A cross-national analysis of the psychometric properties of the Geriatric Anxiety Inventory (GAI)

Helge Molde, PhD,[1] Inger-Hilde Nordhus, PhD,[1,2] Torbjørn Torsheim, PhD,[3] Knut Engedal, PhD,[4] Anette Bakkane Bendixen, MD,[4] Gerard J. Byrne, PhD,[5] María Márquez-González, PhD,[6] Andres Losada, PhD,[7] Lei Feng, PhD,[8] Elisabeth Kuan Tai Ow, PhD,[8] Kullaya Pisitsungkagarn, PhD,[9] Nattasuda Taephant, PhD,[9] Somboon Jarukasemthawee, PhD,[9] Alexandra Champagne, D.Psy,[10] Philippe Landreville, PhD,[10] Patrick Gosselin, PhD,[11] Oscar Ribeiro, PhD,[12], Gretchen J. Diefenbach, PhD,[13] Karen Blank, M.D,[13] Sherry A. Beaudreau, PhD,[14] Jerson Laks, PhD,[15] Narahyana Bom de Araújo, PhD,[15] Rochele Paz Fonseca, PhD,[16] Renata Kochhann, PhD,[16] Analuiza Camozzato, PhD,[17] Rob van den Brink, PhD,[18] Mario Fluiter, MD,[19], Paul Naarding, PhD,[20] Loeki P.R.M. Pelzers, MD,[21] Astrid Lugtenburg, MD,[22] Richard Oude Voshaar, PhD,[18] and Nancy A. Pachana, PhD.[23]

[1]Department of Clinical Psychology University of Bergen, Bergen, Norway

[2]Department of Behavioural Sciences in Medicine, University of Oslo, Oslo, Norway

[3]Department of Psychosocial Science, University of Bergen, Bergen, Norway

[4]Department of Geriatric Psychiatry, Diakonhjemmet Hospital, Oslo, Norway and Norwegian National Advisory Unit on Aging and Health, Vestfold Hospital Trust, Toensberg, Norway

[5]School of Clinical Medicine, The University of Queensland, Brisbane, Australia

[6]Faculty of Psychology, Universidad Autónoma de Madrid, Madrid, Spain

© The Author(s) 2019. Published by Oxford University Press on behalf of The Gerontological Society of America. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

[7]Universidad Rey Juan Carlos, Madrid, Spain

[8]Department of Psychological Medicine, Yong Loo Lin Shchool of Medicine, National University of Singapore, Singapore, Singapore

[9]Faculty of Psychology, Chulalongkorn University, Bangkok, Thailand

[10]School of Psychology, Laval University, Quebec, QC, Canada

[11]Department of Psychology, Université de Sherbrooke, Sherbrooke, Canada

[12]Center for Health Techonology and Services Research (CINTESIS), Department of Education and Psychology, Aveiro University, Aveiro, Portugal

[13]Institute of Living, Hartford, CT, USA

[14]Sierra Pacific Mental Illness Research Education and Clinical Center, VA Palo Alto Health Care System and Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA, USA

[15]Institute of Psychiatry, Universidade Federal do Rio de Janeiro, Brazil

[16]Pontifical Catholic University of Rio Grande do Sul (PUCRS), Brazil

[17]Universidade Federal de Ciências da Saúde de Porto Alegre UFCSPA, Brazil

[18]University of Groningen, University Medical Center Groningen, Rob Giel Research center (RGOc), The Netherlands

[19]Mental Health Center GGZ Noord-Holland Noord, Heerhugowaard , The Netherlands

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

[20]GGNet Mental Health, Division of Old Age Psychiatry, Warnsveld & Apeldoorn, The Netherlands

[21]Mental Health Center GGZ Centraal, Ermelo, The Netherlands.

[22]Mental Health Center GGZ Drenthe, Assen, The Netherlands.

[23] School of Psychology, The University of Queensland, Brisbane, Australia

Corresponding author:

Helge Molde

Department of Clinical Psychology

Christiesgate 12, 5020 BERGEN, Norway

phone: 0047 55588682

e-mail: **helge.molde@uib.no**

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Accepted Manuscript

Abstract

Objectives. Assessing late-life anxiety using an instrument with sound psychometric properties including cross-cultural invariance is essential for cross-national aging research and clinical assessment. To date, no cross-national research studies have examined the psychometric properties of the frequently used Geriatric Anxiety Inventory (GAI) in depth.

Method. Using data from 3,731 older adults from 10 national samples (Australia, Brazil, Canada, The Netherlands, Norway, Portugal, Spain, Singapore, Thailand, USA), this study used bifactor modelling to analyze the dimensionality of the GAI. We evaluated the "fitness" of individual items based on the explained common variance for each item across all nations. In addition, a multigroup confirmatory factor analysis (MG-CFA) was applied, testing for measurement invariance across the samples.

Results. Across samples, the presence of a strong G factor provides support that a general factor is of primary importance, rather than subfactors. That is, the data support a primarily unidimensional representation of the GAI, still acknowledging the presence of multidimensional factors. A GAI score in one of the countries would be directly comparable to a GAI score in any of the other countries tested, perhaps with the exception of Singapore.

Discussion. Although several items demonstrated relatively weak common variance with the general factor, the unidimensional structure remained strong even with these items retained. Thus, it is recommended that the GAI be administered using all items.

Keywords: bifactor, unidimensional, multigroup, measurement, invariance

Introduction

The Geriatric Anxiety Inventory (GAI) is one of most frequently used scales measuring anxiety in older adults in native English speaking countries like Australia, USA and Canada. It has also been translated into several languages including Chinese, Dutch, Finnish, French, Norwegian, Portuguese, Spanish and Thai. The GAI consists of 20 items measuring the severity of anxiety symptoms of older adults (Pachana et al, 2007). The scale has demonstrated excellent reliability and internal consistency, discriminant ability and predictive and convergent validity (Kneebone, Fife-schaw, Lincoln & Harder, 2016). Assessing geriatric anxiety with a sound psychometric instrument is essential for research, and for comparing research findings from different countries and across various languages. For example, to be able to to compare group means from different national samples, observed sumscores should not be compared without first testing that the observed means are equal to the latent mean. Without an empirical demonstration that the GAI has measurement invariance, i.e., that the intercepts and loading of the indicators are equal across the nations compared, it is unknown if GAI sumscores have the same meaning and statistical properties across the nations.

As a special case, configural measurement invariance implies that the GAI should have the same configural factor structure across countries. However, to date, across different countries and samples, independent reports of the factor structure of the GAI have yielded divergent findings. Exploratory factor analysis (EFA) has reported a two factor solution (Bendixen et al, 2016), a three factor solution (Márquez-González et al, 2012; Guan, 2016) and a four factor

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

solution (Diefenbach, Bragdon & Blank, 2014, Loynachan, Lee, Lamb & Pyykkonen, 2016). However, a unidimensional factor structure has also been reported using traditional CFA (Byrne & Pachana, 2011; Jochno, Knight, Tadic & Wuthrich, 2015), and by using item response models (Molde et al, 2017; Yan, Xin, Wang & Tang, 2014). As reported by Molde et al (2017), the lack of factorial consistency across studies might reflect different cultural response styles, differences in semantics due to translation processes, together with varying sample characteristics. However, the variation might also reflect true cultural differences in the structure of anxiety across countries.

The different factor solutions and findings are not necessarily contradictory, as anxiety might consist of both general (unidimensional) and specific factors (Molde et al., 2017; Reise, 2012; Rodriguez, Reise & Haviland, 2016a, 2016b). To identify general and specific factors, EFA is not the ideal methodological approach, because such models typically would view unidimensional and multidimensional models as alternative models, and specific factors as a violation of unidimensionality, rather than integral parts of the model. A situation with general and specific factors can be modelled with bifactor models evaluating the presence and strength of a general unidimensional G factor and simultaneously allowing the presence of specific factors and sub-dimensions (Ss). Bifactor analysis would allow for testing whether the GAI is unidimensional or not across countries, and could provide a model that would identify the overall common structure across countries, while at the same time also including specific national factors. Hence, bifactor analysis of different national samples could establish a more solid evidence base regarding the factor structure of the GAI, and allow us to test if the GAI has measurement invariance across different nations.

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

With this aim in mind, we decided to test if the same configural model exists independently in each national sample. That is, we expected that the GAI would have a strong G (unidimensional) factor in each sample, as bifactor models often provide better fit to data than traditional CFAs, given that the GAI includes a heterogeneous set of scale items, representing different manifestations of anxiety (e.g. somatic or cognitive) (Rodriguez et al., 2016a). We also expected subfactors to differ in different nations, as the latent trait "Anxiety" could have different local cultural variations in expression and responses to the scale items. We also aimed to identify items that functioned poorly within the scale across nations. Importantly, establishing a strong G factor across the national samples would indicate unidimensionality. This would allow us to test the GAI for measurement invariance across the national samples included.

Method

Sample

One large dataset was constructed by merging 10 different GAI samples shared by the coauthors. The national datasets provided were collected from previous published studies; see Table 1 for simple descriptive study details and references to the original studies. We included only adults aged 65 years or older in the analysis, as this age is typically used as indexing the beginning of older adulthood. The grand N overall is 3,731. Age ranged from 65 to 98 across the samples (grand M=74.0, SD=6.57) with the proportion of males ranging from 0.27 to 0.45. Also, see http://gai.net.au/ for further information about the GAI.

Regarding translation, all non-english speaking countires applied the standard formal forward-backward translation procedure for translating instruments (World Health Organization,

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

2016). In addition, the Portuguese, Singapore, Spanish and Thai GAI versions were translated and adapted in collaboration with the original scale creators.

Reportedly, the items referring to somatic sensations, e.g. items 7 ("*I often feel like I have butterflies in my stomach*"), item 12 ("*I can get an upset stomach due to my worrying*") and item 18 ("*I sometimes feel a great knot in my stomach*"), which also are idiomatic expressions, were somewhat difficult to translate into Norwegian, Portuguese and Spanish (Màrquez-Gonzàlez, Losada, Fernàndez-Fernàndez, & Pachana, 2012; Molde et al., 2017; Ribeiro, Paúl, Simões & Firmino, 2011). Notably, the reference to butterflies in item 7 is not common in Spain, hence in the Spanish GAI version, a similar item expression is used referring to ants ("A menudo siento hormigueo en mi estómago"). See the original studies for further information regarding the translation processes.

Regarding cultural identity and preferred language of the participants, the data reported indicate that the majority of the participants speak, read and write the main language of the respective country. For example, in the Spanish sample, all participants are native Spanish citizens (Europeans), while in the Canadian sample, the participants at least could read and write French. For the Singapore sample, the majority of the particpants were ethnic Chinese (> 90 %, preferring one of three different Chinese dialects - Mandarin, Hokkien & Cantonese). The majority of the US sample is Caucasian (80-90 %), while for the Dutch sample, more than 95% of the sample is Caucasian, and native speakers of Dutch. The majority of the Norwegian participants were also Caucasian, speaking native Norwegian. For all details, see Table 1.

Ethical consideration

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

The Regional Committees for Medical and Health Research Ethics, REK-Vest (2016/1015), in Norway, approved the current study. All 10 datasets were collected with Human Ethics Committee approvals in each country, and all dataset custodians agreed in writing to their deidentified data being shared to construct the present large dataset.

Statistics

All the main analyses were conducted using the statistical software R (R Core Team, 2018). First, a confirmatory bifactor model was run in order to evaluate if a (configural) unidimensional model would fit the data across each individual national sample. Evaluating whether the GAI is unidimensional or not, we checked three different criteria proposed by Rodriguez, Reise and Haviland (2016a), using the omegaSem function of the *psych* package in R (Revelle et al, 2018). The first criterion was the presence of a strong general G factor through the index “coefficient omega hierarchical” (ωH). ωH estimates the proportion of variance in the total scores that can be explained by a single general factor. The second criterion was the explained common variance (ECV), defined as the ratio of explained variance by the G factor, divided by the variance explained by both the general factor and the specific factors. A high ECV indicates a strong G factor, even in the presence of multidimensionality. The third criterion applied for this purpose was the “percentage of uncontaminated correlations” (PUC). The PUC values were calculated using a Microsoft-excel based calculator (Hammer, 2016). As PUC values become larger, the general factor in the bifactor model becomes more similar to the single dimension estimated in a unidimensional model, especially when ECV is high. When PUC values are < 0.80, and general ECV values are > 0.60 and $\omega H > 0.70$, this suggests that the presence of some multidimensionality is not severe enough to disqualify the interpretation of the scale as primarily

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

unidimensional (Reise, Scheines, Widaman, & Haviland, 2013). Similarly, when ECV is > 0.70 and PUC > 0.70, it is suggested that the relative bias is small, and the common variance should be interpreted as indicative of an essentially unidimensional scale (Rodriguez et al., 2016b). Thus, essential unidimensionality is stronger than primarily unidimensionality, due to the prescence of some (non-severe) multidimensionality in the latter term.

In order to evaluate the "goodness" of individual GAI items, or items that can be selected to create a more unidimensional scale, the ECV for each item across all nations was computed. I-ECV is the percent of the item's common variance that is attributed to the general factor (G). A cut point of 0.80 is recommended for strong items (Rodriguez et al., 201a).

After evaluating the configural unidimensionality, multi-group confirmatory factor analysis (MG-CFA) was applied. The MG-CFA rely on multivariate normal assumptions. Thus we evaluated the normality assumptions, using skew and kurtosis of the GAI sumscore as indicators. Values for skew and kurtosis between -2 and +2 are often considered acceptable in order to infer normal distribution (George & Mallery, 2010). The MG-CFA analyses were conducted using the lavaan package in R (Rosseel, 2011). MG-CFA is the de-facto standard to investigate if scales are invariant across groups (Chen, 2008). The procedure for measurement invariance testing consists of a series of repeated model comparisons in which a more restricted model is compared to a less restricted model (Byrne & van de Vijver, 2010; Davidov, Meuleman, Cieciuch, Schmidt & Billiet, 2014; Hirschfield & von Brachel, 2014; Pendergast, von der Embse, Kilgus & Eklund, 2017; Putnick & Bornstein, 2016). The nested model comparisons enable tests of configural, metric, and scalar equivalence.

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

First, a baseline model is fitted where the loading pattern is similar across all groups but all parameters are free, meaning that loadings, intercepts and variances are free to vary. Configural invariance, the weakest level of invariance, requires that the strucure of GAI is the same across all countries. Thus, the fit of a model where items load on the same factor would support configural invariance. The next level of invariance, metric invariance, is tested by making the additional assumption that factor loadings are equal across countries. Metric invariance means that every item of the GAI contributes to the latent construct to a similar degree across all countries. This means that the intervals of the scale are equal across all countries, and importantly, any difference or covariance within a country can be compared with differences and covariances in other countries. The highest level of invariance, scalar invariance, adds the assumption that not only factor loadings, but also item intercepts, are equal across the countries. Strong, or scalar, invariance means that person with the same latent score would also have the same expected item score. The implication of scalar invariance is that raw scores of the GAI, and also the mean of the latent construct of the GAI can be compared across all countries. Thus, scale invariance would allow researchers to make the assertion that anxiety scores in country A are lower than in country B, based only on the observed sumscores. This is because the GAI in country A would have the same psychometric properties as the GAI in country B. Thus, any observed differences between the countries would not be due to instrument bias.

Notably, due to the categorical items within the GAI, weak and strong (metric) invariance were tested in one single step in our analysis (Putnick & Bornstein, 2016). Finally, if the strong invariance comparison tests holds, a strict (residual) invariance model is specified. In this model, factor loadings, intercepts, and residual variances are constrained to be equal across all countries.

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

This means that the sum of the specific variance and error variance is similar across all countries. Again, this model is compared with the strong (scalar) invariance model.

If all the model comparisons hold, we may conclude that measurement invariance has been established for the GAI, hence any observed mean differences can be attributed to real differences in anxiety between the countries, and not attributed to instrument bias.

Evaluating the absolute model fit, we apply the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA) and the Standardized Root Mean-square Residual (SRMR). The rule of thumb for absolute fit for CFI is that a model fit is considered adequate if the CFI values are > 0.90, and acceptable if the CFI values are > 0.95. For RMSEA, values < 0.08 represent reasonable fit, while values < 0.05 indicates good fit. SRMR values close to 0.08 indicate acceptable fit (Putnick & Bornstein, 2016).

Most often a cutpoint of ΔCFI < 0.01 is chosen to decide whether a more constrained model, e.g. the weak-invariance model, shows a substantial decrease in model fit compared to a less constrained model, e.g. the baseline model. For ΔRMSEA, a cutpoint of 0.015 is recommended. For ΔSRMR, 0.03 and 0.15 are recommended for weak, and strong and/or residual, respectively. Furthermore, in samples with groups of 10 or more countries, a ΔCFI of 0.02 and RMSEA of 0.03 are considered appropriate for testing metric invariance, while using the ΔCFI < 0.01 and RMSEA < 0.01 for scalar invariance tests (Putnick & Bornstein, 2016).

Results

Table 2 displays results from the bifactor analysis. As seen from the table, the index “coefficient omega hierarchical” (ωH) is strong in all countries, ranging from 0.71 (Singapore) to 0.92

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

(Thailand). The explained common variance (ECV) ranges from 0.62 (Singapore) to 0.85 (Thailand), although most ECV values are above 0.70. The “percentage of uncontaminated correlations”, or PUC values ranges from 0.38 (Singapore) to 0.80 (Canada). Overall, the results provide an empirical base to support the prescence of a strong G factor across all countries. The general and specific factor structure for each nation is available as tables in the supplementary material. 7 countries had three specific, but weak factors in addition to the strong G factor. Portugal and Singapore had two specific factors, while Canada had four specific but weak factors. As expected, for different nations, the items loaded on different subfactors with different factor loadings. Notably, no common subfactors across the countires appeard. Still, the individual items 7, 12 and 18 seemd to cluster themselves together within a subfactor across 9 of 10 samples, but, and importantly, no configural identity appears to exist for the specific subfactors across the nations.

See Table 3 for the i-ECV values of all individual items across each nation. For Norway, Portugal and Spain, the items 7,12, and 18, which reportedly were difficult to translate, had the lowest i-ECV *ranked* values within these countries. As such, there is some pattern that indicate that these three specific items within most countries have i-ECV values among the lower overall ranked items. Otherwise, there seems to be no pattern across the samples. Also, and notably, item 11 (“*My own thoughts make me anxious*”) was the sole item that passed the 0.80 criterion across every nation.

The results of skew and kurtosis evaluation were within the normal range for all countries except Singapore (skew=2.68, kurtosis=7.53). In order to evaluate the influence of this non-

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

normality for one specific sample, we ran two MG-CFA models, one with all countries included, and one with Singapore left out.

For the results of the multigroup confirmatory factor analysis (MG-CFA), see Table 4 and 5. In the analysis using all countires, all model CFI values are above 0.992 and all the RMSEA values vary between 0.044 and 0.054. Thus, the absolute CFI and RMSEA values indicate acceptable and reasonable fit, respectively. The SRMR values indicate values close to acceptable fit for the configural model (0.084), but not for the strong and strict models (0.088). Thus, for two of three fit indices, the absolute model fit was acceptable or reasonable for all models. Regarding the results from the model comparison tests, all steps and models passed the ΔCFI criterion < 0.01. For ΔRMSEA, all model comparisons passed the cutpoint of 0.01, also.

For the models leaving Singapore out, all model CFI values are above 0.993 and all the RMSEA values vary between 0.046 and 0.055. The SRMR vales vary between 0.74 and 0,75. Thus, the absolute CFI, SRMR and RMSEA values indicate acceptable and reasonable fit, for all models respectively. Regarding the results from the model comparison tests, all steps and models passed the ΔCFI criterion < 0.01. For ΔRMSEA, all model comparisons passed the cutpoint of 0.01. Also, using ΔSRMR of 0.15, all model comparisons passed the cutpoint. Thus, for all three fit indices, both the absolue and the relative model fit were acceptable or reasonable.

Thus, overall, for at least 9 of 10 countries, it seems that we may conclude that any differences in anxiety as measured by the GAI are not attributable to instrument bias.

Discussion

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Our first aim was to evaluate the dimensionality of the GAI across the nations using bifactor analysis. The results support the presence of a strong G factor across nations and samples. That means that the majority of the variance in the factor structure is explained by a common or general factor – *Anxiety*. These results are in line with and extend the findings reported by Molde et al (2017). The scoring of the GAI may thus be based on a sum of all items across the scale. Importantly, the bifactor analyses shows that each language included in our study may base its scoring on a total sumscore, because bifactor models were run independently for each sample.

Still, the GAI is a mixture of items that cover different dimensions of anxiety, eg. cognitive and somatic expressions, in a different degree. However, and as expected, the different subfactors across different nations varied, with most items, loading on different subfactors, with different factor loadings. We noticed, however, that the individual items 7, 12 and 18 seemd to cluster themselves together within some form of a subfactor across 9 0f 10 samples. But, and most important, no configural identity appears to exist overall for the weak subfactors across the nations.

Regarding the items 7, 12 and 18, the content of seems to cover somatic sensations, or a somatic factor. However, items 7 and 12 are, as said earlier, also idiomatic expressions, which have varying degrees of difficulty being translated (and often similar idiomatic expressions are used instead of a direct translation). This could also be a contributing factor here.

Evaluating the "goodness" of individual GAI items, or items that can be selected to create a more unidimensional scale, the three items – 7, 12 and 18 - seems to be "candidate" items across the nations. That is, omitting these items would strengthen the G-factor and make the GAI more unidimensional, in most samples. Notably, the Singapore version of the GAI would seem to

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

benefit from this, having the lowest i-ECV values for these items. On the other hand, the results overall (Omega H, ECV and PUC values) support the acceptability of using the scale with all items, thus a revision in light of these results is not absolutely needed, nor recommended. Thus, we do not recommend scoring and use of any subfactors, e.g. a somatic factor score, for the national samples included in this study. Notably, omitting any items might change the content validity, thus we discourage such an idea. One may speculate that the reason that these three "somatic" items are candidate items is due to their content, as older adults commonly experience somatic symptoms related to a chronic medical condition or illness, thus possibly confound the reporting of anxiety symptoms on measures that covers multiple somatic items (Molde et al, 2017). Also, translation issues may play a part, as already mentioned. Thus, refinements in translations of these items could be examined in countries where the items are not performing as well.

Our final aim was to test for measurement invariance, to ascertain if differences in test responses are due to pure chance rather than characteristics of the national groups. For nine of the ten countries, the results strongly supported measurement invariance across the nations sampled. When using all data, with Singapore included, results diverged, as the SRMR absolute fit value indicated poor fit, other than at the configural level. However, the CFI and RMSE absolute and relative fit indices indicated acceptable fit. The mixed results with the Singapore sample included, could be due to non-normality of that specific sample, or also language and translation of the GAI, as the participants in the Singapore sample speak three different Chinese dialects. Hence, for the Singapore version of the GAI, we recommend a closer look at the translation, and also a future reanalysis using a sample with normal distribution of the GAI score.

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

To sum up, and still acknowledging some mixed findings for Singapore, the GAI passed all the MG-CFA test steps, and thus any observed mean differences can be attributed to differences in the underlying construct (self-reported late-life anxiety) between the nations tested. From a research as well as from a clinical point of view, it means that a GAI score in one of the countries in the current study is directly comparable to a GAI score in another of the countries sampled. This however, only applies for the national samples included in this paper. The findings can not be generalized to nations or language samples not tested within this comparison. Further research could extend these findings using different national samples. Future studies could also extend this research, testing the psychometric properties of the Geriatric Anxiety Inventory short form (GAI-SF; Byrne & Pachana, 2011) across cultures and languages. Across samples one may identify how the items cover the latent *anxiety* trait, as there is some indication that the item overlap may be substantial (Molde et al., 2017). Evaluating the items' discriminative threshold for diagnostic anxiety categories, using cognitive diagnostic models, would further enhance the possibility to evaluate each individual item, in this respect. Lastly, as there may be cultural specific expression of anxiety, a network analysis of the GAI items across countries would allow a comparison of the relationship and centrality of different anxiety symptoms, across cultures and samples.

Strengths and limitations

One caveat of this study is the use of collapsed data. The US sample was created by collapsing two independent samples, in order to have a sufficient sample size for the USA dataset. The same was true for the Brazilian sample, which, in addition, also had a small N. However, running the MG-CFA models without the Brazilian data did not change the results in any way. Thus, any

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

issues due to small sample size and power seems not to have affected the results. The MG-CFA procedure is regarded as the most rigorous testing approach when testing for measurement invariance. That our second model was supported and passed all test criteria is a strength, given that even trivial or slight deviation in the parameters would have resulted in lack of invariance. Thus, our results give us confidence that the GAI has measurement invariance across the countries tested. That said, a complex latent trait as “Anxiety” would also be likely to have much local variation in expression. Thus, it might be that the GAI is measuring some common component of anxiety, but still missing important cultural variation, not covered by the items within the scale. However, a potential limitation is that we have no evidence that the samples are “truly” representative of each nation.

Conclusion

The results seem to confirm and support unidimensionality in a bifactor model of the GAI across all national samples tested, due to the presence of a strong unifying G-factor, relative to the multidimensional specific factors. The scoring of the GAI may thus be based on using a sum score of all items across the scale. Furthermore, for the majority of the items, there appears to be no specific configural subfactor structure across the nations. The three items 7, 12 and 18, however, seems to belong to a specific subfactor in nine of the ten samples. The content of these three items seems to cover somatic sensations, or a somatic factor. Thus, omitting item 7, 12, and 18 from the scale would enhance the unidimensionality of the GAI, if wished for, but is not necessary given the data at hand. Notably, omitting any items might change the content validity, thus we discourage any such practice. Importantly, the results imply that a GAI score in one of

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

the countries tested would be directly comparable to a GAI score in one of the other countries tested, with Singapore as an exception.

Conflict of Interest

None.

Funding

Data for part of the USA data set was supported by an Alzheimer’s Association Grant (NIRG-09-133592) to S. A. Beaudreau and by a National Institute of Mental Health Grant (R03 MH071575) to G. J. Diefenbach.

Data for part of the Singapore data set was supported by the National University of Singapore [grant number VG-8] and the National Medical Research Council of Singapore [grant number NMRC/TA/0053/2016].

References

Byrne GJ. (2014) Validation of the informant questionnaire for anxiety in dementia in a memory clinic cohort. *International Psychogeriatrics*, 25 (Suppl. 1): S26.

Byrne, B. M., & van de Vijver, F.J. R. (2010) Testing for Measurement and Structural

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Equivalence in Large-Scale Cross-Cultural Studies: Addressing the Issue of Nonequivalence. *International Journal of Testing*, *10*(2), 107-132, DOI:10.1080/15305051003637306

Byrne, G. J., & Pachana, N. A. (2011). Development and validation of a short form of the Geriatric Anxiety Inventory – the GAI-SF. *International Psychogeriatrics*, *23*, 125–131.

Byrne, G. J., Pachana, N. A., Nur, F., Kua, E-H., Mahenran, R., Jiang, M-J., Feng, L. (2016). Observations on the Geriatric Anxiety Inventory (GAI) in a large Singaporean cohort. *International Journal of Psychology*, *51*, Suppl. 1, 37.

Champagne, A., Landreville, F., Gosselin, P., & Carmichael, P-H. (2016). Psychometric properties of the French Canadian version of the Geriatric Anxiety Inventory. *Aging & Mental Health*, *22*(1), 40-45, DOI: 10.1080/13607863.2016.1226767

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology*, *40*(1), 55-75. https://doi.org/10.1146/annurev-soc-071913-043137

Diefenbach, G. J., Bragdon, L. B., & Blank, K. (2014). Geriatric Anxiety Inventory: Factor

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

structure and associations with cognitive status. *American Journal of Geriatric Psychiatry*, *22*(12), 1418–1426.

Diefenbach, G. J., Tolin, D. F., Meunier, S., & Gilliam, C. (2009). Assessment of Anxiety in Older Home Care Recipients. *The Gerontologist*, 49, 141-153.

Feng, L. (2015). Ageing in a Community Environment Study (ACES) Cohort. In N.A. Pachana (Ed.), E*ncyclopedia of Geropsychology* (pp. 1-6), Springer Singapore, Singapore.

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference*, 17.0 update (10a ed.) Boston: Pearson.

Gould, C. E., Segal, D. L., Yochim, B. P., Pachana, N. A., Byrne, G. J., & Beaudreau, S. A. (2014). Measuring anxiety in late life: A psychometric examination of the Geriatric Anxiety Inventory and Geriatric Anxiety Scale. *Journal of Anxiety Disorders*, 28, 804-811. https://doi.org/10.1016/j.janxdis.2014.08.001

Guan, M. (2016). Factor structure of the Chinese version of the Geriatric Anxiety Inventory. *Annals of General Psychiatry*, *15*(4). DOI 10.1186/s12991-016-0092-4

Hammer, J. H. (2016, October). Percent of uncontaminated correlations (PUC) calculator: a Microsoft Excel-based tool to calculate the PUC statistic. Available from http://DRJosephHammer.com/

Hirschfeld, G., & von Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R – A

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, *19*(7), 1-12.

Kneebone, I. I., Fife-Schaw, C., Lincoln, N. B. & Harder, H. (2016). A study of the validity and the reliability of the Geriatric Anxiety Inventory in screening for anxiety after stroke in older inpatients. *Clinical Rehabilitatio*n. DOI: 10.1177/0269215515619661

Loynachan, C., Lee, M., Lamb, M., & Pyykkonen, B. (2016). Factor structure comparison between the Beck Anxiety Inventory and Geriatric Anxiety Inventory for a non-clinical sample of older adults. Abstract: Poster session, C-65. *Archives of Clinical Neuropsychology*, *31*, 584–637.

Massena, P., De Araújo, N., Pachana, N., Laks, J., & De Pádua, A. (2016). Validation of the Brazilian Portuguese Version of Geriatric Anxiety Inventory – GAI-BR. *International Psychogeriatrics*, *27*(7), 1113-1119. doi:10.1017/S1041610214001021

Màrquez-Gonzàlez, M., Losada, A., Fernàndez-Fernàndez, V., & Pachana, N. (2012). Psychometric properties of the Spanish version of the Geriatric Anxiety Inventory. *International Psychogeriatrics*, 24, 137–144.

Molde, H., Hynninen, K., Torsheim, T., Bendixen, A., Engedal, K., Pachana, N., & Nordhus, I. (2017). A Bifactor and item response analysis of the geriatric anxiety inventory. *International Psychogeriatrics*, *29*(10), 1647-1656. doi:10.1017/S1041610217001004

Oude Voshaar RC, Dhondt ADF, Fluiter M, Naarding P, Wassink S, Smeets MMJ, Pelzers

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

LPRM, Lugtenburg A, Veenstra M, Marijnissen RM, Hendriks GJ, Verlinde AA, Schowevers RA, van den Brink RHS. (submitted). Study design of the Routine Outcome Monitoring for Geriatric Psychiatry & Science (ROM-GPS) project; a cohort study of older patients with affective disorders referred for specialised geriatric mental health care.

Pachana, N., Byrne, G., Siddle, H., Koloski, N., Harley, E., & Arnold, E. (2007). Development and validation of the Geriatric Anxiety Inventory. *International Psychogeriatrics*, *19*(1), 103-114. doi:10.1017/S1041610206003504

Pisitsungkagarn, K., Taephant, N., Jarukasemthawee, S., & Pachana, N. (2016). The development of the Thai version of the Geriatric Anxiety Inventory (GAI). *International Journal of Psychology*, *51*, Suppl. 1, 37. CS28-11-2?

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: a non-technical primer on categorical multi-group confirmatory factor analysis in school psychology. *Journal of school Psychology*, *60*, 65-82. http://dx.doi.org/10.1016/j.jsp.2016.11.002

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review*, 41(9), 71-90. https://doi.org/10.1016/j.dr.2016.06.004

R Core Team. (2018). R: A language and environment for statistical computing. *R Foundation*

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*, 667–696.

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*, 129-140. https://www.ncbi.nlm.nih.gov/pubmed/23030794

Revelle, W. (2018) *psych: Procedures for Personality and Psychological Research*, Northwestern University, Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.8.4.

Ribeiro, O., Paúl, C., Simões, M.R., & Firmino, H. (2011) Portuguese version of the Geriatric Anxiety Inventory: Transcultural adaptation and psychometric validation. *Aging & Mental Health*, 15(6), 742-748, DOI: 10.1080/13607863.2011.562177

Rodriguez, A, Reise, S. P., & Haviland, M. G. (2016a). Evaluaing bifactor modles: calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137-150.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016b). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, *98*, 223-237. doi:10.1080/00223891.2015.1089249

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

World Health Organization (2016). *Process of Translation and Adaption of Instruments*. Retrieved from http://www.who.int/substance_abuse/research_tools/translation/en/

Accepted Manuscript

Table 1. Descreption of all samples

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Country	Authors	Sample variables	N	Mean	SD	Language / ethnicity
Australia	Byrne (2014)	Age		78.5	6.29	Original language 93.5% of the participants had English as their primary language 84.9 % were born in Australia
		Male		0.45[a]	0.50	
		GAI score	192	2.4	2.82	
Brazil	Massena et al. (2016) & data from an ongoing study	Age		72.2	5.18	All participants speech Portugese as their main / first language
		Male		0.45[a]	0.50	
		GAI score	119	4.22	3.0	
Canada (French-Canadian)	Champagne et al (2016).	Age		74.6	6.64	French

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

		Male		0.27[a]	0.44	
		GAI score	301	2.2	3	
The Netherlands	Voshaar et al. (submitted)	Age		74.4	6.13	All native Dutch speakers
		Male		0.39[a]	0.49	95-100 % caucasian
		GAI score	785	5.65	3.0	
Norway	Molde et al. 2017	Age		76.4	7.17	Norwegian
		Male		0.32[a]	0.47	Almost all participsants were Caucasian, with Norwegian as primary language
		GAI score	512	4.41	3.31	
Portugal	Ribeiro et al. (2011)	Age		74.2	6.13	All participants had Portugese as their main language
		Male		0.39[a]	0.49	
		GAI score	204	5.54	3.71	

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Singapore	Feng (2015) & Byrne et al. (2016)	Age		70.8	5.0	Chinese languge (Mandarin, Hokkien and Cantonese[b])
		Male		0.33[a]	0.47	
		GAI score	611	0.72	1.54	
						➢ 90 % ethnic Chinese ➢ Less than 10 % English speaking
Spain	Màrquez-Gonzàlez et al. (2012)	Age		73.7	5.98	Spanish
		Male		0.29[a]	0.45	Spanish citizens (Caucasian)
		GAI score	436	4.33	3.1	
Thailand	Pisitsungkagarn et al (2016)	Age		71.7	6.24	Thai / Thai citizens

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

		Male		0.37[a]	0.47	
		GAI score	275	3.80	3.77	
USA	Diefenbach et al. (2009; 2014) + unpublished data & Gould, Segal, Yochim, Pachana, Byrne, & Beaudreau (2014)	Age		74.6	6.65	Original language ➢ 80-90 % Caucasian,
		Male		0.37[a]	0.49	
		GAI score	235	2.22	2.80	

---

[a] Proportion of males; [b] a form of Chinese spoken by over 54 million people, mainly in southeast China (including Hong Kong)

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Table 2. Configural bifactor model across countries

	Aust.	Brazil	Can.	The Netherl.	Nor.	Port.	Singap.	Spain	Thail.	USA
Omega H (ωH)	0.88	0.83	0.90	0.85	0.87	0.89	0.71	0.85	0.92	0.88
ECV	0.75	0.69	0.80	0.66	0.82	0.83	0.62	0.74	0.85	0.79
PUC	0.62	0.72	0.80	0.72	0.55	0.41	0.38	0.71	0.65	0.72

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Table 3. The i-ECV values of all individual items across each nation

	Country									
	Australia	Brazil	Canada	The Netherl.	Norway	Portugal	Singapore	Spain	Thailand	USA
GAI item	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV	i-ECV
1	0.97	0.89	0.99	0.83	1.00	0.96	0.87	0.59	0.78	1.00
2	1.00	0.95	0.72	1.00	0.68	0.92	0.69	0.96	0.65	0.84
3	0.78	1.00	0.89	0.97	0.78	0.83	0.74	0.98	0.94	0.70
4	0.83	0.98	0.99	0.92	0.91	0.90	0.88	0.80	0.92	0.73
5	**0.35**	0.86	0.71	0.70	0.83	0.97	**0.38**	0.62	0.93	0.71
6	0.99	0.55	0.97	0.93	0.65	0.96	0.89	0.90	0.59	0.99
7	**0.38**	0.47	0.98	0.54	**0.39**	0.60	**0.16**	0.50	0.83	0.42
8	0.85	0.96	0.68	0.96	0.95	0.75	0.97	0.79	1.00	0.97
9	1.00	0.79	0.77	1.00	0.72	0.88	0.99	0.93	0.97	0.96
10	0.56	0.96	0.55	**0.22**	0.81	0.72	1.00	0.45	0.98	0.91

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

11	1.00	0.89	0.94	0.92	0.99	0.86	0.98	0.90	0.99	0.90
12	0.61	**0.33**	**0.32**	**0.38**	0.63	0.67	**0.34**	0.41	0.83	0.46
13	0.55	0.83	0.69	0.85	0.89	0.65	0.96	0.53	1.00	0.98
14	0.48	0.63	0.85	0.96	0.99	0.97	0.40	0.99	0.92	1.00
15	0.81	0.76	0.85	0.99	0.62	0.86	0.85	0.74	0.99	0.77
16	0.92	0.58	0.94	0.49	0.99	0.92	**0.24**	0.85	1.00	0.81
17	0.95	0.44	0.94	0.60	1.00	0.99	0.71	0.96	0.70	0.67
18	0.73	**0.29**	0.53	0.41	0.43	0.41	**0.10**	0.56	0.61	0.61
19	1.00	0.74	0.62	0.96	0.96	0.97	**0.17**	0.96	0.70	0.82
20	0.96	0.77	0.96	0.93	0.85	0.98	0.69	0.87	0.84	0.77

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Table 4. Multiple-Group confirmatory factor analysis (MG-CFA) – all countries

Model	RMSEA	SRMR	CFI	ΔCFI	Desicion
Configural	0.044	0.084	0.995	-	Accept
Strong	0.054	0.088	0.992	0.003	Accept
Strict	0.054	0.088	0.992	0.000	Accept

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019

Table 5. Multiple-Group confirmatory factor analysis (MG-CFA) – without Singapore

Model	RMSEA	SRMR	CFI	ΔCFI	Decision
Configural	0.046	0.074	0.995	-	Accept
Strong	0.055	0.075	0.993	0.002	Accept
Strict	0.055	0.075	0.993	0.000	Accept

Accepted Manuscript

Downloaded from https://academic.oup.com/psychsocgerontology/advance-article-abstract/doi/10.1093/geronb/gbz002/5280714 by Universitetsbiblioteket i Bergen, helge.molde@uib.no on 09 January 2019