in the Korean Longitudinal Study on Health and Aging (LSHA; mean age = 75.03) and 106 Korean elders receiving treatment at a

Key Points

- Mental health professionals around the world increasingly need empirically supported instruments for older adult psychopathology.
- The study tested the degree to which syndromes (groupings of co-occurring problems) derived from self-ratings of psychopathology by older adults in the United States would fit self-ratings by older adults in 19 other societies.
- The syndromes fit the data well across the tested societies or provided an accurate model of how the elders' mental health problems grouped together.
- The syndromes can be used to assess diverse older adults and as a taxonomic framework in geriatric psychiatry.

university psychiatry clinic (mean age = 74.05). EFA indicated five factors that were designated as *sad mood and agitation*, *positive mood*, *lack of energy*, *cognitive inefficiency*, and *social withdrawal*. The Korean factor structure was similar to the US factor structure, except for the combination of the US *sad mood* and *agitation* factors into a single Korean factor and the derivation of an additional Korean *cognitive inefficiency* factor. Salamero and Marcos¹¹ factor analyzed self-ratings by 234 Spanish elders (mean age = 77.5) who resided mostly in residential institutions (184 or 79%). While EFA indicated eight factors, only the first three were interpretable (*depressed mood*, *cognitive impairment*, and *social withdrawal and avoidance*).

Two other international studies found a single-factor structure for the GDS. Ertan and Eker¹² conducted EFA on GDS ratings by 276 Turkish elders, including 179 community dwellers (mean age = 66.6) and 97 retirement home residents (mean age = 76.7). EFAs indicated two factors, which were designated as *depressive affect and thought content* and *other symptoms*. Chiu et al¹³ factor analyzed GDS ratings by 183 community dwelling elders in Hong Kong, including 113 elders recruited from a senior center (mean age = 73.8) and 80 elders recruited from a university psychiatry clinic (mean age = 71.4). EFA indicated a dominant first factor comprising 28 of the 30 items that was designated as *general depressed mood* and a two-item grouping that was difficult to interpret. In summary, while the Korean and US GDS factor structures were quite similar, Spanish, Turkish, and Hong Kong structures were quite different.

The HADS⁶ is a 14-item self-report questionnaire that was developed in the United Kingdom to detect anxiety and depressive symptoms in adults receiving medical care. Like the GDS, the HADS omits physical symptoms associated with mental health conditions. As its name suggests, the factor structure of the HADS comprises two first-order factors designated as *anxiety* and *depression*.¹⁴ Spinhoven et al tested the factor structure of the HADS using self-ratings by Dutch elders recruited from a general population registry.¹⁵ EFAs supported the British two-factor structure for younger (N = 1901, ages 57-65; mean age = 61.3) and older (N = 3293, ages >65; mean age = 74.3) subsamples.

Taking a multisociety approach, Prince et al factor analyzed data from 14 centers in 11 European countries to compare depressive symptoms among elders in different societies.¹⁶ Because different assessment instruments were used in different centers, the authors employed probabilistic modeling and expert opinion to construct a scale that was generalizable to all centers, which they named EURO-D. As not all EURO-D items had counterparts on all instruments, Prince et al imputed the data for the missing items. EFAs conducted separately in each society indicated two factors designated as *depressed affect* and *motivation* in each society. In addition, factors designated as *somatic*, *irritability*, and *guilt* emerged in subgroups of societies. However, because differences between the instruments accounted for 63% of the between-center variance in the EURO-D scores, it is difficult to draw firm conclusions from these findings.¹⁷

To summarize, several factor analytic studies have tested the factor structures of standardized assessment instruments of elder psychopathology in different societies. With the exception of Prince et al, these studies were conducted in single societies. All these studies assessed narrow spectrum constructs of psychopathology, such as anxiety and depression.

1.2 | The present study

We tested the cross-society generalizability of seven factors derived from self-ratings of psychopathology by US elders using the Older Adult Self-Report (OASR)¹⁸ in 19 societies. The OASR was developed as a broad measure of elder psychopathology, as well as cognitive and adaptive functioning, substance use, illnesses, and disabilities.¹⁸ The OASR is a standardized, self-report questionnaire that can be self-administered in under 20 minutes on paper or online, or read aloud by an interviewer. OASR items were generated from research and clinical work with older adults and from suggestions by older adults and people who work with them, as described by Achenbach, Newhouse, and Rescorla.¹⁸

The seven OASR factors (referred to as "syndromes" because they comprise co-occurring problems) are designated as *Anxious/Depressed*, *Worries*, *Somatic Complaints*, *Functional Impairment*, *Memory/Cognition Problems*, *Thought Problems*, and *Irritable/Disinhibited*. The syndromes were derived via EFAs and CFAs from self-ratings by 1048 US 60- to 98-year-olds and span a broad spectrum of elder psychopathology. Analogous syndromes were derived from ratings by collateral informants such as spouse/partner or adult offspring using the parallel Older Adult Behavior Checklist (OABCL).¹⁸

Ivanova et al tested the fit of the seven-syndrome OASR model to self-ratings by 352 adults who were 60 to 102 years old in community and residential care settings in Porto, Portugal.¹⁹ The OASR model showed good fit, as indicated by the primary model fit index, the root mean square error of approximation (RMSEA),²⁰ and acceptable fit, as indicated by secondary fit indices, the comparative fit index (CFI)²¹ and Tucker-Lewis index (TLI).²² Loadings were statistically significant for 98% of the items, indicating that they measured the syndromes well. The results thus supported the OASR syndrome structure among Portuguese elders.

Extending the single-society CFA study of the Portuguese OASR,¹⁹ the present study used a multisociety design to test the seven-syndrome OASR model in 20 societies from Asia, North and South America, and Europe. We used translations of the same standardized assessment instrument in each society, which allowed us to avoid the methodological challenges of having to combine data from different instruments.^{16,17}

We tested how well the seven-syndrome OASR model fit the data obtained in the 20 societies using two CFA approaches. First, to test the configural measurement invariance of the seven-syndrome model across societies, we tested the model separately in each society using single-society CFA. We predicted that the seven-syndrome model would demonstrate configural invariance in each society. Second, to test higher levels of measurement invariance, we used multi-sample alignment CFA by fitting the seven-syndrome model simultaneously in all societies. Alignment CFA was developed to overcome the computational issues associated with testing complex models such as ours across a large number of societies using traditional CFA.^{23,24} Marsh et al²⁵ found that alignment CFA outperformed traditional measurement invariance testing by yielding more accurate parameter estimates, including estimates of latent factor means. We predicted that the seven-syndrome model would demonstrate metric invariance for most items but did not expect to find scalar invariance because of the general consensus that it is unattainable when testing complex models such as ours across a large number of societies.^{26,27}

2 | METHODS

2.1 | Instrument

The OASR measures a broad spectrum of emotional, behavioral, social, and cognitive problems using 113 items written at a fifth grade reading level (eg, "I cry a lot" and "I worry too much about my memory").¹⁸ The OASR also has items for assessing personal strengths (eg, "I make good use of my time"), relationships with friends and spouse/partner, substance use (alcohol, tobacco, drugs), illnesses and disabilities, and residential accommodations. It is part of a system of transdiagnostic dimensional assessment forms spanning ages 1.5 to 90+ years, the Achenbach System of Empirically Based Assessment (ASEBA).²⁸ The ASEBA provides tools for the multi-informant assessment of psychopathology and adaptive functioning in terms of empirically derived taxonomic constructs scored in relation to age, gender, self-versus-collateral informant, and multicultural norms.

The OASR was translated and back-translated by bilingual native speakers of the languages of non-Anglophone societies (all, except the United States). The back-translations were checked extensively for content and readability against the English-language originals by both the indigenous researchers and T.M. Achenbach. Elders rated their emotional, behavioral, social, and cognitive problems on 113 OASR items that describe a broad spectrum of problems. Each item is rated 0 = *not true*, 1 = *somewhat or sometimes true*, or 2 = *very true or often true*, based on the preceding 2 months.

2.2 | Samples

Table 1 describes the samples and sampling procedures. The OASR was completed by 12 826 adults who were 60 to 102 years old in 20 societies. Following their respective institutional ethics guidelines, indigenous researchers arranged to have OASRs completed by consenting participants.

2.3 | Tested model

Using EFAs and CFAs, Achenbach et al derived the seven-syndrome OASR model on self-ratings by 1048 US elders who were selected from a larger sample for having total problem scores (sum of ratings on the 113 items) that were at or above the median for the national survey sample.¹⁸ The sample included participants in a national household survey, plus residents of 29 residential and day facilities and outpatients in four mental health/substance abuse services.¹⁸

Supporting the construct and criterion-related validity of the syndromes, OASR/OABCL syndrome scores have been significantly ($p < .01$) associated with elders' cognitive performance, psychopathology, and adaptive functioning on multiple measures.^{18,31} These include the Neuropsychiatric Inventory,³² Mini-Mental State

Exam,³³ Clock Drawing Test,³⁴ Alzheimer's Disease Assessment Schedule,³⁵ Geriatric Depression Scale,⁵ Clinical Dementia Rating Scale,³⁶ Dementia Severity Rating,³⁷ Trail Making Test,³⁸ and Activities of Daily Living.³⁹ OASR/OABCL scales have also discriminated significantly between elders diagnosed as having Alzheimer versus affective disorders^{18,31} and between clinically referred versus non-referred Brazilian elders.⁴⁰

With the OASR items listed by syndrome (factor), Table 2 presents our tested model. Table 2 lists the 97 items that are scored on the syndromes, with each item assigned to only one syndrome. An additional 16 of the 113 OASR items are counted in the Total Problems score (the sum of all items) but do not load on any of the seven syndromes. The name of each syndrome reflects the content of the items comprising the syndrome. In the tested model, the factors representing the syndromes were modeled as first-order correlated factors, with no hierarchical relations between factors specified.

2.4 | Data analyses

For consistency with the Achenbach et al analyses, we deleted OASRs with eight or more unrated items from each sample and transformed

TABLE 1 Descriptive information about the samples and sampling procedures

Society	Reference	N	Mean Age (SD)	Age Range	% Male	Sample	Completion Rate
1. Albania	Sokoli ^a	892	70.02 (7.47)	60-93	48	Community	Not available
2. Brazil	da Silva Oliveira ^a	306	70.62 (8.23)	60-93	34	Community	86%
3. China	Liu, Shi, Sun, et al ^b	686	69.70 (7.35)	60-99	45	Community	93%
4. Germany	Müller, Turner, Tüscher, et al ^a	300	66.9 (5.26)	60-84	44	Regional household	57%
5. Iceland	Guðmundsson, Sigurðardóttir ^a	340	71.09 (7.91)	60-96	45	National household	45%
6. Italy	Frigerio, Sangiorgio, Colombo, et al ^a	312	75.48 (9.38)	60-97	42	Regional household	32%
7. Japan	Funabiki ^a	1693	75.41 (9.43)	60-99	46	National household	Not available
8. Korea	Kim et al ²⁹	1032	70.68 (7.45)	60-95	42	National community	47%
9. Latvia	Sebre, Bite ^a	301	72.13 (8.09)	60-96	33	Community	90% ^a
10. Lithuania	Šimulionienė, Gedutienė, Rugevičius et al ^a	328	71.91 (8.49)	60-97	36	National household	82%
11. Mexico	Portillo-Reyes, Capps ^a	292	71.91 (8.49)	60-97	36	Community	60%
12. Netherlands	Willemsen et al ³⁰	2149	68.74 (5.23)	60-92	48	Community	54%
13. Poland	Zasępa, Misiec ^a	304	70.69 (8.76)	60-97	31	Community	75%
14. Portugal	Caldas, Almeida, Leite ^a	352	72.90 (8.95)	60-102	39	Community	90%
15. Romania	Dobrea ^a	779	68.49 (6.43)	60-89	36	National	78%
16. Serbia	Markovic, Tomasevic, Milijasevic, et al ^a	303	70.97 (8.52)	60-97	32	Community	65%
17. Singapore	Yeo, Heo, Sng, et al ^a	300	66.82 (4.66)	60-87	35	Community	67%
18. Taiwan	Chen ^a	318	70.61 (7.45)	60-93	48	National household	95%
19. Turkey	Anafarta-Sendag, Erol ^a	697	69.71 (7.58)	60-98	37	Regional household	NA
20. US	Achenbach et al ¹⁸	1142	73.66 (8.18)	60-102	37	National household	90%

^aUnpublished data collected in 2018.

^bUnpublished data collected in 2017.

TABLE 2 Descriptive statistics for factor loadings and invariance results for aligned loadings and thresholds across 20 societies by OASR syndrome. Items are designated with summary labels for their content.

Syndromes and Items	Mean Factor Loading (SD)	Median Factor Loading	N(%) of Societies with Invariant Aligned OASR Loadings	N(%) of Societies with Invariant Aligned OASR Thresholds
<i>Anxious/Depressed</i>	.65(.10)	.64	18(90)	11(58)
8. Can't get mind off thoughts	.57(.07)	.57	19(95)	13(65)
9. Can't sit still	.42(.14)	.43	16(80)	9(45)
11. Lonely	.63(.09)	.63	20(100)	15(75)
13. Cries	.58(.11)	.59	19(95)	11(55)
14. Concerned about getting old	.58(.09)	.58	20(100)	10(50)
21. Worries about future	.58(.15)	.62	15(75)	12(60)
23. Feels too guilty	.69(.09)	.70	20(100)	15(75)
26. Fears	.57(.12)	.55	17(85)	11(55)
28. Fears doing bad	.66(.12)	.63	18(90)	13(65)
32. Feels worthless	.77(.08)	.78	17(85)	11(55)
34. Restless, fidgety	.69(.14)	.69	17(85)	11(55)
40. Nervous	.73(.06)	.73	19(95)	13(65)
42. Lacks self-confidence	.71(.06)	.71	20(100)	6(30)
45. Fearful, anxious	.74(.08)	.73	17(85)	13(65)
47. Guilty conscience	.70(.11)	.71	19(95)	9(45)
62. Self-conscious	.64(.09)	.64	17(85)	14(70)
91. Thinks about past	.64(.06)	.65	19(95)	12(60)
93. Sad	.77(.07)	.77	19(95)	16(80)
100. Worries	.66(.11)	.68	16(80)	9(45)
109. Concerned about death	.63(.08)	.63	17(85)	12(60)
<i>Worries</i>	.51(.14)	.54	18(91)	11(56)
51. Worries about appearance	.47(.23)	.54	18(90)	10(50)
72. Worries about family	.36(.23)	.41	17(85)	7(35)
89. Concerned about neatness	.31(.16)	.30	18(90)	7(35)
90. Trouble sleeping	.55(.06)	.54	19(95)	10(50)
101. Wakes up early	.36(.14)	.36	17(85)	13(65)
102. Worries about health	.64(.12)	.67	19(95)	16(80)
117. Get too tired	.63(.10)	.63	19(95)	13(65)
121. Feels burdensome	.73(.06)	.73	19(95)	14(70)
<i>Somatic Complaints</i>	.63(.10)	.63	18(92)	13(67)
5. Too much medication	.52(.09)	.50	17(85)	8(40)
33. Feels sick	.78(.10)	.80	16(80)	12(60)
46. Dizzy	.69(.11)	.69	18(90)	10(50)
49a. Aches	.61(.13)	.61	19(95)	12(60)
49b. Headaches	.58(.10)	.58	20(100)	14(70)
49c. Nausea	.74(.12)	.74	18(90)	18(90)
49d. Eye problems	.55(.08)	.55	18(90)	13(65)
49e. Itching, rashes	.48(.11)	.50	19(95)	12(60)
49f. Stomachaches	.63(.07)	.63	19(95)	15(75)
49g. Vomits	.66(.12)	.64	19(95)	18(90)
49h. Heart pounds	.63(.10)	.62	20(100)	10(50)
49i. Numbness	.66(.09)	.67	18(90)	15(75)
49j. Short of breath	.66(.07)	.66	20(100)	17(85)
103. Nightmares	.69(.10)	.69	19(95)	14(70)

(Continues)

TABLE 2 (Continued)

Syndromes and Items	Mean Factor Loading (SD)	Median Factor Loading	N(%) of Societies with Invariant Aligned OASR Loadings	N(%) of Societies with Invariant Aligned OASR Thresholds
<i>Functional Impairment</i>	.65(.10)	.67	18(90)	13(69)
3. Difficulty getting things done	.63(.08)	.66	18(90)	13(65)
10. Too dependent	.67(.04)	.67	19(95)	15(75)
16. Sits around	.62(.10)	.61	18(90)	14(70)
29. Difficulty preparing meals	.55(.13)	.56	18(90)	9(45)
54. Poor task performance	.76(.08)	.75	18(90)	9(45)
55. Clumsy	.76(.08)	.77	17(85)	16(80)
68. Sleeps more than most	.49(.10)	.48	20(100)	15(75)
92. Lacks energy	.73(.09)	.74	16(80)	17(85)
104. Trouble dressing	.67(.16)	.73	16(80)	15(75)
106. Trouble bathing	.67(.14)	.68	19(95)	16(80)
111. Soiling accidents	.58(.14)	.63	20(100)	13(65)
<i>Memory/Cognition Problems</i>	.67(.08)	.67	18(90)	12(63)
7. Can't concentrate	.67(.06)	.67	18(90)	14(70)
12. Confused	.82(.07)	.81	19(95)	9(45)
20. Forgets names	.55(.08)	.55	20(100)	9(45)
52. Can't finish things	.73(.08)	.76	17(85)	11(55)
69. Trouble with decisions	.72(.08)	.72	16(80)	11(55)
70. Can't talk	.67(.09)	.65	19(95)	14(70)
110. Can't remember	.67(.08)	.67	18(90)	15(75)
114. Forgets if not written down	.51(.12)	.53	18(90)	20(100)
122. Worries about memory	.68(.08)	.69	18(90)	11(55)
<i>Thought Problems</i>	.59(.13)	.56	18(94)	13(67)
24. Jealous	.66(.13)	.64	17(85)	14(70)
27. Bad relations with neighbors	.49(.16)	.47	18(90)	14(70)
30. Feels no one cares	.75(.07)	.75	19(95)	14(70)
31. Feels others out to get him/her	.74(.13)	.74	20(100)	10(50)
36. Hears things	.59(.18)	.56	20(100)	17(85)
38. Rather be alone	.43(.10)	.41	20(100)	9(45)
41. Twitches	.57(.09)	.57	19(95)	13(65)
57. Repeats acts	.59(.14)	.56	20(100)	14(70)
58. No friends	.64(.11)	.60	20(100)	13(65)
60. Secretive	.44(.15)	.45	14(70)	10(50)
61. Sees things	.58(.17)	.55	20(100)	16(80)
74. Strange behavior	.61(.12)	.56	20(100)	16(80)
75. Strange ideas	.59(.18)	.55	20(100)	15(75)
77. Mood changes	.71(.06)	.70	18(90)	14(70)
99. Withdrawn	.55(.14)	.51	18(90)	12(60)
<i>Irritable/Disinhibited</i>	.60(.12)	.63	18(92)	11(59)
2. Argues	.49(.10)	.48	18(90)	7(35)
15. Mean	.67(.10)	.68	19(95)	13(65)
18. Seeks attention	.57(.09)	.57	19(95)	10(50)
19. Damages things	.67(.19)	.68	20(100)	15(75)

(Continues)

TABLE 2 (Continued)

Syndromes and Items	Mean Factor Loading (SD)	Median Factor Loading	N(%) of Societies with Invariant Aligned OASR Loadings	N(%) of Societies with Invariant Aligned OASR Thresholds
22. Doesn't get along	.71(.10)	.71	19(95)	11(55)
25. Gets along badly with family	.68(.09)	.66	17(85)	15(75)
35. Wants own way	.27(.18)	.22	16(80)	11(55)
37. Impulsive	.64(.09)	.63	19(95)	12(60)
39. Does things others don't like	.65(.11)	.63	20(100)	12(60)
43. Not liked	.71(.11)	.74	17(85)	7(35)
59. Screams	.68(.11)	.65	19(95)	11(55)
65. Shows off	.46(.14)	.46	19(95)	14(70)
67. Irresponsible	.72(.10)	.72	20(100)	16(80)
76. Stubborn	.68(.09)	.66	19(95)	13(65)
79. Suspicious	.66(.10)	.64	18(90)	6(30)
83. Talks too much	.44(.10)	.44	20(100)	16(80)
84. Irritates people	.73(.10)	.76	17(85)	17(85)
85. Loses temper	.65(.11)	.62	19(95)	8(40)
86. Thinks about sex	.37(.20)	.36	18(90)	11(55)
94. Loud	.50(.16)	.51	17(85)	14(70)

Note: Values in italics are descriptive statistics for syndromes. For syndrome means and SDs, they were calculated as means of mean loadings and of SDs across societies. For syndrome medians, they were calculated as medians of median loadings across societies. For alignment values, they were calculated as means of the corresponding values for the items comprising the syndromes.

unrated items on retained OASRs to zero. The median percent of excluded OASRs per society was 0.33, ranging from 0.00 (Japan, Korea, Mexico, Portugal, Romania, Serbia, Taiwan) to 4.44 (Albania) percent. Like Achenbach et al, we transformed the 0-1-2 item ratings to 0 vs 1 (transforming 2s into 1s), and computed tetrachoric correlations between the items.¹⁸

Because our data violated assumptions of multivariate normality, we used robust estimators: the WLSMV for the single-sample CFAs and the Bayes estimator for the multisample alignment CFA. All analyses were carried out in Mplus.⁴¹ For single-sample CFAs, the RMSEA was our primary index of model fit because Yu and Muthén identified it as the best performing fit index for the WLSMV estimator, with $\leq .06$ indicating good fit.⁴² The CFI and TLI were considered secondary to the RMSEA. Hu and Bentler proposed $>.95$ as a criterion for good model fit.⁴³ However, Marsh et al criticized this criterion as too stringent, thereby risking unjustified rejection of well-defined complex models.⁴⁴ Because our model was complex, we followed Marsh's recommendations by adopting a more liberal criterion of $>.90$ for good model fit and $.80$ to $.90$ for acceptable fit.

Multisample CFA was carried out using alignment CFA.^{23,24} Alignment CFA estimates the entire factor model without requiring scalar invariance by allowing modest parameter noninvariance. Because alignment models can absorb some noninvariance, they estimate all model parameters simultaneously, rather than terminating the estimation process and requiring post hoc sequential model modification, as done in nonalignment CFA.

3 | RESULTS

3.1 | Single-society CFA

The seven-syndrome model converged for all 20 societies (including the United States). The RMSEA ranged from .018 (Serbia) to .032 (Singapore), indicating good fit for all societies (see Table 3). The RMSEA equaled .022, .023, and .028 at the 25th, 50th, and 75th percentiles, respectively. CFI and TLI values were very similar within societies (Pearson $r > .99$). CFI/TLI values ranged from .821/.816 for Singapore to .957/.956 for China. For all societies, CFIs and TLIs thus indicated acceptable to good fit, with 12 (60%) societies falling in the good fit category.

As Table 3 shows, the median of factor loadings in each society ranged from .59 (Mexico and Latvia) to .79 (China), with an overall median of .62. The OASR items thus demonstrated strong loadings on their assigned factors for each society. Most loadings were also strong when considered by item across the 20 societies. Across all societies, the median item loadings ranged from .22 (item 35. *Wants own way*) to .81 (item 12. *Confused*), with an overall median of .63. By syndrome, the overall median item loadings ranged from .54 for *Worries* to .67 for *Functional Impairment and Memory/Cognition Problems* (Table 2).

Of the 1940 item loadings (97 items * 20 societies), 1914 (98.7%) were statistically significant. For seven societies, all items had statistically significant loadings on their specified factors (Table 3). For Germany, Iceland, and Lithuania, one item had a

nonsignificant loading. For Mexico, Poland, Portugal, Serbia, and Turkey, two items had nonsignificant loadings. For Albania, Italy, and Latvia, three items had nonsignificant loadings. Finally, four items had nonsignificant loadings for Brazil. Of the 26 nonsignificant loadings, seven were for item 35. *Wants own way*; five for item 72. *Worries about family*; four for item 86. *Thinks about sex*; three for item 89. *Concerned about neatness*; two each for items 51. *Worries about appearance* and 101. *Wakes up early*; and one each for items 27. *Bad relations with neighbors*, item 36. *Hears things*, and item 90. *Trouble sleeping*.

For 19 societies, all items were identified. For Taiwan, two items were unidentified (eg, had negative residual item variances): item 36. *Hears things* and item 49c. *Nausea*. Only two out of 1940 tested items (<.01%) were thus unidentified.

3.2 | Alignment CFA

The 97 items had invariant loadings on a mean of 18.4 (92%) of the 20 societies. The number of invariant loadings ranged from 14 societies (60. *Secretive*) to 20 societies (11. *Lonely*; 14. *Concerned about getting old*; 19. *Damages things*; 20. *Forgets names*; 23. *Feels too guilty*; 31. *Feels others out to get him/her*; 36. *Hears things*; 38. *Rather be alone*; 39. *Does things others don't like*; 42. *Lacks self-confidence*; 49b. *Headaches*; 49h. *Heart pounds*; 49j. *Shortness of breath*; 57. *Repeats acts*; 58. *No friends*; 61. *Sees things*; 68. *Sleeps more than most*; 74. *Strange*

behavior; 75. *Strange ideas*; 67. *Irresponsible*; 83. *Talks too much*; 111. *Soiling accidents*). Full metric invariance (ie, invariance of loadings across all societies) was found for 22 items, and approximate metric invariance (which we defined as invariance across 80-99% of societies) was found for 73 additional items. For the remaining two items, metric invariance was found for 70% and 75% of societies. Overall, approximate to full metric invariance was found for 95 (98%) of the items.

The number of invariant thresholds was smaller than the number of invariant loadings. The 97 thresholds had invariant thresholds on a mean of 12.6 (63%) of societies. The number of invariant thresholds ranged from six societies (42. *Lacks self-confidence*; 79. *Suspicious*) to 20 societies (114. *Forgets if not written down*). Full invariance of thresholds was found for one item, and approximate invariance of thresholds (which we again defined as invariance for 80-99% of societies) was found for 14 additional items. For all 15 items that demonstrated approximate or full invariance of thresholds, the loadings were all also approximately to fully invariant (36. *Hears things*; 49c. *Nausea*; 49g. *Vomits*; 49j. *Short of breath*; 55. *Clumsy*; 61. *Sees things*; 67. *Irresponsible*; 74. *Strange behavior*; 83. *Talks too much*; 84. *Irritates people*; 92. *Lacks energy*; 93. *Sad*; 102. *Worries about health*; 106. *Trouble bathing*; 114. *Forgets if not written down*). These items thus demonstrated approximate scalar invariance across the tested societies. Fifty-eight (60%) additional items demonstrated scalar invariance across more than half of societies. With one exception, these were found in the context of

TABLE 3 CFA results

Society	RMSEA (95% CI)	CFI	TLI	Items with Nonsignificant Loadings	Empirically Unidentified Items	Median Factor Loading
1. Albania	.028 (.027-.029)	.876	.873	Items 51, 72, 89		.60
2. Brazil	.019 (.015-.022)	.903	.900	Items 51, 72, 89, 101		.61
3. China	.024 (.023-.026)	.957	.956			.79
4. Germany	.023 (.019-.025)	.843	.839	Item 35		.61
5. Iceland	.022 (.020-.025)	.890	.887	Item 35		.64
6. Italy	.020 (.016-.023)	.911	.908	Items 27, 35, 86		.60
7. Japan	.027 (.026-.027)	.928	.926			.78
8. Korea	.030 (.029-.031)	.877	.873			.62
9. Latvia	.022 (.018-.024)	.906	.903	Item 35, 36, 86		.59
10. Lithuania	.022 (.019-.025)	.927	.925	Item 35		.62
11. Mexico	.020 (.016-.023)	.926	.924	Item 35, 72		.59
12. Netherlands	.023 (.022-.023)	.920	.918			.61
13. Poland	.023 (.020-.025)	.917	.915	Item 72, 90		.64
14. Portugal	.031 (.029-.033)	.869	.866	Item 72, 86		.66
15. Romania	.026 (.025-.027)	.910	.908			.62
16. Serbia	.018 (.014-.021)	.948	.947	Item 35, 86		.64
17. Singapore	.032 (.030-.034)	.821	.816			.63
18. Taiwan	.030 (.028-.033)	.875	.872		Items 36, 49c	.69
19. Turkey	.025 (.024-.026)	.899	.897	Items 89, 101		.61
20. US	.022 (.021-.023)	.911	.908			.63

approximate to full metric invariance, indicating that 59% of items demonstrated approximate to full scalar invariance across a large proportion of societies.

Across the OASR syndromes, the percentage of invariant loadings ranged from 90.3% (*Anxious/Depressed*) to 94.3% (*Thought Problems*), and the percentage of invariant thresholds ranged from 56.3% (*Worries*) to 69.1% (*Functional Impairment*).

4 | DISCUSSION

We tested how well the US seven-syndrome OASR model fit self-ratings in 19 societies. The OASR model was originally derived from self-ratings by US elders and tested in this study on data obtained in societies from North and South America, Asia, and Eastern, Northern, Southern, and Western Europe.

When tested individually in each society, our findings supported the configural invariance of the OASR. The model converged for all societies, and the primary model fit index indicated good fit, while the secondary model fit indices indicated acceptable to good fit for all societies. The items loaded strongly on their respective factors, with the cross-society median item loading being .63. When considered by syndrome, the overall median loadings ranged from .54 (*Worries*) to .67 (*Functional Impairment* and *Memory/Cognition Problems*), with 98.7% being statistically significant. Of the 26 loadings that were not statistically significant, 16 (62%) were for only three items: 35. *Wants own way* (*Irritable/Disinhibited* syndrome), 72. *Worries about family* (*Worries* syndrome), and 86. *Thinks about sex* (*Irritable/Disinhibited* syndrome). These three items may not be as generalizable across societies as indices of their respective syndromes as the other OASR items. For items 35 and 72, the relatively low item loadings may have been due to the high prevalence of the symptoms they measure. Endorsed by 77% and 69% of the overall sample, respectively, the symptoms of *wanting one's way* and *worrying about family* may have been too typical of older adults to warrant their association with latent constructs of psychopathology. For item 86, the reason may have been its low prevalence rate (15.7%), which may have been insufficient to establish a robust item-factor correlation.

Multisample CFAs indicated approximate to full metric invariance for 95 (98%) of the items. This suggested that OASR items generally reflect the same degree of the underlying construct of psychopathology across societies. Also, 15 items approached scalar invariance by showing approximate to full invariance of both thresholds and loadings. Fifty-seven (59%) additional items demonstrated approximate to full scalar invariance across more than half of the tested societies. This indicated that many OASR items are affected by comparable systematic factors across societies. These factors include both the underlying constructs of psychopathology and other nonrandom factors (eg, societal influences and sampling variations). We were surprised to find this much evidence of scalar invariance, as it is typically unattainable in multicultural studies of complex models.²⁵⁻²⁷

4.1 | Implications of the findings

Our findings have significant implications for understanding the structure of elder psychopathology. The strong evidence for the configural and metric invariance of the seven-syndrome OASR model supports its generalizability across the 20 tested societies. Although evidence of its scalar invariance was modest, it is probably as strong as can be expected for such a complex model tested across so many societies. Our prior alignment CFA study of an eight-syndrome model of child psychopathology yielded similar results.⁴⁵ We found strong evidence for configural and metric invariance but modest evidence for scalar invariance in ratings by 61 703 parents in 30 societies using the Child Behavior Checklist for Ages 6-18 (CBCL)⁴⁶ and in ratings by 29 486 youths in 19 societies using the Youth Self Report (YSR).⁴⁶

The present findings are consistent with evidence for the generalizability of the seven-syndrome model derived from collateral ratings of elder psychopathology on the OABCL across 11 societies.⁴⁷ They are also consistent with our previous CFA findings for the Adult Self-Report for Ages 18-59, for which an empirically derived syndrome model was supported by data from 29 societies.⁴⁸

Rescorla and colleagues compared distributions of scores on the seven syndrome scales across the tested societies, finding small to medium effects of society.⁴⁹ Based on the distribution of the Total Problems scores (sum of all problem items), Achenbach and Rescorla constructed multicultural norms for societies with relatively low, medium, or high scale scores.⁵⁰ Software is available to generate profiles of elders' scores on the seven syndromes in relation to age, gender, and multicultural norms.

Our findings also have significant clinical implications for psychiatrists and other health professionals. The seven syndromes offer a concise set of empirically based clinical constructs spanning a broad spectrum of problems reported for elders diverse backgrounds. While psychiatric assessment of elders often focuses on depression and dementia, mental health issues in later life are diverse and multidimensional. The seven OASR syndromes capture dimensions that include anxiety, worry, somatic complaints without known medical cause, functional impairment, irritability, and disinhibition, as well as depression, memory problems, and other cognitive and thought problems. Together with scales for adaptive functioning, personal strengths, and substance use, the seven syndromes and six DSM-oriented scales assessed by the OASR offer clinicians a comprehensive picture of adaptive and maladaptive aspects of functioning.

Initial assessment and subsequent monitoring of older adult psychopathology and functioning are included in core competencies of the geriatric psychiatry subspecialty established by the US Accreditation Council of Graduate Medical Education.⁵¹ As a practical tool for assessment of elder emotional, behavioral, social, and cognitive functioning, the OASR can help geriatric psychiatry trainees achieve their professional milestones in assessment and monitoring. Because the OASR syndromes are empirically derived taxonomic constructs, they can also sharpen trainees' understanding of the structure of self-reported geriatric psychopathology.

As argued by Eyre, Baune, and Lavretsky,⁵² geriatric psychiatry must innovate to address the approaching surge in needs for services. Eyre et al especially recommend prevention of psychiatric disorders through resilience-building interventions. As a practical, normed instrument for broad-spectrum assessment of psychopathology and adaptive functioning, the OASR can assist in identifying elders at risk for psychiatric disorders and in documenting aspects of their adaptive functioning that need strengthening.

4.2 | Limitations

A limitation of this study is that we could not test the generalizability of the OASR in all societies, nor randomly select societies from all societies. However, the diversity of the included societies with respect to numerous sociocultural and geopolitical factors that could have affected the data (language, religion, economic development, political structure, societal views on aging, systems of elder care) argues for the wide generalizability of the syndrome model. Another limitation is imposed by CFA, which tests a single specified syndrome model. Other configurations of problems might be found via other analytic or assessment methods. Finally, although the number of OASRs omitted for having eight or more unrated items and the number of retained OASRs having unrated items were very small, it is possible that other ways of addressing missing data could have produced somewhat different results.

5 | CONCLUSION

As far as we know, our study was the first multisociety study of elder psychopathology to assess a broad spectrum of emotional, behavioral, social, and cognitive problems using the same standardized instrument. Our findings suggest that we can be reasonably confident that the seven OASR syndromes measure similar constructs and the OASR items perform similarly as indicators of these constructs across the tested societies. The OASR and OABCL can be used for harmonized multi-informant assessment in clinical practice, research, and training, and their empirical syndromes can be used as a taxonomic framework in geriatric psychiatry. We plan further research on the factors that might affect the multisociety generalizability of empirically based syndromes of psychopathology across the lifespan.

DATA AVAILABILITY STATEMENT

The data used in this study are available from the corresponding author upon reasonable request.

ORCID

Masha Y. Ivanova  <https://orcid.org/0000-0002-1537-0982>

Meltem Anafarta-Sendag  <https://orcid.org/0000-0002-4810-893X>

Paola Colombo  <https://orcid.org/0000-0003-4687-5913>

Halldór S. Guðmundsson  <https://orcid.org/0000-0002-5917-5206>

Min Quan Heo  <https://orcid.org/0000-0002-6970-9136>

Sigurveig H. Sigurðardóttir  <https://orcid.org/0000-0002-5052-887X>

REFERENCES

1. He W, Goodkind D, Kowal PR. *An Aging World: 2015*. Washington, DC: United States Census Bureau; 2016.
2. Dolberg P, Sigurðardóttir SH, Trummer U. Ageism and older immigrants. In: Ayalon L, Tesch-Römer C, eds. *Contemporary Perspectives on Ageism*. Vol 19. Cham: Springer; 2018:177-191.
3. Mindt MR, Arentoft A, Coulehan K, et al. Neuropsychological evaluation of culturally/linguistically diverse older adults. In: Ravdin L, Katzen H, eds. *Handbook on the Neuropsychology of Aging and Dementia. Clinical Handbooks in Neuropsychology*. Cham: Springer; 2019: 25-48.
4. Millsap R. *Statistical approaches to measurement invariance*. New York, NY: Routledge; 2012.
5. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1982-83;17:37-49.
6. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67:361-370.
7. Beck AT, Steer RA, Brown GK. *BDI-II: Beck Depression Inventory Manual*. 2nd ed. San Antonio, TX: Psychological Corp; 1996.
8. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56-62.
9. Sheikh JI, Yesavage JA, Brooks JO, et al. Proposed factor structure of the Geriatric Depression Scale. *Int Psychogeriatr*. 1991;3:23-28.
10. Kim JY, Park JH, Lee JJ, et al. Standardization of the Korean version of the Geriatric Depression Scale: reliability, validity, and factor structure. *Psychiatry Investig*. 2008;5:232-238.
11. Salamero M, Marcos T. Factor study of the Geriatric Depression Scale. *Acta Psychiatr Scand*. 1992;86:283-286.
12. Ertan T, Eker E. Reliability, validity, and factor structure of the Geriatric Depression Scale in Turkish elderly: are there different factor structures for different cultures? *Int Psychogeriatr*. 2000;12:163-172.
13. Chiu HF, Lee HC, Wing YK, Kwong PK, Leung CM, Chung DW. Reliability, validity, and structure of the Chinese Geriatric Depression Scale in a Hong Kong context: a preliminary report. *Singapore Med J*. 1994;35:477-480.
14. Cosco TD, Doyle F, Ward M, McGee H. Latent structure of the hospital anxiety and depression scale: a 10-year systematic review. *J Psychosom Res*. 2012;72:180-184.
15. Spinhoven PH, Ormel J, Sloekers PP, et al. A validation study of the Hospital Anxiety and Depression Scale (HADS) in different groups of Dutch subjects. *Psychol Med*. 1997;27:363-370.
16. Prince MJ, Reischies F, Beekman ATF, et al. Development of the EURO-D scale—a European Union initiative to compare symptoms of depression in 14 European centres. *Br J Psychiatry*. 1999;174: 330-338.
17. Prince MJ, Beekman ATF, Deeg DJH, et al. Depression symptoms in late life assessed using the EURO-D scale. *Br J Psychiatry*. 1999;174: 339-345.
18. Achenbach TM, Newhouse PA, Rescorla LA. *Manual for the ASEBA Older Adult Forms & Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families; 2004.
19. Ivanova MY, Achenbach TM, Leite M, et al. Empirically derived dimensional syndromes of self-reported psychopathology: cross-cultural comparisons of Portuguese and US elders. *Int J Geriatr Psychiatry*. 2018;33:695-702.
20. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Newbury Park, CA: Sage; 1993:136-162.
21. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107(2):238-246.

22. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973;38:1-10.
23. Asparouhov T, Muthén B. Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary J*. 2014;21:495-508.
24. Muthén B, Asparouhov T. IRT studies of many groups: the alignment method. *Frontiers Psychol*; 2014;12:978-985.
25. Marsh HW, Guo J, Parker PD, et al. What to do when scalar invariance fails: the extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychol Methods*. 2018;23:525-545.
26. Seddig D, Leitgöb H. Approximate measurement invariance and longitudinal confirmatory factor analysis: concept and application with panel data. *Survey Res Methods*. 2018;12:29-41.
27. Muthén B, Asparouhov T. Recent methods for the study of measurement invariance with many groups: alignment and random effects. *Sociol Methods Res*. 2018;47:637-664.
28. Achenbach TM. *The Achenbach System of Empirically Based Assessment (ASEBA): Development, Findings, Theory, and Applications*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families; 2009.
29. Kim MY, Lee J, Chun JH, et al. Self-other congruence in assessing older adults' behavior problems across informant groups. *J Korean Gerontological Society*. 2015;35:395-410.
30. Willemsen G, Vink JM, Abdellaoui A, et al. The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. *Twin Res Human Genet*. 2013;16:271-281.
31. Brigid BD, Achenbach TM, Dumenci L, Newhouse PA. Broad spectrum assessment of psychopathology and adaptive functioning with the Older Adult Behavior Checklist: a validation and diagnostic discrimination study. *Int J Geriatr Psychiatry*. 2010;25:1177-1185.
32. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. *Neurology*. 1994;44:2308-2314.
33. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.
34. Brodaty H, Moore CM. The clock drawing test for dementia of the Alzheimer's type: a comparison of three scoring methods in a memory disorders clinic. *Int J Geriatr Psychiatry*. 1997;12:619-627.
35. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984;141:1356-1364.
36. Morris J. The CDR: current version and scoring rules. *Neurology*. 1993;4:2412-2413.
37. Reisberg B, Ferris SH, de Leon MJ, Crook T. The Global Deterioration Scale (GDS) for assessment of primary degenerative dementia. *Am J Psychiatry*. 1982;139:1136-1139.
38. Reitan RM, Wolfson D. *The Halstead-Reitan Neuropsychological Battery: Theory and Clinical Interpretation*. Tucson: Tucson Neuropsychological Press; 1993.
39. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*. 1969;9:179-186.
40. De Oliveira CR, Paloski LH, Farina M, et al. Older adult self report construct and criterion validity evidence for the Brazilian elderly. *Anales De Psicologia/Annals of Psychology*. 2017;33:277-282.
41. Muthén LK, Muthén BO. *Mplus: User's guide*. Muthén & Muthén: Los Angeles, CA; 2001-2019.
42. Yu CY, Muthén BO. *Evaluation of Model Fit Indices for Latent Variable Models with Categorical and Continuous Outcomes: Technical Report*. Muthén & Muthén: Los Angeles, CA; 2002.
43. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1-55.
44. Marsh HW, Hau KT, Wen Z. In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct Equ Modeling*. 2004;11:320-341.
45. Ivanova MY, Achenbach TM, Rescorla LA, et al. Testing syndromes of psychopathology in parent and youth ratings across societies. *J Clin Child Psychol*. 2019;48:596-609.
46. Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families; 2001.
47. Ivanova MY, Achenbach TM, Rescorla LA, et al. The generalizability of empirically-derived syndromes of collateral-reported elder psychopathology across 11 societies. *Int Psychogeriatr*. Submitted.
48. Ivanova MY, Achenbach TM, Rescorla LA, et al. Syndromes of self-reported psychopathology for ages 18-59 in 29 societies. *J Psychopathol Behav Assess*. 2015;37:171-183.
49. Rescorla LA, Achenbach TM, Ivanova MY, et al. Multicultural comparisons of self- and collateral reports of elder psychopathology. 2019; in preparation.
50. Achenbach TM, Rescorla LA. *Multicultural Supplement to the Manual for the ASEBA Older Adult Forms & Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families; 2019.
51. Swantek SS, Maixner SM, Llorente MD, et al. The road ahead in education: milestones for geriatric psychiatry subspecialty training. *Am J Geriatr Psychiatry*. 2016;24:675-689.
52. Eyre H, Baune B, Lavretsky H. Clinical advances in geriatric psychiatry. *A Focus on Prevention of Mood and Cognitive Disorders Psychiatr Clin North Am*. 2015;38:495-514.

How to cite this article: Ivanova MY, Achenbach TM, Rescorla LA, et al. The generalizability of Older Adult Self-Report (OASR) syndromes of psychopathology across 20 societies. *Int J Geriatr Psychiatry*. 2020;35:525-536. <https://doi.org/10.1002/gps.5268>