

Older adult psychopathology: international comparisons of self-reports, collateral reports, and cross-informant agreement

L.A. Rescorla,¹ M.Y. Ivanova,² T.M. Achenbach,² Vera Almeida,^{3,4} Meltem Anafarta-Sendag,⁵ Ieva Bite,⁶ J. Carlos Caldas,^{3,7} John William Capps,⁸ Yi-Chuen Chen,⁹ Paola Colombo,¹⁰ Margareth da Silva Oliveira,¹¹ Anca Dobrean,¹² Nese Erol,¹³ Alessandra Frigerio,¹⁰ Yasuko Funabiki,¹⁴ Reda Gedutienė,¹⁵ Halldór S. Guðmundsson,¹⁶ Min Quan Heo,¹⁷ Young Ah Kim,¹⁸ Tih-Shih Lee,¹⁷ Manuela Leite,³ Jianghong Liu,¹⁹ Jasminka Markovic,²⁰ Monika Misiec,²¹ Marcus Müller,²² Kyung Ja Oh,²³ Verónica Portillo-Reyes,²⁴ Wolfgang Retz,^{22,25} Sandra B. Sebre,⁶ Shupeng Shi,¹⁹ Sigurveig H. Sigurðardóttir,¹⁶ Roma Šimulionienė,¹⁵ Elvira Sokoli,²⁶ Dragana Milijasevic,²⁷ and Ewa Zasepa²¹

¹Bryn Mawr College, Bryn Mawr, PA, USA

²Department of Psychiatry, Larner College of Medicine, University of Vermont, Burlington, VT, USA

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

⁷CINTESIS – Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

⁸Department of Psychology, The University of Texas at El Paso, El Paso, TX, USA

⁹Department of Psychology, National Chung Cheng University, Chiayi, Taiwan

¹⁰Scientific Institute IRCCS E. Medea, Child Psychopathology Unit, Bosisio Parini, Italy

¹¹Faculty of Psychology, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brazil

¹²Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania

¹³Department of Child and Adolescent Mental Health, Ankara University, School of Medicine, Ankara, Turkey

¹⁴Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

¹⁵Department of Psychology, Klaipėda University, Klaipėda, Lithuania

¹⁶Faculty of Social Work, University of Iceland, Reykjavik, Iceland

¹⁷Neuroscience and Behavioral Disorders Programme, Duke-National University of Singapore Medical School, Singapore, Singapore

¹⁸Huno Inc., Seoul, Republic of Korea

¹⁹Department of Family and Community Health, University of Pennsylvania School of Nursing, Philadelphia, PA, USA

²⁰Private Practice, Novi Sad, Serbia

²¹Department of Psychology, Akademia Pedagogiki Specjalnej, University of Warsaw, Warszawa, Poland

²²Department of Psychiatry and Psychotherapy, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

²³Department of Psychology, Yonsei University, Seoul, Republic of Korea

²⁴Departamento de Psicología, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Mexico

²⁵Institute for Forensic Psychology and Psychiatry, Saarland University, Homburg/Saar, Germany

²⁶Department of Psychology, University of Tirana, Tirana, Albania

²⁷Faculty of Medicine, University of Novi Sad, Institute of Public Health of Vojvodina, Novi Sad, Serbia

ABSTRACT

Objectives: To conduct international comparisons of self-reports, collateral reports, and cross-informant agreement regarding older adult psychopathology.

Participants: We compared self-ratings of problems (e.g. *I cry a lot*) and personal strengths (e.g. *I like to help others*) for 10,686 adults aged 60–102 years from 19 societies and collateral ratings for 7,065 of these adults from 12 societies.

Measurements: Data were obtained via the Older Adult Self-Report (OASR) and the Older Adult Behavior Checklist (OABCL; Achenbach *et al.*, 2004).

Correspondence should be addressed to: Leslie A. Rescorla, Department of Psychology, Bryn Mawr College, 101 N. Merion Ave, Bryn Mawr, PA 19010, USA. Phone: +1 610 526 7318. Email: lrescorl@brynmawr.edu. Received 23 Mar 2020; revision requested 29 May 2020; revised version received 08 Jul 2020; accepted 15 Jul 2020.

Results: Cronbach's alphas were .76 (OASR) and .80 (OABCL) averaged across societies. Across societies, 27 of the 30 problem items with the highest mean ratings and 28 of the 30 items with the lowest mean ratings were the same on the OASR and the OABCL. *Q* correlations between the means of the 0–1–2 ratings for the 113 problem items averaged across all pairs of societies yielded means of .77 (OASR) and .78 (OABCL). For the OASR and OABCL, respectively, analyses of variance (ANOVAs) yielded effect sizes (ESs) for society of 15% and 18% for Total Problems and 42% and 31% for Personal Strengths, respectively. For 5,584 cross-informant dyads in 12 societies, cross-informant correlations averaged across societies were .68 for Total Problems and .58 for Personal Strengths. Mixed-model ANOVAs yielded large effects for society on both Total Problems (ES = 17%) and Personal Strengths (ES = 36%).

Conclusions: The OASR and OABCL are efficient, low-cost, easily administered mental health assessments that can be used internationally to screen for many problems and strengths.

Key words: OASR/OABCL, mental health assessment, multicultural, international

Introduction

As birth rates decline and longevity increases, more countries are experiencing an increase in the proportion of the population aged ≥ 65 years (Kitayama *et al.*, 2020). Karel *et al.* (2012) report that people at the age of 65 years or older will increase from 13% of the U.S. population in 2010 to 16% by 2020. The number of older adults worldwide is predicted to grow annually by >25 million people, such that by 2050, almost 17% of the world's population will be older than 65 (He *et al.*, 2016).

Although older adults contribute to society in many positive ways (e.g. economic and artistic productivity, family participation, community engagement, economic consumption), they also present health care burdens because they often have numerous medical problems as well as cognitive impairment. Karel *et al.* (2012) report that 80% of older adults have at least one chronic condition, with about 60%–65% having two or more conditions. Many older adults also have mental health problems, although mental disorders are less likely to be recognized than physical disorders and/or dementia. Furthermore, older adult immigrants are particularly vulnerable to the mental health stresses that forced migration, immigration, and dislocation entail (Dolberg *et al.*, 2018). Mental health problems in people aged ≥ 65 years may be overlooked because their medical problems and/or dementia may seem more severe and pressing and psychiatric taxonomies such as the American Psychiatric Association's (APA) *Diagnostic and Statistical Manual, Fifth Edition* (DSM-5; APA, 2013) were not developed to focus on people aged ≥ 65 years.

Epidemiological studies of the prevalence of mental disorders in European older adults have yielded widely varying rates. Andreas *et al.* (2017) suggested that their study, using the Composite International Diagnostic Interview (CIDI) 65+ with >3,000 participants in six countries, yielded

higher rates than many previous studies for several disorders (e.g. depression, agoraphobia, and alcohol disorders) because they used a version of the CIDI adapted to meet the specific needs of people aged ≥ 65 years.

Tomlinson *et al.* (2009) note that mental disorders account for about 14% of the global health burden, are linked to many other health problems, and are among the most costly disorders to treat. However, because most mental health research has been done in Western countries, global implementation of evidence-based mental health practices is limited. This is even more true for people aged ≥ 65 years than for young and middle-aged adults. Consequently, mental health needs assessment for older adults should be an important international public health priority.

Although diagnostic interviews may be useful for large multinational epidemiological surveys, their cost may be prohibitive for indigenous investigators conducting epidemiological research in their own societies. Moreover, as Mindt *et al.* (2019) note, most assessment instruments have been developed in rather similar high-income societies and few have been tested with culturally and linguistically diverse older adults. To determine whether an assessment instrument is valid for international applications, it needs to be tested in different societies.

Research that compares findings from the same instrument in different societies exemplifies an *etic* approach, whereas *emic* research examines constructs specific to particular societies (Pike, 1967). Etic epidemiological research compares the prevalence, distribution, and correlates of mental health problems in different populations. Such research can reveal similarities and differences among societies in the prevalence of categorically defined disorders and/or in scores on quantitative scales.

Many theorists have proposed methods for assessing the international equivalence of assessment instruments. For example, Geisinger (1994) argued

that “substantial evidence of the comparability” of a translated/adapted instrument and the original instrument is needed for cross-national comparisons. Poortinga (1989) suggested comparisons of correlations among scales, factor structure, and item difficulty in order to test for invariance of instruments across cultures. Butcher and Han (1996) proposed testing instrument equivalence across different societies by examining whether items perform similarly, using confirmatory factor analysis (CFA), and comparing levels and correlates of scale scores.

The OASR and the OABCL

The Older Adult Self-Report (OASR) and Older Adult Behavior Checklist (OABCL) completed by collaterals (Achenbach *et al.*, 2004) are low-cost, standardized rating forms developed in the U.S. to assess behavioral, emotional, social, and thought problems and personal strengths for adults aged 60 years and above. The OASR and the OABCL also have a few items assessing substance use and relations with friends and spouse/partner, but these data were not available for the current study. The English-language OASR and OABCL are written at a fifth-grade reading level according to readability indices and thus can be read by adults with only a primary education. Each form takes about 15–20 minutes to complete on paper or online and includes 113 problem items and 20 Personal Strengths items. Respondents rate items as 0 = *not true*, 1 = *somewhat or sometimes true*, and 2 = *very true or often true* based on the preceding 2 months. Collaterals such as a spouse, an adult child, a sibling, a friend, or a caregiver/therapist can complete the OABCL to describe an older adult’s functioning.

The sum of 0–1–2 ratings on the 113 problem items provides a broad spectrum Total Problems score for each form. Factor analyses of the OASR and OABCL’s problem items for U.S. samples yielded seven syndromes: *Anxious/Depressed*, *Worries*, *Somatic Complaints*, *Functional Impairment*, *Memory/Cognition Problems*, *Thought Problems*, and *Irritable/Disinhibited*. OASR and OABCL items are also scored on six DSM-oriented scales: *DSM-Depressive Problems*, *DSM-Anxiety Problems*, *DSM-Somatic Problems*, *DSM-Dementia Problems*, *DSM-Psychotic Problems*, and *DSM-Antisocial Personality Problems*. OASR/OABCL problem scores are significantly associated with elders’ cognitive performance, psychopathology, and adaptive functioning on multiple measures, indicating good construct- and criterion-related validity of the syndromes (Achenbach *et al.*, 2004; Brigidi *et al.*, 2010). The 20-item *Personal Strengths* scale comprises items such as *I make good use of my time*.

Indigenous mental health researchers requested permission from the OASR and OABCL’s authors

(Achenbach *et al.*, 2004) to develop non-English versions of the two instruments to use in their own societies. Many of these collaborators had already contributed data for previous international comparisons of problems in adults aged 18 to 59 years (Rescorla *et al.*, 2016a,b). After a translation and independent back-translation process, the indigenous researchers collected OASR self-ratings in 20 societies and OABCL collateral ratings in 12 societies.

Ivanova *et al.* (2020a) performed CFAs of within-society self-ratings of OASR problem items by 12,826 of 60- to 102-year-olds in 20 societies to test the fit of OASR problem item ratings to the 7-syndrome model derived from factor analyses of U.S. data. *Configural* invariance tests cross-society similarities of item arrangements into factors, *metric* invariance tests the degree to which items reflect the same underlying syndrome constructs across societies, and *scalar* invariance tests the invariance of item intercepts/thresholds across societies (Ivanova *et al.*, 2010). Items loaded strongly on their respective factors, with a median item loading of .63 across 20 societies; 98.7% of the loadings were statistically significant. Additionally, 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, supporting the generalizability of OASR–OABCL syndromes across societies. In a parallel study, Ivanova *et al.* (2020b under review) used within-society CFA procedures to test the 7-syndrome model in OABCLs rated by 7,283 collaterals of 60- to 102-year-olds from 12 societies. The items loaded strongly on their respective factors, with a median item loading of .69 across the 12 societies. By syndrome, the overall median item loadings ranged from .47 for *Worries* to .77 for *Functional Impairment*, indicating that the syndrome structure was generalizable across the tested societies.

Ivanova *et al.*’s CFA studies are important steps in testing similarities and differences across many societies in performance of the U.S.-derived OASR and OABCL. However, demonstrating international robustness of the OASR and the OABCL requires more than just testing similarities and differences in factor structure across societies. Because the OASR and OABCL are broad instruments that tap many different aspects of functioning, they afford testing of similarities and differences across various societies not only in the empirically derived syndromes but also the DSM-oriented scales, the Total Problems scale, the Critical Items scale, and the Personal Strengths scale. We also used different statistical approaches than CFAs, namely *Q* correlations

for mean item ratings, as well as Pearson correlations and analyses of variance (ANOVAs) to test scale scores.

Goals of our study

To address issues not addressed by Ivanova *et al.*'s CFA studies, we analyzed OASR data from 19 of the 20 societies analyzed by Ivanova *et al.* (2020a) and OABCL data from the same 12 societies analyzed by Ivanova *et al.* (2020b). We excluded Ivanova's 20th society because the older adults were all parents of twins participating in the Netherlands Twin Register or else twins themselves and hence did not comprise a representative population sample. Our overarching purpose was to identify both similarities and differences across societies in reports of older adults' behavioral, emotional, social, and thought problems and personal strengths by use of correlation analyses and ANOVAs

Based on results from our previous international comparisons of reports of problems for adults aged 18–59 years using many of the same kinds of analyses (Rescorla *et al.*, 2016a,b), we hypothesized (a) that the pattern of Cronbach's alphas of OASR/OABCL scales would be similar to those reported for the U.S.; (b) that there would be strong consistency across societies with respect to ratings of specific problems based on means of the 0–1–2 ratings, as tested by *Q* correlations; (c) that societal effects on OASR/OABCL scores would be significant but relatively modest for most problem scales and large for Personal Strengths; (d) that age group and gender effects would be small but consistent across societies; and (e) that OASR/OABCL cross-informant agreement would be moderate as measured by correlations and ANOVAs.

Method

Samples

We used data from 19 societies for the OASR and from 12 of those societies for the OABCL (see Supplement Table 1). The data for each society were collected by indigenous investigators after obtaining informed consent as required by their home research institutions. Rigorous random sampling methods were used in some societies but convenience sampling was used in other societies. Following Nunnally and Bernstein (1994), we required a minimum *N* of 300 per society. Our cross-informant sample included 5,584 dyads from 12 societies for whom we obtained matched OASR/OABCL data.

Our OASR sample comprised 10,686 elders (45% male), with a mean age of 71.5 (*SD* = 8.3) years. Our OABCL sample comprised 7,065 elders

(39.7% male), with a mean age of 72.7 (*SD* = 8.7) years. We subdivided OASR and OABCL samples roughly in half by age group, with 59% and 53% of participants ≤ 72 years and 41% and 47% of participants >72 years.

Instrument

The OASR/OABCL's 113 problem items tap a wide variety of emotional, behavioral, social, and thought problems. The 20 Personal Strengths items (e.g. *I make good use of opportunities*) are rated on the same 0–1–2 scale as the problem items. For OASR and OABCL U.S. data, respectively, Achenbach *et al.* (2004) reported alphas of .96 and .97 for the broad spectrum Total Problems scale, .69 to .92 and .66 to .92 for the syndromes, .63 to .88 and .75 to .89 for the *DSM*-oriented scales, and .83 and .90 for the Personal Strengths scale. One-week test–retest correlations (*r*s) were .95 for Total Problems on both forms, .74 to .94 and .92 to .95 for the OASR/OABCL syndromes, .78 to .93 and .90 to .95 for the OASR/OABCL *DSM*-oriented scales, and .91 and .95 for the OASR/OABCL Personal Strengths scale. OASR/OABCL scales significantly discriminated between demographically similar clinically referred and non-referred samples of adults, accounting for 20% and 29% of the variance in Total Problems and 11% and 30% of the variance in Personal Strengths, respectively.

Data analysis

We had very few missing values in our data because, following standard practice with the OASR and OABCL, the few cases with >8 problem items left blank were excluded as invalid. OASR/OABCL problem scale scores were positively skewed in all our samples, but general linear models are very robust with respect to deviations from normality, especially with our very stringent $p < .001$ criterion for significance and large samples having similar skew (Kirk, 1995). Accordingly, we used untransformed raw scores for all analyses.

We first computed Cronbach's alphas to test for internal consistencies of the OASR/OABCL problem scales and Personal Strengths scales in each society. Next, we calculated mean item ratings for each of the 113 problem items on the OASR/OABCL for each society. These 19 rank orderings of endorsement for the OASR and 12 rank orderings for the OABCL (1 per society) served as our measure of relative problem frequency/severity in each society for each form. For each form, we averaged these mean item ratings to obtain the omnicultural mean item rating for each form. We also computed *Q* correlations between the mean 0–1–2 problem item ratings from each society and the mean item

Table 1. Omnicultural means (*SDs*) of alpha coefficient for the OASR (19 societies, *N* = 10,686) and the OABCL (12 societies, *N* = 7,065)

SCALE (NUMBER OF ITEMS)	OASR <i>M</i> (<i>SD</i>), RANGE	OABCL <i>M</i> (<i>SD</i>), RANGE
<i>Broad-band scale</i>		
Total problems (114)	.95 (.01), .93–.98	.96 (.01), .93–.98
<i>Syndromes</i>		
Anxious/depressed (20)	.87 (.02), .83–.93	.87 (.04), .79–.93
Worries (8)	.62 (.07), .48–.76	.61 (.10), .45–.75
Somatic complaints (14)	.80 (.04), .74–.86	.82 (.05), .71–.87
Functional impairment (11)	.77 (.07), .59–.88	.84 (.03), .78–.88
Memory/cognition problems (9)	.78 (.02), .75–.82	.82 (.02), .78–.84
Thought problems (15)	.73 (.06), .62–.89	.78 (.05), .70–.89
Irritable/disinhibited (20)	.82 (.04), .75–.90	.88 (.04), .77–.91
<i>DSM-oriented scales</i>		
Depressive problems (18)	.82 (.03), .78–.90	.84 (.04), .73–.89
Anxiety problems (9)	.75 (.03), .69–.85	.75 (.05), .63–.86
Somatic problems (8)	.72 (.06), .60–.80	.74 (.08), .59–.82
Dementia problems (7)	.73 (.03), .68–.80	.79 (.02), .76–.82
Psychotic problems (6)	.55 (.11), .36–.84	.62 (.11), .44–.85
Antisocial personality problems (11)	.63 (.09), .47–.81	.72 (.08), .57–.83
Critical items (31)	.85 (.03), .81–.93	.87 (.03), .80–.93
Personal strengths (20)	.81 (.04), .73–.88	.85 (.03), .80–.91

OASR = Older Adult Self-Report; OABCL = Older Adult Behavior Checklist.

ratings from each other society. A *Q* correlation is a Pearson *r* calculated between groups across item scores rather than between item scores across participants (Stephenson, 1936). *Qs* are more appropriate for multicultural comparisons than *rs* because they reflect agreement at the group rather than individual level of analysis. Next, we tested each OASR/OABCL scale score with society, gender, and age group as factors in ANOVAs. Finally, for our cross-informant sample of 5,584 dyads from 12 societies, we assessed inter-rater reliability by calculating Pearson *rs* between scale scores for participants based on self-report and collateral ratings. We also used mixed-models ANOVAs for Total Problems and Personal Strengths, with informant (self vs. collateral), society, gender, and age group as predictors. Because our very large sample size made even very small effects significant at $p < .05$, we used $p \leq .001$ for all analyses to avoid Type 1 errors. We report effect sizes (ESs) measured by η^2 characterized using Cohen's (1988) criteria (*small* = .01 to .059, *medium* = .06 to .139, *large* $\geq .14$), rather than *F* and *p* values.

Results

Internal consistency of OASR/OABCL scales in different societies

Alphas for Total Problems were $\geq .93$ for each of the 19 OASR societies and each of the 12 OABCL

societies. The omnicultural mean Total Problems alpha (average of the societal alphas) was .95 for the OASR and .96 for the OABCL (see Table 1). Averaged across the 16 OASR/OABCL scales, the omnicultural mean alphas were .76 and .80, respectively, indicating good internal consistency.

Mean problem item ratings

The mean ratings for each of the 113 OASR/OABCL problem items were averaged to yield an omnicultural mean item rating. As shown in Table 2, 27 of the 30 problem items with the highest omnicultural mean ratings were the same on the OASR and the OABCL. The most commonly endorsed items included problems from the Anxious/Depressed, Worries Memory/Cognition, Irritable/Disinhibited, Somatic Complaints, and Thought Problems syndromes. Furthermore, 28 of the 30 problem items with the lowest omnicultural mean ratings on the OASR and the OABCL (Table 3) were the same on the 2 forms. These rarely endorsed problems include stealing, breaking the law, self-injury, seeing or hearing things that aren't there, threatening others, strange behavior or ideas, drinking too much alcohol, repeating acts over and over, and attacking people. Given that there was very strong concordance between self-reports and collateral reports on 60 of the 113 OASR/OABCL items, it can be inferred that there was also strong concordance between self-reports and collateral reports for the remaining 53 items. Overall, the rankings of omnicultural mean

Table 2. OASR and OABCL problem items with highest 30 omnicultural means

	OASR ITEM	OASR OC MEAN	OABCL OC MEAN
1 ^a	35. I like to have things my own way	1.27	1.16
2	72. I worry about my family	1.18	1.24
3	89. I am too concerned about being neat or clean	0.87	0.79
4	101. I wake up too early	0.82	0.73
5	100. I worry a lot	0.74	0.73
6	114. If I don't write things down, I forget them	0.73	0.57
7	102. I worry too much about my health	0.72	0.70
8	20. I forget people's names	0.70	0.48
9	60. I am secretive or keep things to myself	0.65	0.55
10	92. I don't have much energy	0.64	0.58
11	91. I think about the past too much	0.64	0.59
12	90. I have trouble sleeping	0.63	0.56
13	117. I get too tired from doing my daily tasks	0.62	0.59
14	5. I use too much medication	0.61	0.63
15	21. I worry about my future	0.6	0.62
16	7. I have trouble concentrating or paying attention	0.57	0.50
17	110. I have trouble remembering things I am told	0.56	0.45
18	122. I worry too much about my memory	0.55	0.46
19	38. I would rather be alone than with others	0.53	
20	51. I worry about my appearance	0.52	0.59
21	48. I feel tired without good reason	0.52	0.45
22	3. I have difficulty getting things done	0.51	0.45
23	2. I argue a lot	0.49	0.53
24	8. I can't get my mind off certain thoughts (describe)	0.48	
25	76. I am stubborn, sullen, or irritable	0.48	0.55
26	49a. I have aches or pains not due to physical cause, medication	0.48	0.48
27	40. I am nervous or tense	0.48	0.46
28	11. I feel lonely	0.47	0.47
29	105. I don't like to use the telephone	0.47	
30	53. There is very little that I enjoy	0.46	0.43

OASR = Older Adult Self-Report; OABCL = Older Adult Behavior Checklist; OC = omnicultural mean, obtained by averaging the mean item ratings from each society.

^aItems are listed in descending order of mean scores for OASR descending order.

item ratings indicate that collateral and self-ratings agreed very well.

Mean problem item rating *Q* correlations

Q correlations computed between the means of the 0–1–2 ratings for each of the 113 problem items in each pair of societies yielded bi-society *Q*s from .66 (Korea) to .82 (Serbia) for the OASR and from .71 (Japan) to .82 (Poland) for the OABCL. When the mean bi-society *Q*s were averaged for each instrument, the omnicultural mean bi-society *Q* was .77 for the OASR and .78 for the OABCL. These large effects (Cohen, 1988) indicate that older adults in very different societies agreed strongly with respect to which problem items tended to receive low, medium, or high ratings, whether based on self-reports or collateral reports.

Problem scale scores

For each OASR/OABCL problem scale, Table 4 displays the smallest and the largest societal mean, the omnicultural mean (and its *SD*), and the omnicultural mean of the societal *SD*s for each scale. Ten societies had mean Total Problems scores within 1 *SD* of the omnicultural mean. There was more within-societal variation in OASR/OABCL scores than between-society variation, as indicated by comparing the *SD* of the omnicultural mean vs. the omnicultural mean of the within-society *SD*s (i.e. 10.9 vs. 24.7 for OASR Total Problems).

As shown in Table 5, Total Problems had societal ESs of 15% and 18%, respectively, on the OASR and the OABCL. For the syndrome scales, Worries had the largest ESs on both forms (30% and 32%) and Functional Impairment (6% and 5%) and

Table 3. OASR and OABCL problem items with lowest 30 omnicultural means

	OASR ITEM	OASR OC MEAN	OABCL OC MEAN
1 ^a	73. I steal things	0.03	0.03
2	82. I do things that may cause trouble with the law	0.03	0.03
3	50. I physically attack people	0.03	0.05
4	17. I deliberately try to hurt or kill myself	0.03	0.04
5	81. I talk (think) about killing myself	0.04	0.04
6	87. I threaten to hurt people	0.04	0.05
7	19. I damage or destroy things	0.06	0.06
8	36. I hear sounds or voices that others think are not there	0.07	0.06
9	49g. I vomit or throw up, not due to physical cause/medication	0.07	0.09
10	61. I see things other people think are not there	0.08	0.05
11	74. I do things that other people think are strange	0.09	0.06
12	67. My behavior is irresponsible	0.10	0.13
13	31. I feel that others are out to get me	0.10	0.13
14	75. I have thoughts that other people (would) think are strange	0.11	0.06
15	80. I drink too much alcohol or get drunk	0.11	0.11
16	63. I am being punished for what I have done	0.13	0.13
17	104. I have trouble dressing myself	0.14	0.18
18	111. I have soiling accidents	0.14	0.16
19	24. I am jealous of others	0.16	
20	106. I have trouble bathing or grooming	0.16	0.21
21	25. I get along badly with my family	0.16	0.20
22	15. I am mean to others	0.16	0.17
23	28. I am afraid I might think or do something bad	0.16	0.13
24	41. Parts of my body twitch or make nervous movements	0.16	0.41
25	86. I think about sex too much	0.18	0.12
26	49c. I feel nauseous/sick, not due to physical cause/medication	0.18	0.18
27	70. I have trouble talking	0.19	0.15
28	57. I repeat certain acts over and over	0.19	0.12
29	32. I feel worthless or inferior	0.20	
30	23. I feel too guilty	0.21	0.19

OASR = Older Adult Self-Report; OABCL = Older Adult Behavior Checklist; OC = omnicultural mean, obtained by averaging the mean item ratings from each society.

^aItems are listed in ascending order of OASR mean scores.

Memory/Cognition (5% and 5%) had the smallest OASR/OABCL ESs. The 19 societies varied less on the scales tapping aspects of dementia than on the scales tapping emotional problems.

OASR/OABCL ESs for gender were either not significant or <2%. Women tended to have slightly higher scores than men except on Irritable/Disinhibited and DSM-Antisocial Personality Problems. ESs for age group were either not significant or <2%, except for Functional Impairment, where they were 3% and 5% for the OASR and OABCL, respectively. Scores were higher in the older than the younger age group for most scales on both forms, except for Personal Strengths, where younger adults had higher scores on both forms.

To test reports for differences in scores derived from ratings by different kinds of collaterals (e.g. spouse/partner, child, and "other"), we used within-society ANOVAs on OABCL Total Problems scores

for the 11 societies reporting the collateral's relationship to the OASR target (missing for Japan). Differences were significant for only 2 of the 11 societies at $p < .001$ (U.S. and Korea, both with scores from children of the OASR target yielding the highest mean score).

The OASR/OABCL societal ESs of 42% and 31% for Personal Strengths were much larger than the ESs found for the problem scales. On both the OASR and the OABCL, Korea (12.1, 14.0) and Japan (16.4, 16.6) had much lower mean Personal Strengths scores than the other societies, with all the other societies all scoring within 1 *SD* of the omnicultural mean.

OASR/OABCL cross-informant agreement

To assess OASR/OABCL cross-informant agreement, we first calculated Pearson *r*s between scale scores for the 5,584 dyads in each of the 12 societies

Table 4. OASR/OABCL descriptive statistics

SCALE	MINIMUM MEAN	MAXIMUM MEAN	OMNICULTURAL MEAN (SD)	OMNICULTURAL SD OASR/OABCL
Total problems	26.5/24.3	59.5/59.6	39.8 (10.9)/37.5 (11.9)	24.7/25.9
<i>Syndromes</i>				
Anxious/depressed	4.5/3.4	11.8/11.6	7.8 (2.4)/7.2 (2.6)	6.4/6.2
Worries	3.3/3.1	9.0/8.9	5.6 (1.9)/5.5 (2.0)	2.8/2.8
Somatic complaints	2.1/2.3	8.6/8.2	4.6 (1.8)/4.6 (1.8)	4.2/4.3
Functional impairment	1.8/2.3	4.8/5.2	3.5(1.0)/3.6 (0.9)	3.5/4.0
Memory/cognition	3.0/2.8	5.6/5.0	4.4 (0.7)/3.6 (0.7)	3.3/3.4
Thought problems	2.3/2.1	5.2/5.2	3.6 (1.0)/3.2 (1.1)	3.3/3.5
Irritable/disinhibited	4.0/3.7	9.9/10.6	6.5 (1.9)/6.5 (2.4)	4.9/5.8
Other problems	2.3/2.1	5.8/5.6	3.7 (1.1)/3.3 (1.1)	3.2/3.2
<i>DSM-oriented scales</i>				
Depressive problems	3.9/3.4	9.7/9.3	6.6 (1.8)/6.1 (1.9)	5.1/5.3
Anxiety problems	1.8/1.4	6.2/6.0	3.9 (1.4)/3.5 (1.4)	3.2/3.0
Somatic problems	1.4/1.6	5.6/5.5	2.9 (1.1)/2.9 (1.2)	2.6/2.7
Dementia problems	2.6/2.3	4.6/3.9	3.7 (0.5)/3.0 (0.5)	2.7/2.9
Psychotic problems	0.2/0.3	1.4/1.3	.8 (0.3)/0.7 (0.3)	1.3/1.3
Antisocial Problems	1.7/1.6	4.1/4.4	2.9 (.07)/2.8 (0.9)	2.2/2.6
Critical items	5.7/5.2	12.8/12.9	9.0 (2.3)/8.2 (2.4)	6.7/7.0
Personal strengths	12.2/14.0	30.7/28.8	26.3 (4.9)/24.1 (4.8)	6.5/7.4

Omnicultural mean = average of the scale means across all societies for each form; Omnicultural SD = mean of the OASR and OABCL SDs for each scale across all societies. OASR = Older Adult Self-Report; OABCL = Older Adult Behavior Checklist.

in the cross-informant sample. The omnicultural mean cross-informant r for Total Problems was (.68, $SD = 12.2$, range = .54–.89), large by Cohen's (1988) benchmarks. The r was also large for Personal Strengths (.58, $SD = 15.9$, range = .32–.88). For the syndromes and DSM scales, omnicultural mean cross-informant r s ranged from .46 (DSM-Psychotic Problems) to .69 (Functional Impairment).

Mixed-model ANOVAs testing the effects of society, informant (within-subjects), gender, and age group for OASR/OABCL Total Problems scores yielded significant but very small ESs for informant, gender, and age group ($ES \leq 1\%$), but a society ES of 17%, indicating that Total Problems scores varied considerably across our 12 societies having cross-informant data. The results for Personal Strengths were similar, namely very small ($ES < 1\%$) effects for informant, gender ($ES < 1\%$), and age group ($ES = 1\%$), and an ES of 36% for society.

Discussion

Our international comparisons of 19 societies for self-reports and 12 societies for collateral reports are quite unique. To our knowledge, no previous research has made international comparisons of problems and strengths in large population samples of older adults from multiple continents based on

self-reports and collateral reports using parallel instruments. Furthermore, we tested cross-informant agreement using both correlations and ANOVAs for 5,485 dyads, which is rare in mental health research on older adults. By combining data for all the societies into a single data set, we could directly test similarities and differences between societies.

Our study is similar in many ways to our international comparisons of self-report and collateral reports for adults aged 16 to 59 years, which utilized the ASR and the ABCL (Rescorla *et al.*, 2016a,b). Because many items had counterparts on the ASR and the OASR (and on the ABCL and the OABCL), and because we conducted many of the same analyses as conducted for adults aged 18 to 59 years, viewing our current OASR/OABCL findings in relation to previous findings, ASR/ABCL findings is particularly informative.

Our international comparisons of data for people aged ≥ 65 years in 19 societies using the OASR and in 12 societies using the OABCL revealed many similarities across societies that differed widely in economic, political, religious, geographic, linguistic, and ethnic characteristics. For example, although Cronbach's alphas varied across scales, the omnicultural mean alphas for the 16 OASR/OABCL scales indicated good overall internal consistency, comparable with those reported by Rescorla *et al.* (2016a) for 10,197 adults aged 18 to 59 years in 17

Table 5. Significant effect sizes (η^2) for society \times gender \times age group ANOVAs on OASR/OABCL scales

SCALE	SOCIETY	GENDER	AGE	S \times G	S \times A	G \times A	S \times G \times A
Total problems	15/18	<1 \ll 1	<1/1	<1/NS	NS/<1	NS/NS	NS/NS
<i>Syndromes</i>							
Anxious/depressed	13/16	2/<1	<1/<1	<1/<1	NS/<1	NS/NS	NS/NS
Worries	30/32	2/2	<1/NS	<1/NS	NS/NS	NS/NS	NS/NS
Somatic complaints	14/15	1/1	1/2	<1/<1	<1/<1	NS/NS	NS/NS
Functional impairment	6/5	NS/NS	3/5	<1/<1	<1/1	NS/NS	<1/NS
Memory/cognition	5/5	<1/<1	<1/2	NS/<1	NS/<1	NS/NS	NS/NS
Thought problems	8/8	NS/NS	NS/<1	NS/NS	NS/<1	NS/NS	NS/NS
Irritable/disinhibited	12/14	<1/<1	NS/NS	<1/<1	NS/1	NS/NS	NS/NS
<i>DSM-Oriented Scales</i>							
Depressive problems	12/14	<1/<1	<1/1	<1/<1	<1/<1	NS/NS	NS/NS
Anxiety problems	16/20	2/<1	<1/<1	1/<1	NS/<1	NS/NS	NS/NS
Somatic problems	15/16	1/1	<1/1	<1/<1	<1/<1	NS/NS	NS/NS
Dementia problems	5/3	<1/NS	1/2	NS/<1	<1/<1	NS/NS	NS/NS
Psychotic problems	5/6	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS	NS/NS
Antisocial problems	9/10	<1/<1	NS/NS	<1/<1	NS/1	NS/NS	NS/NS
Critical items	10/11	<1/<1	<1/1	<1/NS	NS/<1	NS/NS	NS/NS
Personal strengths	42/31	<1/<1	<1/2	NS/NS	<1/<1	<1/NS	NS/NS

OASR = Older Adult Self-Report; OABCL = Older Adult Behavior Checklist.

Cell entries reflect percent of OASR/OABCL variance accounted for by scale; OASR: 19 societies ($N = 10,686$); OABCL: 12 societies, ($N = 7,065$); NS = not significant at $p < .001$.

societies, where alphas were $>.70$ for 9 scales and $>.60$ for 8 other scales.

We found that the majority of most and least commonly endorsed problem items were the same on the OASR and OABCL across very different societies. Furthermore, our large omnicultural mean bi-society Q correlations of $.77$ and $.78$ for the OASR and OABCL problem items indicated strong agreement across societies with respect to which items tended to receive low, medium, and high ratings. These findings suggest that the older adults completing the OASR and their collateral informants interpreted the OASR/OABCL items very similarly across 19 societies, despite the fact that the societies differed widely in race/ethnicity, religion, economic/political characteristics, geographical region, and language. These findings for people aged ≥ 65 years are very consistent with findings by Rescorla *et al.* (2016a,b) of an omnicultural mean Q correlation of $.77$ for mean item self-ratings by adults aged 18 to 59 years across 17 societies and of $.76$ for collateral ratings in 14 societies.

The large ESs for society yielded by ANOVAs indicated significant societal variation in OASR and OABCL problem scores. Interestingly, societies varied less on scales tapping dementia (Functional Impairment, Memory/Cognition, and DSM-Dementia Problems) than on scales tapping emotional problems, suggesting that the prevalence

and/or perceptions of emotional problems are influenced more by culture than the prevalence and/or perceptions of cognitive problems. Despite societal differences in Total Problems scores, a diverse set of 10 societies (China, Korea, U.S., Italy, Serbia, Brazil, Latvia, Romania, Portugal, and Poland) scored within 1 SD of the OASR omnicultural mean, whereas 8 did so on the OABCL (Korea, Iceland, U.S., Serbia, Brazil, Poland, Portugal, and Turkey). For both the OASR and the OABCL, the societies scoring within 1 SD of the omnicultural mean differ in many ways (e.g. race/ethnicity, religion, economic/political characteristics, geographical region, language). For both forms, there was more within-societal variation in OASR and OABCL scores than between-society variation. These findings are very consistent with findings by Rescorla *et al.* (2016a) for the ASR, where nine quite different societies had mean Total Problems scores within 1 SD (6.1) of the omnicultural mean of 42.7.

The five societies with Total Problems scores >1 SD below the omnicultural mean on the OASR included three Asian societies (Japan, Taiwan, and Singapore, but not China or Korea) and two European countries (Germany and Iceland, but not Poland, Italy, Latvia, Serbia, Romania, or Portugal). Japan and Taiwan also had Total Problems scores >1 SD below the omnicultural mean on the ABCL. There is no obvious commonality among these

low-scoring societies that is also not found in the larger number of societies scoring within 1 *SD* of the omnicultural mean beyond the fact that they tend to report somewhat fewer problems than the other societies. The same can be said for the four societies with OASR scores and the two societies with OABCL scores >1 *SD* above the omnicultural mean (OASR: Mexico, Turkey, Lithuania, and Albania; OABCL: Turkey, Albania).

An important finding is that the very large societal ESs for Personal Strengths were much larger than those found for the problem scales. On both forms, Korea and Japan had the lowest mean Personal Strengths scores, but other Confucian societies such as China, Singapore, and Taiwan did not have exceptionally low scores. Interestingly, Rescorla *et al.* (2016a,b) found that Japan and Korea also had low mean Personal Strengths scores on both the ASR and the ABCL. These findings suggest considerable cultural consistency in both self-report and collateral report endorsement of personal strengths across the ages 18–59 and 60–90+ years in Japan and Korea. It appears that not only do adults in both societies tend to have a strong tendency to avoid endorsing favorable statements about themselves but people who know them well also tend to avoid endorsing favorable statements about them more than do collaterals in other societies, including neighboring Confucian societies. An emic investigation (Pike, 1967) involving interviews and focus groups might help to elucidate cultural values in Japan and Korea (e.g. modesty, collectivism, conformity) that could potentially explain these findings.

Our large omnicultural mean cross-informant *r* for Total Problems (.68) indicated that collateral reports and self-reports were quite concordant for 5,584 dyads in 12 societies. This was larger than Rescorla *et al.*'s (2016b) ASR/ABCL cross-informant *r* of .50 for Total Problems, indicating better agreement in dyads where the target was of age ≥ 65 years than ages 18 to 59 years. Furthermore, our mixed-model ANOVAs indicated that OASR and OABCL mean Total Problems scores were very similar (informant ES < 1%), whereas 18- to 59-year-old's ASR mean Total Problems were considerably higher than ABCL mean scores (ES = 7%). It is possible that because collaterals in a dyad with an older adult might have known each other for more years than dyads with an adult aged 18–59 years, collateral and self-reports might be more similar in the former than the latter. However, our study did not provide data relevant to this speculation. Another possible reason for greater concordance in older than younger adults is that the brain circuitry underlying the assessed problems may become more entrenched with age.

Clinical applications

Our findings have implications for practical assessment in specific societies. That ratings by collaterals were highly consistent with self-ratings for most OASR/OABCL items suggests that large discrepancies in a particular dyad might have clinical relevance. That some societies manifested a low-scoring tendency, others a medium-scoring tendency, and others a high-scoring tendency for problems is also important. Multicultural norms are currently available to take account of the tendencies of some societies to have relatively low, medium, or high scores for ages 18–59 years (Achenbach and Rescorla, 2015), but not yet for ages 50–90+ years. Clinicians in “high-scoring” societies, such as Albania and Turkey, may therefore choose to be more conservative in interpreting high OASR/OABCL problem scores than clinicians in “low-scoring” societies such as Japan and Taiwan. Additionally, our findings that adults of the ages 18–59 and 60–90+ years in both Japan and Korea obtained relatively low Personal Strengths scores may help clinicians in those societies to evaluate whether a particular client has especially low perceptions of his or her personal strengths or whether the scores reflect a more general low-scoring tendency in that client's society regarding personal strengths.

Limitations

A limitation of our study is that some of our samples were of unknown representativeness because they were obtained using convenience rather than random sampling methods. Because ours was an etic study, the same instrument (in translation) was used in every society. Inclusion of society-specific items might yield different results. It should also be noted that our obtained societal differences in scores may involve societal differences in response styles.

Conclusions

Older adults comprise an increasingly large percentage of the world's population (He *et al.*, 2016), and most older adults have at least one chronic health condition (Karel *et al.*, 2012). Many older adults also have mental health problems, although mental disorders are often lumped together with dementia in epidemiological estimates. Moreover, current taxonomies such as *DSM-5* (American Psychiatric Association, 2013) were not developed to focus on psychiatric problems for people aged ≥ 65 years, beyond various forms of dementia. Assessing mental health problems in older adults is therefore going to be an increasingly important public health need in the coming years.

Our study demonstrates that international epidemiological data on mental health problems in people aged ≥ 65 years can be obtained for large samples at relatively low cost when indigenous investigators use the same standardized assessment instruments that do not require professional time for administration or scoring. Data were easily obtained in diverse societies (e.g. Albania, Brazil, Iceland, Kenya, Latvia, and Serbia), including some with turbulent recent histories. Because the same assessment instruments were used in all societies (following a rigorous translation and back-translation process) and the data comprised quantitative ratings, they could be easily merged to enable international comparisons. OASR and OABCL items appeared to be interpreted quite similarly across the societies we studied, as reflected in our large omnicultural mean Q_s for mean item ratings. The list of items with the highest and lowest mean ratings should be useful for researchers who survey mental health problems in people aged ≥ 65 years. In sum, because the OASR and the OABCL assess problems and personal strengths (as well as adaptive functioning and substance use, not analyzed here), they are efficient, low-cost, easily administered mental health instruments that can be used internationally to screen for a wide range of problems and strengths.

Conflict of interest

The first three authors receive compensation from the University of Vermont non-profit 501(3)© Research Center for Children, Youth, and Families, which publishes the instruments used in this research

Description of authors' roles

Leslie A. Rescorla formulated the research questions, designed the study, analyzed the data, and wrote the article. Masha Y. Ivanova and Thomas M. Achenbach provided feedback on drafts of the manuscript. The international collaborators provided the data for the study and reviewed the final manuscript prior to submission.

Supplementary material

To view supplementary material for this article, please visit <https://doi.org/10.1017/S1041610220001532>

References

- Achenbach, T. M., Newhouse, P. A. and Rescorla, L. A.** (2004). *Manual for the ASEBA Older Adult Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- Achenbach, T. M. and Rescorla, L. A.** (2015). *Multicultural Supplement to the Manual for the ASEBA Adult Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- American Psychiatric Association.** (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.
- Andreas, S. et al.** (2017). Prevalence of mental disorders in elderly people: the European MentDis_ICF65+ study. *British Journal of Psychiatry*, 210(2), 125–131.
- Brigidi, B. D., Achenbach, T. M., Dumenci, L. and Newhouse, P. A.** (2010). Broad spectrum assessment of psychopathology and adaptive functioning with the Older Adult Behavior Checklist: a validation and diagnostic discrimination study. *International Journal of Geriatric Psychiatry*, 25(11), 1177–1185.
- Butcher, J. N. and Han, K.** (1996). Methods of establishing cross-cultural equivalence. In: J. N. Butcher (Ed.), *International Adaptations of the MMPI-2: Research and Clinical Applications* (pp. 44–63). Minneapolis, MN: University of Minnesota Press.
- Cohen, J.** (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Academic Press.
- Dolberg, P., Sigurðardóttir, S. H. and Trummer, U.** (2018). Ageism and older immigrants. In: L. Ayalon and C. Tesch-Römer (Eds.), *Contemporary Perspectives on Ageism* (pp 177–191). Cham: Springer.
- Geisinger, K. F.** (1994). Cross-cultural normative assessment: translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6, 304.
- He, W., Goodkind, D. and Kowal, P.** (2016). *An Aging World: 2015: International Population Reports* (P95/16-1). Suitland, MD: US Census Bureau.
- Ivanova, M. Y. et al.** (2010). Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(12), 1215–1224.
- Ivanova, M. Y. et al.** (2020a). The generalizability of Older Adult Self-Report (OASR) syndromes of psychopathology across 20 societies. *International Journal of Geriatric Psychiatry*, 35(5), 525–536. doi: [10.1002/gps.5268](https://doi.org/10.1002/gps.5268).
- Ivanova, M. Y., et al.** (2020b). The generalizability of empirically-derived syndromes of collateral-reported elder psychopathology across 11 societies (under review).
- Karel, M. J., Gatz, M. and Smyer, M. A.** (2012). Aging and mental health in the decade ahead: what psychologists need to know. *American Psychologist*, 67(3), 184.
- Kirk, R. R.** (1995) *Experimental Design Procedures for the Behavioral Sciences*. Pacific Grove, CA: Brookside, Cole.
- Kitayama, S., Berg, M. K. and Chopik, W. J.** (2020) Culture and well-being in later adulthood: theory and evidence. *American Psychologist*, 75, 576–576.

- Mindt, M. R. *et al.*** (2019). Handbook on the neuropsychology of aging and dementia. In: L. Ravdin and H. Katzen (Eds.), *Clinical Handbooks in Neuropsychology* (pp. 25–48). Cham: Springer.
- Nunnally, J. C. and Bernstein, H.** (1994). *Psychometric Theory* (3rd ed.). New York: McGraw Hill.
- Pike, K. L.** (Ed.) (1967). *Language in Relation to a Unified Theory of the Structure of Human Behavior* (2nd ed.). The Hague: Mouton De Gruyter.
- Poortinga, Y. H.** (1989). Equivalence of cross-cultural data: an overview of basic issues. *International Journal of Psychology*, 24, 737–756.
- Rescorla, L. A. *et al.*** (2016a). Problems and adaptive functioning reported by adults in 17 societies. *International Perspectives in Psychology: Research, Practice, Consultation*, 5(2), 91.
- Rescorla, L. A. *et al.*** (2016b). Collateral reports and cross-informant agreement about adult psychopathology in 14 societies. *Journal of Psychopathology and Behavioral Assessment*, 38(3), 381–397.
- Stephenson, W.** (1936). The inverted factor technique. *British Journal of Psychology*, 26(4), 344.
- Tomlinson, M. *et al.*** (2009). Setting priorities for global mental health research. *Bulletin of World Health Organization*, 87, 438–446. doi: [10.2471/BLT.08.054353](https://doi.org/10.2471/BLT.08.054353).