Measurement Invariance of Personal Well-being Index in 26 Countries

Magdalena Żemojtel-Piotrowska, University of Gdańsk;

Jarosław Piotrowski; University of Social Sciences and Humanities;

Jan Cieciuch; Zurich University;

Byron Adams, Utrecht University

Rakhman Ardi;

Sergiu Baltatescu; University of Oradea;

Arbinda Bhomi;

Amanda Clinton; University of Puerto Rico

Gisela T. de Clunie

Valdiney Gouveia;

Ashrafalsadat Hosseini; University of Melbourne

Narine Kachatryan; Erevan University

Shanmukh Kamble;

Anna Kawula; Jagiellonian University;

Eva Letovancova, Comenius University

Kadi Liik, Talinn University

Sara Malo Cerrato; University of Girona

John Maltby, Leicester University

Sofya Nartova-Bouchaver;

Marija Nikolic;

Evgeny Osin; University of Moscow

Joonha Park; University of Nagoya

Elena Paspalanova

Gyozo Pek; University of Debrecen

Joanna Różycka-Tran; University of Gdańsk

Truong Thi Khanh Ha, National Vietnam University

Melanie Vauclair;

Anna Włodarczyk

Abstract: Current report presents examination of measurement invariance for Personal Well-being Index, 8-item version. Basing on data obtained in 26 countries from students samples (n = 5,731) the single Confirmatory Factor Analyses (CFA) and multi-group CFA (MGCFA) was performed to assess the goodness of fit for unidimensional solution in each one country separately and to assess the level of measurement invariance. Analyses confirmed good or moderate model fit in each country. MGCFA indicated partial metric invariance, suggesting usefulness of PWI8 in cross-cultural studies on correlates and predictors of life satisfaction measured by PWI8.

Key words: Personal Well-being Index; cross-cultural studies, measurement invariance

Measurement Invariance of Personal Well-being Index in 26 Countries

Quality of life has become an important, well researched topic over the last few years. More specifically in terms of well-being, which is often assessed at national levels for international comparisons. These international comparisons, however require measures which have been shown to be invariant across different cultural groups and countries. The objective of this study is to examine the measurement invariance of just such a measure, the Personal Well-being Index (PWI) which is considered one of the most popular measures for evaluating subjective well-being (International Well-being Group, 2013; Sirgy, 2012).

Life satisfaction is defined as a cognitive representation of subjective well-being and has a general character (Diener, 1984). According to Diener, Horowitz and Emmons (1985) life satisfaction a judgmental process conditioned by "a comparison of one's circumstances with what is thought to be appropriate standard" (.p.71). Therefore it refers to some standards of evaluation, which could be related to different life domains. The PWI evaluates cognitive aspects of subjective well-being by measuring life satisfaction in different domains (Cummins, Eckersley, Pallant, Van Vugt, & Misajon, 2003), namely: (1) standard of living, (2) personal health, (3) life achievements, (4) personal relationships, (5) personal safety, (6) community connectedness, (7) future security, and (8) religion and spirituality. The last item was added recently (International Well-being Group, 2013). The PWI has been used as an assessment of life satisfaction in child (Casas, Bello, González, & Aligué, 2012), adolescent and student (Tomyn, Norrish, & Cummins, 2011), aging (Bricker-Katz, Lincoln, & McCabe**,** 2009; Forjaz et al., 2011), and clinical populations (e.g. Engel & Cummins, 2011; Werner, 2012). The scale is intended to be inclusive of all important life domains which could contribute to the general level of life satisfaction and to serve as a tool in cross-cultural comparisons on relative importance of particular domain in life satisfaction (International Wellbeing Group, 2013). The idea behind developing the PWI was to include the most

important predictors of general life satisfaction. A selection of domains were done by international team and was based on several criteria: to include only basic domains importnant for predicting "life satisfaction as a whole", each domain refers to broad aspect of life, and each domain need to represent indicator not causal variable of general life satisfaction (see International Wellbeing Group, 2013).

Despite the many cultural adaptations of this measure (see International Wellbeing Group, 2013 for details) and its increasing popularity among cross-cultural researchers, to the best of our knowledge there is little evidence that the PWI is invariant across different countries, with the exceptions of adolescent samples in Chile and Brasil (Sarriera et al., 2014) or general populations in Hong Kong and Australia (Lau, Cummins, & McPherson, 2005). Sometimes the levels of PWI are compared without examination of measurement invariance, like in case of Romania and Hungary (Baltatescu, 2014). This report intends to fill this gap by examining the measurement invariance of the PWI in a university student samples across 26 countries.

Measurement Invariance as Means of Cross-Cultural Inquiry

Oishi (2010) pointed towards several important methodological and conceptual issues related to cross-cultural studies on subjective well-being, these are: conceptual equivalence, translation issues, desirability of the concept, response style, item functioning, differences in self-presentations, memory bias, and validity criteria. For instance, single-item measures of subjective well-being like Cantrill's ladder or general item on life satisfaction are less reliable than longer scales and do not allow for more in-depth examination of cross-cultural equivalence in terms of measurement equivalence (see Oishi, 2010). As the PWI is the multi-item scale it is particularly useful in cross-cultural research. Using simple and easy to translate items based on the work from an international group of well-being researchers, which allows for minimization of problems with conceptual equivalence and translations (see International Wellbeing Group, 2006, 2013).

Multi-group Confirmatory Factor Analysis (MGCFA) serves as a statistical tool for assessment of cross-cultural equivalence of a measure. Such analysis is fundamental for establishing the usefulness of any measure intended for cross-cultural research. There are three levels of measurement invariance which are most commonly used to establish whether a measure is equivalent: a) Configural invariance provides indication that the general factor structure of the measure is the same across different groups. At this level, the construct is measured similarly in different samples. b) Metric invariance indicates that the factor loading of items are similar, (i.e., load in the same way in assumed factor) across groups. At this level, measure correlates and/or predictors may be compared across samples. c) Scalar invariance indicates that item intercepts are equal intercepts across groups. At this level means may be compared across samples (Davidov, Meuleman, Cieciuch, Schmidt, & Billiet, 2014). Scalar invariance is rarely found in large cross-cultural comparisons (see Davidov et al., 2014). In general there are very few studies examining measurement invariance of scales examining subjective well-being. For instance, only metric invariance was reported for Satisfaction with Life Scale (SWLS) across the US, England and Japan (Wishman, & Judd, 2016), and mixed results for SWLS invariance across Russian and US samples (Tucker, Ozer, Lyubomirski, & Boehm, 2006). Most studies typically focus on comparisons between several national groups, rarely examining large representation of countries (see also Ponizovsky, Dmitrova, Sachner, & Van de Schoot, 2013).

The current study

Despite the many cultural adaptations of the PWI (see International Wellbeing Group, 2013 for details) and its increasing popularity among cross-cultural researchers, to the best of our knowledge there is little evidence that the PWI is invariant across different countries, with the exceptions of adolescent samples in Chile and Brasil (Sarriera et al., 2014) or general populations in Hong Kong and Australia (Lau, Cummins, & McPherson, 2005). Sometimes

the levels of PWI are compared without examination of measurement invariance, like in case of Romania and Hungary (Baltatescu, 2014). This report intends to fill this gap by examining the measurement invariance of the PWI in a university student samples across 26 countries.

In the current study we examine the measurement invariance of PWI across countries from different regions of the world: Europe (10), Asia (10), Africa (2); and Latin America (4). Among them there are the most affluent and developed countries like UK or Japan and less affluent, agrarian societies like Iran or Kenya. In terms of cultural regions we had representatives for all Huntington (1996) cultural groups, in terms of religion we had countries representing all main world religions. This selection of countries are not exhaustive, but it allows for examining measurement invariance of PWI across different languages and cultures. The aim of the study was to investigate the measurement invariance of the PWI across different countries and languages. Given the large number of countries compared, we expected to find support for the metric level of invariance.

Method

Sample and Procedure

Data were collected in a paper-pencil or online format between April 2014 and August 2015. The sample included 5,530 university students (42.4 % men, $M = 21.29$, $SD = 3.15$, age ranged from 16 to 39). We excluded all participants above the age of 40 (1.7% of total sample) from the analyses, as in most countries the respondents' age was in the 18-25 range, and rarely exceeded 30 years. We also asked students to indicate the socioeconomic status of their families on a 7-point Likert-type scale (from 1 = *significantly below average* to 7 = *significantly over average*). The students majored in different fields (e.g., social sciences, technical sciences and medical sciences) and originated from 26 countries (see Table 1 for a sample breakdown). They were recruited to the study during their classes and participated on

a voluntary basis. They completed PWI as part of a broader project on entitlement and subjective wellbeing. The paper-and-pencil surveys were administered in small groups. In countries where the survey was administered in a non-native language, a researcher assisted students and explained the meaning of particular words.

--- Table 1 about here ---

Measure

The Personal Well-being Index (PWI, Cummins et al., 2003; International Well-being Group, 2013) measures satisfaction with different life domains: (1) standard of living, (2) health, (3) life achievements, (4) personal relationships, (5) personal safety, (6) community connectedness, (7) future security, and (8) religion and spirituality. Previous studies suggest that in different countries relative importance of religiosity and spirituality vary as a function of cultural differences (Norris & Inglehart, 2005) and that they both significantly contribute to subjective well-being (Piedmont, & Friedman, 2012). Therefore, we used one combined question about religiosity and spirituality (*How much are you satisfied with your spirituality or religion?*) as suggested by manual for PWI-8 despite some researchers postulated two parallel version for item 8 (Sarriera et al., 2014). Participants responded on an 11 point Likert-type scale (0 = *not at all satisfied* to 10 = *totally satisfied*). National versions of scale were authorised versions or they were obtained by repeating back translations procedure with bilingual researchers and with the participation of Robert Cummins (see Table 1 for information on language of administration).

Statistical Analyses

We started by using Mplus 7.4 to perform Confirmatory Factor Analysis (CFA) in order to test for a unidimensional structure of the PWI-8 in each country sample. Because the score distributions were not perfectly normal and Mardia's multivariate skewness and kurtosis statistics were significant in all samples, we used the robust Satorra-Bentler $\chi^{2.}$ (Satorra, &

Bentler, 1994; referred as estimator MLM in Mplus). Because the country samples differed in gender distribution, we used weighting in all analyses to equalize the contribution of male and female respondents within each country to the model.

The model fit was examined using the most common fit indices: the Chi-square (χ^2), the CFI (Comparative Fit Index), the RMSEA (Root Mean Squared Error of Approximation), and the SRMR (Standardized Root Mean Squared Residual). In larger samples (N > 200), practical fit indices (CFI, RMSEA, and SRMR) are preferred to the χ^2 as they are less sensitive to sample size (Chen, 2007; Davidov et al., 2014). CFI values above .90 were considered as evidence of an acceptable model fit and those above .95 as evidence of a good fit. Because in smaller samples RMSEA tends to over-reject correct models (Hu & Bentler, 1999), we used RMSEA values of .10 and .06 as thresholds for acceptable and good fit, respectively. For SRMR, .08 and .05 were used (Brown, 2015). In samples where the fit of the one-factor theoretical model was outside the acceptable range and a pronounced and interpretable outlier was found among the modification indices suggesting an error covariance, the latter was added and the model was retested. We aimed to introduce as few modifications as possible in order to achieve acceptable fit without over-complicating the model.

After establishing the measurement model for each country, we proceeded by conducting multi-group CFA (MGCFA) to test for configural, metric, and scalar measurement invariance. We tested the invariance based on modified measurement models using a conventional approach (Byrne, 2012). We used the ΔCFI and ΔRMSEA values of .010 and .015, respectively, as evidence of pronounced difference between nested models (Chen, 2007; Cheung & Rensvold, 2002). We looked for outliers among the modification indices and introduced them into the model one-by-one, until the difference in practical fit indices

between the configural invariance and partial metric invariance models became small enough (ΔCFI $\leq$ .01, ΔRMSEA $\leq$ .015). The procedure was repeated for scalar invariance.

A potential drawback of the manual approach is that each modification results in a different model, and the exact resulting list of non-invariant parameters is dependent on the sequence in which modifications are entered into the model. In case of a long sequence of modifications, the conventional approach (addressing the strongest modification index at each step) does not guarantee that the resulting model will be optimal. This problem is overcome by the alignment procedure (Asparouhov & Muthén, 2014), which evaluates the fit of different combinations of non-equivalent parameters. We tried to cross-validate our findings using the alignment procedure, based on the same modified measurement model. Finally, we tested the invariance of the PWI across genders. Because the sample sizes were not large enough to test the invariance across genders in each country separately, we tested a single-factor model in the combined sample with robust chi-square (MLR) and standard errors computed using the sandwich estimator for clustered samples to account for non-independence of observations within countries.

Results

The internal consistency (Cronbach's alpha) values of the PWI in each national sample are presented in Table 1. Cronbach's alpha values above .70 indicating good reliability (Lance, Butts, & Michels, 2006) were found in all samples.

Table 2 presents the results of single-sample CFA analyses for the initial (theoretical) model. In most countries the theoretical model showed acceptable fit, based on the combination of practical fit indices. The fit of the model was outside the acceptable range in Spain, Poland, South Korea, Hungary, Romania, Indonesia, and Panama.

--- Table 2 about here ---

We explored the modification indices in countries with unacceptable and marginal fit and introduced additional covariances in cases where they were theoretically justified. The error covariance for items 4 and 5 (relationships and safety) was found in three Hispanic countries, in line with previous studies (Sarriera et al., 2014). The error covariance of items 4 and 6 (relationships and feeling part of community) was peculiar to two post-Communist Central European countries (Poland, Hungary). The other error covariances were explained by back-translation analysis. For instance, the error covariance of items 5 and 7 was found in countries (Poland, Brasil) where local translations used the same word for "safety" and "security".

South Korea was the only country where modification indices revealed a pattern suggesting higher dimensionality. Using exploratory factor analysis (robust statistics with oblique Geomin rotation), we found that items 1, 3, and 7 tended to form a separate dimension reflecting satisfaction with financial success. However, in 2- and 3-factor models the dimensions were highly correlated and showed numerous cross-loadings, suggesting that a single-factor model is optimal. We added two error covariances to address the subdimension. The introduction of additional error covariances resulted in acceptable fit in all countries (shown in Table 3).

Comment [EO1]: If it looks too cumbersome, we can just say something about translation artifacts of cultural features.

No, it is very interesting and I would keep this. In fact, I`m surprised that we did not find something similar in Vietnam

--- Table 3 about here ---

We proceeded by conducting invariance analyses. Multi-group model included modified measurement models for 10 countries and theoretical model for the remaining 16 countries. The configural model showed good fit to the data, the fit of the metric model was acceptable, and the fit of the scalar model was poor (see Table 4). The chi-square differences were also significant ($p < .001$) between the three models. The difference in practical fit indices between the configural and metric model was very small for RMSEA (ΔRMSEA = .003), but above the recommended .01 threshold for the CFI (ΔCFI = .018) and we followed by establishing partial metric invariance. After 6 constraints for non-invariant loadings (listed in

supplementary material) were relaxed, the difference in practical fit indices between the configural invariance and partial metric invariance models became small enough (ΔCFI = .01). We followed by establishing the partial scalar invariance. After relaxing 74 constraints for equal intercepts, the corresponding modification indices became non-significant at $p <$. 01 level and we stopped the procedure to reduce the risk of false positives. Even though the ΔCFI criterion was not reached (ΔCFI = .024), the ΔRMSEA and ΔSRMR were quite small (< .010), and practical fit indices (Table 4) were within acceptable limits. The complete list of non-invariant parameters obtained from the final partial scalar invariance model is given in Supplementary Information (table SI.1). The parameters of the resulting model are also given in Supplementary Information (tables SI.4 and SI.5).

Comment [MŻ2]: Here you mean SI.1 and SI.2?

--- Table 4 about here ---

The number of non-equivalent intercepts ranged from 6 to 12 per PWI item. Some of the intercepts revealed meaningful patterns. For instance, non-equivalence of the intercept of item 3 (“achieving in life”) was more often found in Asian, collectivistic cultures. Non-equivalence of the intercept of item 6 (“feeling part of your community”) was typically found in Latin American countries, but not in post-Communist ones.

--- Table 5 about here---

The estimates of latent factor means and variances obtained from the model are presented in Table 5. The latent factor means were highly correlated with the observed means (r = .95).

The results of the alignment procedure are presented in Supplementary Information (table SI.2). The alignment procedure has identified a smaller number of non-equivalent parameters, 1 loading and 37 intercepts, suggesting that the manual approach might be more conservative. Thirty-three of these intercepts were also identified as non-equivalent using conventional approach. The latent factor means estimated using the alignment procedure were

highly correlated with the observed means (r = .98) and those obtained using conventional approach (r = .96). These data indicate a fairly good convergence of the findings from the two procedures.

Finally, to test the contribution of gender to non-invariance, we performed a multi-group CFA for females and males in the combined sample using country as a cluster variable in order to account for non-independence of observations within each country. The resulting fit indices and model comparison results are shown in Table 5. All three models showed good fit to the data. The difference between the nested models in terms of practical fit indices was below the thresholds suggested by Chen (2007), indicating that gender does not have any uniform effects on measurement invariance across countries.

Discussion

Our objective was to establish measurement invariance of the PWI-8 across 26 countries. We found that the PWI was unidimensional, with Cronbach's alphas indicating acceptable internal consistency in all countries. MGCFA confirmed that the basic construct structure of the PWI-8 is similar across groups. Although additional covariances between items can improve the fit in some countries, in most countries the fit of the theoretical model was very close to the acceptable range. The same was true for configural and metric-invariant MGCFA models. Although the difference between these two nested models was significant in terms of chi-square difference, the difference in practical fit indices was not strong, suggesting that the comparison of effects (e.g., correlations) obtained using the PWI in different languages across countries would not be strongly biased by non-equivalent loadings. One should only be cautious while performing the comparison of raw scores, as partial scalar invariance model indicated good model fit but the differences in CFI between the partial metric and partial scalar model was higher than recommended value of .01. These findings suggest that it is

possible to examine the predictors and correlates of the subjective well-being phenomenon using the PWI-8 across countries.

The results of analysis using alignment approach based on the theoretical model suggest that despite the presence of some weak non-invariance of loadings and pronounced non-invariance of item intercepts in some countries, the observed mean scores are very similar to unbiased latent factor scores. Although removal of bias can slightly change the rank ordering of countries, these effects were only pronounced for a few countries (Kenya, Japan, Israel, South Africa).

We developed modified measurement models by introducing theoretically interpretable and strongly significant ($p < .001$) modification indices for some countries. However, in most cases these modification indices accounted for translation artifacts, which can be removed by improving translations of the instrument into certain languages.

The attempts to develop partial scalar invariance models using the older, manual approach and alignment approach result in different, although largely overlapping, sets of non-invariant parameters. Items referring to more objective realities, such as one's living standard and health turned out to be more invariant, compared to the items referring to more subjective phenomena, such as one's spirituality or future security.

Finally, gender does not seem to contribute to non-invariance of loadings and intercepts in any uniform manner across countries. However, because our samples did not allow to evaluate gender invariance in each country separately, this analysis does not rule out the possibility of country-specific non-invariance associated with gender.

Limitations and recommendation for future studies

The current report has several limitations: the use of student samples, the lack of several important cultures and countries (such as Chinese or American) and the overrepresentation of European countries. For practical reasons we used both online and paper-pencil surveys and

in some countries the questionnaire was distributed in English, rather than in native language, which could lead to increased measurement error. Also, the student samples were not representative of their respective countries, which precludes us from interpreting the substantial differences in the mean score estimates. Finally, although we decided for using a combined item for measuring satisfaction with religion and spirituality, these constructs are not interchangeable (Piedmont & Friedman, 2012). As this solution occurred to work good both in CFA and MGCFA it could be used in cross-cultural comparisons, however for further exploration of importance of religion and spirituality as separate factors in shaping overall life satisfaction two separate items should be used (see Sarriera et al., 2014).

In terms of specific recommendations, PWI researchers could use this tool in all countries as indicator of one general life satisfaction, measured by particular items. However, we have found some consistently repeating cultural differences, which could be further explored. For instance, in South Korean sample we have found evidence for existing additional factor representing concern about financial success. It could suggest that in this population life satisfaction is somewhat affected by materialism. We have found also several differences in intercepts for PWI items. It means that particular populations vary in their style of responding on particular items, what could suggest that the meaning of item is different in different cultures. For instance, satisfaction for achievement in life indicates specificity in collectivistic countries. As collectivistic countries are typically "face-saving" cultures (Bond, 1991) life achievement could means that individual just fits to social environment, contrary to individualistic countries, where life achievements would mean developing unique characteristics. As PWI statements are typically very general (as they are aimed to represent broad life domains, see International Well-being Group, 2013), the cultural meaning of these broad statements could be affected by cultural context. Our study provides some suggestions, where it is the case (e.g. life achievement, feeling part of community).

In conclusion, the current report provided information about the possibility of cross-cultural research among university students based on PWI-8 scores, providing the evidence of metric invariance and partial scalar invariance, allowing the cross-country comparison of effects, but not of group or individual raw scores. We also compared the results of different approaches to establishing unbiased factor means across countries. This provides valuable information on the further development of subjective well-being research in different cultural contexts. As the main goal of the International Well-being group is to explore the importance of satisfaction with particular domains in shaping overall life satisfaction our findings indicate that this research goal could be realised successfully in cross-cultural research.

References

Asparouhov, T. & Muthén, B. (2014). Multiple-Group Factor Analysis Alignment. *Structural Equation Modeling: A Multidisciplinary Journal, 21*, 1-14. DOI: 10.1080/10705511.2014.919210

Bricker-Katz, G., Lincoln, M., & McCabe, P. (2009). A Life-time of Stuttering: How emotional reactions to stuttering impact activities and participation in older people. *Disability and Rehabilitation, 31*, 1742–1752.

Baltatescu, S. (2014). Romanian-Hungarian cross-border region, Personal Well-being Index. In: Alex C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 5598-5605). Netherlands: Springer.

Bond, M. H. (1991). *Beyond the Chinese face: Insights from psychology*. Oxford University Press.

Brown, T. A. (2015). Confirmatory Factor Analysis for Applied Research. 2nd Ed. N. Y.: The Guilford Press.

Byrne, B. N. (2012). *Structural Equation Modelling with Mplus*. N.Y.: Routledge.

Casas, F., Bello, A., González, M., & Aligué, M. (2012). Personal Well-Being among Spanish Adolescents. *Journal of Social Research & Policy, 3*, 19-45.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.

Cummins, R. A., Eckersley, R., Pallant, J., Van Vugt, J., & Misajon, R. (2003). Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. *Social Indicators Research, 64*, 159–190.

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55-75. DOI: 10.1146/annurev-soc-071913-043137

Engel, L., & Cummins, R. A. (2011). Impact of dose adjustment for normal eating in Australia (OzDAFNE) on subjective wellbeing, coping resources and negative affects in adults with type 1 diabetes: A prospective comparison study. *Diabetes Research and Clinical Practice, 91*, 271–279.

Forjaz, F. M., Prieto-Flores, M. E., Ayala, A., Rodriguez-Blazquez, C., Fernandez-Mayoralas, G., Rojo-Perez, F., & Martinez-Martin, P. (2011). Measurement properties of the Community Wellbeing Index in older adults. *Quality of Life Research, 20*, 733-743. DOI: 10.1007/s11136-010-9794-2.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*, 1-55.

Huntington, S. (1996). *The clash of civilizations and remaking of world order*. New York: Touchstone.

International Wellbeing Group (2006). *Personal Wellbeing Index*. Melbourne, Australia: Australian Centre on Quality of Life, Deakin University. Retrieved from http://www.deakin. edu.au/research/acqol/instruments/wellbeing_index.htm

International Wellbeing Group (2013). *Personal Wellbeing Index: 5th edition*. Melbourne, Australia: Australian Centre on Quality of Life, Deakin University. Retrieved from http:// www.deakin.edu.au/research/acqol/instruments/wellbeingindex/index.php

Lance, C.E., Butts, M.M., & Michels, L.C. (2006). The sources of four commonly reported cut-off criteria: What did they really say? *Organizational Research Methods, 9*, 202-220. DOI: 10.1177/1094428105284919

Lau, A. L. D., Cummins, R. A., & McPherson, W. (2005). An Investigation into the Cross-Cultural Equivalence of the Personal Wellbeing Index. *Social Indicators Research, 72*, 403–430.

Oishi, S. (2010). Culture and well-being: Conceptual and methodological issues. In: E. Diener, J. F. Helliwell, & D. Kahneman (Eds.), *International differences in well-being* (pp. 34-69). Oxford University Press.

Piedmont, R.L., & Friedman, P. (2012). Spirituality, religiosity and subjective well-being. In: K. C. Land, A. C. Michalos, & M. J. Sirgy, (Eds.), *Handbook of social indicators and quality of life research* (pp.313-329). Netherlands: Springer. DOI: 10.1007/978-94-007-2421-1

Ponizovsky, Y., Dimitrova, R., Schachner, M. K., & Van de Schoot, R. (2013). The satisfaction with life scale: measurement invariance across immigrant groups. European Journal of Developmental Psychology, 10(4), 526-532. DOI: 10.1080/17405629.2012.707778

Sarriera, J. C., Casas, F., Alfaro, J., Bedin, L., Strelhow, M. R., Abs, D., Valdenegro, B., Garcia, C., & Oyarzun, D. (2014). Psychometric properties of the Personal Wellbeing Index in Brazilian and Chilean adolescents including spirituality and religion. *Psychology/ Psicologia Refl exão e Crítica, 27*, 710-719. DOI: 10.1590/1678-7153.201427411

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In: A. von Eye, C. C. Clogg (Eds.), *Latent Variables Analysis: Applications for developmental research* (pp 399–419). Thousand Oaks, CA: Sage.

Sirgy, M.J. (2012). *The psychology of quality of life. Hedonic well-being, life satisfaction, and eudaimonia*. New York, London: Springer. DOI: 10.1007/978-94-007-4405-9.

Tomyn, A. J., Norrish, J. M., & Cummins, R. A. (2011). The subjective wellbeing of indigenous Australian adolescents: Validating the personal Wellbeing Index – School Children. *Social Indicators Research*, *103,* 405-411. DOI 10.1007/s11205-011-9970-y

Werner, S. (2012). Subjective well-being, hope, and needs of individuals with serious mental illness, *Psychiatry Research, 196*, 214-219. doi:10.1016/j.psychres.2011.10.012

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*, 10–18.

Table 1

Sample information and internal consistency coefficients for PWI-8 in 26 countries

Country	*N*	Female %	Age *M*	SES	Language	Procedure	Cronbach's α
Armenia	223	48	19.00	4.98	Armenian	Paper-pencil	.84
Brasil	225	64	20.50	4.38	Portuguese	Online	.82
Bulgaria	197	66	23.70	4.66	Bulgarian	Paper-pencil	.82
Chile	241	52	22.00	4.34	Spanish	Paper-pencil	.82
Estonia	289	68	22.22	4.37	Estonian	Online	.79
Hungary	206	69	21.01	n/a	Hungarian	Paper-pencil	.81
India	200	69	22.59	4.37	English	Paper-pencil	.76
Indonesia	200	50	21.38	4.70	Indonesian	Online	.89
Iran	201	50	21.28	4.46	English	Paper-pencil	.75
Israel	200	40	24.58	4.88	Hebrew	Online	.82
Japan	202	23	18.91	4.14	Japanese	Paper-pencil	.82
Kenya	161	53	23.39	4.07	English	Paper-pencil	.90
Nepal	199	51	22.70	4.08	English	Paper-pencil	.74
Panama	176	32	22.03	4.13	Spanish	Online	.88
Poland	258	60	21.85	4.69	Polish	Paper-pencil	.86
Portugal	187	77	22.79	4.11	Portuguese	Online	.82
Puerto Rico	300	43	20.26	4.16	Spanish	Paper-pencil	.84
Romania	210	48	21.49	4.72	Romanian	Paper-pencil	.85
Russia	227	83	21.03	3.11	Russian	Online	.87
Serbia	199	61	22.46	3.77	Serbian	Paper-pencil	.84
Slovakia	202	72	21.13	4.76	Slovakian	Paper-pencil	.77

S. Africa	188	67	20.17	4.45	English	Paper-pencil	.79
S. Korea	215	55	22.20	3.90	Korean	Paper-pencil	.83
Spain	196	51	21.20	4.01	Spanish	Online	.72
UK	302	81	19.44	4.21	English	Online	.84
Vietnam	259	53	20.52	4.25	Vietnamese	Paper-pencil	.86

Note. SES = Subjective economic status of family (1-7). Due to technical error data for socio-economic status of family are not available for Hungary.

Table 2

Goodness-of-fit indices for PWI8 one-factor model (no error covariances)

Country	S-B χ^2	SCF	CFI	RMSEA (90% CI)	SRMR
Armenia	59.03***	1.67	.901	.095 (.068-.124)	.056
Brasil	51.72***	1.35	.912	.089 (.059-.119)	.052
Bulgaria	47.66**	1.24	.924	.084 (.053-.115)	.050
Chile	34.81*	1.39	.966	.056 (.022-.086)	.036
Estonia	55.76***	1.35	.902	.079 (.055-.104)	.054
Hungary	53.97***	1.47	.883	.091 (.062-.121)	.061
India	23.17	1.65	.981	.028 (.000-.070)	.044
Indonesia	60.52***	1.35	.931	.101 (.072-.130)	.047
Iran	33.88*	1.15	.953	.059 (.020-.092)	.045
Israel	37.49*	1.57	.931	.067 (.032-.100)	.050
Japan	26.22	1.73	.975	.042 (.000-.082)	.046
Kenya	28.99	3.14	.962	.055 (.000-.096)	.053
Nepal	35.05*	1.47	.916	.064 (.025-.098)	.055
Panama	55.70***	1.51	.932	.101 (.070-.133)	.045
Poland	65.24***	1.40	.904	.103 (.076-.131)	.053
Portugal	32.57*	1.62	.949	.058 (.014-.094)	.050
Puerto Rico	71.21***	1.69	.908	.092 (.070-.116)	.051
Romania	48.18***	2.17	.898	.084 (.054-.115)	.060
Russia	40.46**	1.85	.935	.077 (.042-.111)	.047
Serbia	47.56**	1.38	.927	.084 (.053-.115)	.051
Slovakia	33.80*	1.49	.939	.059 (.020-.092)	.051
South Africa	37.82**	1.57	.917	.069 (.034-.103)	.056

South Korea	99.00***	1.35	.844	.137 (.111-.164)	.064
Spain	86.24***	1.30	.756	.131 (.104-.160)	.076
UK	33.69*	1.71	.973	.048 (.016-.075)	.033
Vietnam	44.88**	1.37	.956	.070 (.042-.097)	.041

Note. S-B χ^2 = Satorra-Bentler chi-square (df = 20), *** $p < .001$, ** $p < .01$, * $p < .05$; SCF = Scaling Correction Factor for Satorra-Bentler χ^2, CFI = Comparative Fit Index, RMSEA = Root Mean Square of Approximation, SRMR = Standardized Root Mean Square Residual.

Table 3.

Goodness-of-fit Indices for PWI8 One-factor Solution, modified measurement models

Country	S-B χ^2	*df*	SCF	CFI	RMSEA (90% CI)	SRMR	Error Covariances
Armenia	25.81	19	1.62	.983	.041 (.000-.077)	.041	6-8
Brasil	36.01*	19	1.27	.953	.067 (.032-.099)	.043	5-7
Hungary	38.73**	19	1.45	.932	.071 (.038-.103)	.055	4-6
Indonesia	40.52**	19	1.36	.963	.075 (.043-.107)	.041	3-7
Panama	46.32***	19	1.55	.948	.090 (.058-.124)	.043	4-5
Poland	33.86*	18	1.36	.966	.064 (.029-.097)	.039	5-7, 4-6
Puerto Rico	51.12***	19	1.69	.942	.075 (.051-.100)	.043	4-5
Romania	35.51*	19	2.18	.940	.066 (.030-.099)	.047	1-2
South Korea	54.32***	18	1.35	.928	.098 (.069-.128)	.048	1-3, 3-7
Spain	39.24**	19	1.33	.925	.074 (.041-.108)	.052	4-5

Note. S-B χ^2 = Satorra-Bentler robust chi-square (df = 20), *** $p < .001$, ** $p < .01$, * $p < .05$; SCF = Scaling Correction Factor for Satorra-Bentler χ^2, CFI = Comparative Fit Index, RMSEA = Root Mean Square of Approximation, SRMR = Standardized Root Mean Residual.

Table 4

Fit indices for Multi-group Confirmatory Factor Analysis (MGCFA) models of the PWI8 across 26 countries

Model	χ^2	*df*	SCF	CFI	RMSEA (90 % CI)	SRMR
Configural	975.95	508	1.57	.948	.066 (.060-.073)	.047
Metric	1316.10	683	1.53	.930	.067 (.061-.072)	.084
Partial Metric	1237.03	677	1.52	.938	.063 (.057-.068)	.075
Scalar	3455.53	852	1.22	.711	.121 (.111-.125)	.120
Partial Scalar	1554.46	777	1.34	.914	.069 (.064-.074)	.079
Difference tests	$\Delta\chi^2$	Δdf		ΔCFI	ΔRMSEA	ΔSRMR
Metric vs. Configural	340.47	175		.018	.001	.037
Partial metric vs. Configural	254.10	169		.010	.003	.028
Scalar vs. Partial metric	39298.62	175		.227	.058	.045
Partial scalar vs. Partial metric	1669.61	100		.024	.006	.004

Note. Satorra-Bentler χ^2, all $p < .001$. SCF = Scaling Correction Factor, CFI = Comparative Fit Index, RMSEA = Root Mean Square of Approximation, SRMR = Standardized Root Mean Square Residual. The list of error covariances included in the model is given in Table 2. The list of non-invariant loadings and intercepts is given in Table 4.

Table 5

Comparison of observed country mean scores and latent factor estimates

	Observed Score			MGCFA Estimates		
	M	*SD*	Rank	*M*	Variance	Rank
Romania	8.07	1.33	1	0.07	0.92	1
India	7.67	1.36	2	-0.27	1.23	2
Chile	7.25	1.37	8	-0.38	1.04	3
Slovakia	7.43	1.16	5	-0.40	0.75	4
Panama	7.31	1.69	6	-0.46	2.05	5
Spain	7.15	1.20	10	-0.47	0.64	6
Hungary	7.58	1.29	3	-0.50	1.10	7
Puerto Rico	7.29	1.62	7	-0.54	1.55	8
Armenia	7.52	1.55	4	-0.55	1.44	9
Brasil	7.13	1.46	11	-0.59	1.25	10
Israel	7.15	1.31	9	-0.67	1.00	11
Portugal	6.89	1.27	16	-0.87	0.93	12
Bulgaria	7.02	1.54	13	-0.92	1.38	13
South Africa	6.99	1.40	14	-1.01	1.09	14
Vietnam	6.78	1.42	19	-1.12	1.17	15
Estonia	7.10	1.21	12	-1.13	0.82	16
Poland	6.91	1.54	15	-1.14	1.27	17
Serbia	6.80	1.65	18	-1.14	1.63	18
Indonesia	6.45	1.60	20	-1.25	1.56	19
UK	6.85	1.38	17	-1.29	1.20	20

Russia	6.40	1.64	22	-1.30	1.74	21
South Korea	6.15	1.34	23	-1.61	1.07	22
Nepal	6.44	1.41	21	-1.62	1.08	23
Japan	5.38	1.45	26	-2.06	1.41	24
Kenya	5.86	2.15	24	-2.26	3.03	25
Iran	5.72	1.61	25	-2.33	1.28	26

Table 6

Fit indices for Multi-group Confirmatory Factor Analysis (MGCFA) models of the PWI8 across gender

Model	S-B χ^2	*df*	CFI	RMSEA (90 % CI)	SRMR
Configural	218.12*	40	.968	.040 (.035-.046)	.028
Metric	231.90*	47	.967	.038 (.033-.043)	.031
Scalar	265.05*	54	.962	.038 (.033-.043)	.035
Difference tests	$\Delta\chi^2$	Δdf	ΔCFI	ΔRMSEA	ΔSRMR
Metric vs. configural	6.03	7	.001	.002	.003
Scalar vs. metric	32.84*	7	.005	<.001	.004

Note. Satorra-Bentler χ^2, * $p < .001$. CFI = Comparative Fit Index, RMSEA = Root Mean Square of Approximation, SRMR = Standardized Root Mean Square Residual.

Supplementary Information

Table SI.1

Results of Partial Scalar Invariance Analyses: List of Non-Invariant Parameters

Country	Loadings, item # ($\Delta\chi^2$)	Intercepts, item # ($\Delta\chi^2$)
Armenia		3 (34.49), 4 (8.27)
Brazil		5 (41.46), 6 (26.79), 7 (17.93), 8 (8.22)
Bulgaria		
Chile		5 (43.19), 6 (31.96), 2 (28.05), 8 (19.32)
Estonia		8 (63.17), 5 (46.16), 4 (11.32)
Hungary	3 (12.65)	7 (15.63), 5 (8.19)
India	5 (9.12)	3 (18.74), 8 (12.83), 2 (7.66), 6 (7.87)
Indonesia		3 (23.74), 4 (8.51)
Iran	1 (11.38)	2 (26.60), 4 (17.17), 3 (10.03), 1 (12.13)
Israel		6 (57.76)
Japan	8 (15.21)	8 (24.12), 3 (20.26), 7 (10.97)
Kenya		7 (19.93), 8 (22.09), 3 (7.69)
Nepal		4 (21.65), 8 (15.71)
Panama	2 (9.74)	6 (37.78), 5 (13.82), 4 (12.74)
Poland		5 (7.73), 1 (8.60), 7 (8.18)
Portugal		7 (16.55), 6 (6.77)
Puerto-Rico		6 (30.35), 5 (25.85)
Romania	4 (19.12)	1 (14.18), 3 (7.35), 4 (7.65)
Russia		3 (23.10), 2 (18.45)

Serbia	1 (73.40), 4 (28.04), 2 (21.46)
Slovakia	7 (17.99), 3 (11.91), 8 (11.34), 4 (8.16), 2 (6.72)
S. Africa	6 (26.48), 8 (10.92), 7 (6.65)
S. Korea	1 (23.86), 3 (17.13), 7 (17.92), 5 (6.67)
Spain	4 (39.13), 6 (33.17), 7 (31.52), 8 (26.49)
UK	6 (35.02), 5 (18.32), 1 (19.45), 7 (7.32)
Vietnam	7 (43.19), 3 (18.45)

Table SI.2

Results of Alignment Analyses: List of Non-Invariant Parameters and Mean Estimates

	Non-invariant parameters		Latent factor estimates	
	Loadings	Intercepts	Mean	Variance
Armenia		3, 4	-0.15	1.25
Brazil		5, 8	-0.35	1.09
Bulgaria			-0.43	1.17
Chile		5, 7*	-0.33	0.94
Estonia		1*, 5, 8	-0.45	0.69
Hungary		5	-0.06	0.94
India		3, 8	0.00	1.00
Indonesia		3, 6*	-0.95	1.42
Iran			-1.32	1.20
Israel		6	-0.26	0.84
Japan		3, 8	-1.47	1.26
Kenya	2	7, 8	-1.52	2.58
Nepal		8	-0.95	0.89
Panama		6	-0.27	1.78
Poland		5	-0.51	1.06
Portugal		7	-0.53	0.92
Puerto Rico		5	-0.28	1.36
Romania		1	0.33	0.88
Russia			-1.01	1.52
Serbia		1, 4	-0.51	1.43

Slovakia	6*, 8	-0.22	0.68
S. Africa	7, 8	-0.57	0.94
S. Korea	3, 7	-1.03	0.90
Spain	4	-0.31	0.61
UK	5	-0.69	1.04
Vietnam	3, 7	-0.67	1.03

Note: non-invariant intercepts that were not discovered using the manual approach are marked with an asterisk.

Table SI.3 Results of Mardia's test for multivariate normality

Country	Multivariate skewness			Multivariate kurtosis		
	Sample statistic	M	SD	Sample statistic	M	SD
Armenia	18.70	3.28	0.47	123.06	79.32	1.60
Brasil	14.94	3.51	0.48	105.79	79.26	1.66
Bulgaria	13.07	3.62	0.50	100.71	79.28	1.73
Chile	9.06	3.00	0.40	102.50	79.44	1.57
Estonia	9.18	2.44	0.34	99.03	79.49	1.38
Hungary	19.10	3.43	0.49	116.22	79.25	1.62
India	25.06	3.48	0.49	132.86	79.24	1.71
Indonesia	11.05	3.51	0.48	108.80	79.25	1.66
Iran	5.93	3.49	0.49	83.04	79.25	1.78
Israel	13.58	3.53	0.52	106.45	79.25	1.85
Japan	9.60	3.54	0.73	102.73	79.08	2.19
Kenya	47.77	4.69	0.70	183.93	79.06	1.90
Nepal	13.28	3.84	0.55	104.81	79.31	1.79
Panama	18.63	3.72	0.67	115.45	79.12	1.89
Poland	12.23	3.33	0.44	103.44	79.31	1.65
Portugal	11.40	3.75	0.61	100.21	79.19	1.75
Puerto Rico	16.54	2.32	0.34	124.12	79.52	1.38
Romania	43.27	3.52	0.50	172.15	79.36	1.76
Russia	8.05	3.85	0.69	101.90	79.22	1.80
S. Africa	12.33	3.82	0.59	108.91	79.28	1.77

S. Korea	8.21	3.38	0.43	103.68	79.30	1.61
Serbia	13.96	3.61	0.53	105.85	79.28	1.72
Slovakia	19.75	3.45	0.49	113.64	79.25	1.81
Spain	14.23	3.63	0.50	106.72	79.28	1.72
UK	7.78	2.18	0.41	103.59	79.47	1.51
Vietnam	8.47	2.77	0.38	105.07	79.50	1.42

Note: the statistics calculated by Mplus are based on Mardia, Kent, & Bibby (1979) definitions.

All the sample statistics are significant at $p < .001$.

Table SI.4. Estimates of factor loadings based on partial scalar invariance model

	PWI1	PWI2	PWI3	PWI4	PWI5	PWI6	PWI7	PWI8
Armenia	.67	.73	.72	.63	.63	.65	.58	.47
Brasil	.68	.58	.73	.60	.59	.65	.63	.42
Bulgaria	.69	.59	.71	.64	.63	.68	.61	.37
Chile	.68	.52	.75	.68	.66	.63	.69	.35
Estonia	.50	.48	.62	.53	.65	.59	.63	.45
Hungaria	.61	.57	**.49**	.63	.70	.68	.57	.49
India	.62	.54	.58	.61	**.38**	.63	.65	.46
Indonesia	.78	.64	.74	.71	.74	.78	.75	.50
Iran	**.80**	.45	.60	.55	.55	.56	.50	.34
Israel	.61	.56	.69	.64	.66	.57	.66	.44
Japan	.69	.55	.69	.62	.70	.68	.75	**.12**
Kenya	.71	.72	.78	.68	.78	.80	.74	.58
Nepal	.58	.51	.56	.52	.55	.51	.53	.44
Panama	.81	**.53**	.86	.78	.78	.69	.77	.56
Poland	.70	.54	.72	.68	.72	.67	.63	.39
Portugal	.68	.56	.71	.60	.66	.60	.59	.41
Puerto Rico	.75	.64	.75	.60	.67	.68	.68	.41
Romania	.57	.60	.75	**.87**	.74	.70	.65	.39
Russia	.76	.69	.76	.69	.73	.75	.67	.51
S. Korea	.67	.52	.66	.72	.66	.75	.74	.37
Serbia	.59	.58	.71	.71	.68	.67	.65	.46
Slovakia	.59	.47	.64	.50	.58	.63	.60	.42
South Africa	.53	.54	.64	.59	.55	.59	.65	.42

Spain	.62	.53	.61	.59	.58	.56	.49	.23
UK	.70	.61	.75	.62	.72	.72	.74	.40
Vietnam	.60	.62	.71	.66	.71	.72	.74	.47

Note: non-invariant parameters are marked.

Table SI.5. Intercepts based on partial scalar invariance model

	PWI1	PWI2	PWI3	PWI4	PWI5	PWI6	PWI7	PWI8
Armenia	3.94	4.86	4.14	3.41	3.53	3.60	2.72	3.14
Brasil	4.27	4.10	4.20	3.73	2.94	3.48	2.89	3.26
Bulgaria	4.12	4.00	3.88	3.80	3.61	3.87	2.92	2.55
Chile	4.69	3.72	4.72	4.61	3.83	3.62	3.80	2.51
Estonia	3.91	4.23	4.44	4.22	5.34	4.36	3.94	4.75
Hungaria	4.11	4.28	5.19	4.14	4.74	4.27	2.84	3.76
India	3.94	3.61	2.97	3.81	3.85	3.57	3.28	3.61
Indonesia	4.40	4.05	3.44	3.70	4.00	4.15	3.37	3.22
Iran	3.79	3.69	3.70	3.75	3.28	3.31	2.50	2.44
Israel	4.29	4.44	4.44	4.43	4.42	2.99	3.70	3.53
Japan	4.09	3.66	3.28	3.63	3.97	3.79	3.36	2.65
Kenya	2.88	3.30	3.06	2.69	3.01	3.03	2.74	3.13
Nepal	3.97	3.89	3.48	3.85	3.58	3.28	2.85	3.75
Panama	4.01	4.59	3.88	3.51	3.42	2.59	3.00	3.13
Poland	4.64	3.81	4.11	4.15	4.63	3.97	3.39	2.77
Portugal	4.95	4.63	4.70	4.31	4.63	3.93	3.08	3.44
Puerto Rico	4.24	4.05	3.87	3.31	3.44	3.27	3.06	2.60
Romania	3.94	4.99	4.78	3.85	5.22	4.82	3.80	3.26
Russia	4.09	3.83	3.25	3.61	3.73	3.75	2.84	3.07
S. Korea	4.81	4.03	3.74	4.79	4.47	4.81	3.75	2.86
Serbia	2.66	3.97	3.59	4.25	3.60	3.45	2.83	2.85
Slovakia	4.80	4.13	4.39	3.77	4.52	4.80	3.45	4.13
South Africa	3.56	4.15	3.97	3.91	3.59	3.46	3.73	3.56

Spain	5.44	5.22	4.92	5.39	4.89	3.99	2.90	1.87
UK	4.83	4.40	4.38	3.94	4.82	4.16	3.94	2.94
Vietnam	3.92	4.53	3.92	4.25	4.43	4.43	4.17	3.49