Received: 18 July 2016 | Revised: 15 August 2017 | Accepted: 24 October 2017
DOI: 10.1002/jclp.22570

WILEY

RESEARCH ARTICLE

The mental health continuum-short form: The structure and application for cross-cultural studies–A 38 nation study

Magdalena Żemojtel-Piotrowska[1] | Jarosław P. Piotrowski[2] | Evgeny N. Osin[3] | Jan Cieciuch[1,4] | Byron G. Adams[5,6] | Rahkman Ardi[7] | Sergiu Bălţătescu[8] | Sergey Bogomaz[9] | Arbinda Lal Bhomi[10] | Amanda Clinton[11] | Gisela T. de Clunie[12] | Anna Z. Czarna[13] | Carla Esteves[14] | Valdiney Gouveia[15] | Murnizam H.J. Halik[16] | Ashraf Hosseini[17] | Narine Khachatryan[18] | Shanmukh Vasant Kamble[19] | Anna Kawula[20] | Vivian Miu-Chi Lun[21] | Dzintra Ilisko[22] | Martina Klicperova-Baker[23] | Kadi Liik[24] | Eva Letovancova[25] | Sara Malo Cerrato[26] | Jaroslaw Michalowski[27] | Natalia Malysheva[28] | Alison Marganski[29] | Marija Nikolic[30] | Joonha Park[31] | Elena Paspalanova[32] | Pablo Perez de Leon[33] | Győző Pék[34] | Joanna Różycka-Tran[35] | Adil Samekin[36] | Wahab Shahbaz[37] | Truong Thi Khanh Ha[38] | Habib Tiliouine[39] | Alain Van Hiel[40] | Melanie Vauclair[41] | Eduardo Wills - Herrera[42] | Anna Włodarczyk[43] | Illia Yahiiaev[44] | John Maltby[45]

[1]Cardinal Stefan Wyszynski University, Warsaw, Poland
[2]University of Social Sciences and Humanities, Poznan Campus, Poland
[3]National Research University Higher School of Economics, Moscow, Russia
[4]University of Zurich, University Research Priority Program Social Networks, Switzerland
[5]Tilburg University, The Netherlands, Poland
[6]University of Johannesburg, South Africa
[7]Airlangga University, Indonesia
[8]University of Oradea, Romania
[9]National Research Tomsk State University
[10]Tribhuvan University
[11]University of Puerto Rico, American Psychological Association
[12]Universidad Tecnológica de Panamá

[13] Jagellonian University, Poland
[14] Instituto Universitário de Lisboa (ISCTE-IUL), CIS-IUL, Lisboa, Portugal
[15] Federal University of Paraíba, Brasil
[16] University of Malaysia Sabah, Malaysia
[17] University of Melbourne, Australia
[18] Yerevan State University, Armenia
[19] Karnatak University, India
[20] Pedagogic University in Cracov, Poland
[21] Lingnan University, China
[22] Daugavpils University, Latvia
[23] Academy of Sciences of the Czech Republic, Czech
[24] Tallinn University, Estonia
[25] Comenius University in Bratislava, Slovakia
[26] University of Girona, Spain
[27] University of Social Sciences and Humanities, Poznan Campus
[28] Lomonosov Moscow State University, Russia
[29] Le Moyne College, US
[30] LUM University
[31] Nagoya University of Commerce and Business, Japan
[32] New Bulgarian University, Bulgaria
[33] Universidad Católica del Uruguay, Uruguay
[34] University of Debrecen, Hungary
[35] University of Gdansk, Poland
[36] S. Toraighyrov Pavlodar State University, Kazakhstan
[37] Massey University, Australia
[38] University of Social Sciences and Humanities, VNU, Hanoi, Vietnam
[39] Oran University, Algeria
[40] Ghent University, Belgium
[41] Instituto Universitário de Lisboa (ISCTE-IUL)
[42] Universidad de Los Andes, Colombia
[43] Universidad Católica del Norte (Chile)
[44] Taras Shevchenko National University of Kyiv, Ukraine
[45] University of Leicester, United Kingdom

Correspondence
Magdalena A. Żemojtel-Piotrowska, Institute of Psychology, Wóycickiego 1/3 budynek 14; 01-938 Warsaw, Poland
Email: psymzp@ug.edu.pl

Funding information
The work of Jarosław P. Piotrowski was supported by research grant rewarded by University of Social Sciences and Humanities, Poznan Faculty. The work of Jan Cieciuch was supported by grants 2014/14/M/HS6/00919 from the National Science Centre, Poland. The work of Evgeny Osin was funded by the Russian Academic Excellence project '5-100'. Order of authors reflects their contribution to the work.

Abstract

Objective The Mental Health Continuum-Short Form (MHC-SF) is a brief scale measuring positive human functioning. The study aimed to examine the factor structure and to explore the cross-cultural utility of the MHC-SF using bifactor models and exploratory structural equation modelling.

Method Using multigroup confirmatory analysis (MGCFA) we examined the measurement invariance of the MHC-SF in 38 countries (university students, $N = 8{,}066$; 61.73% women, mean age 21.55 years).

Results MGCFA supported the cross-cultural replicability of a bifactor structure and a metric level of invariance between student samples. The average proportion of variance explained by the general factor was high (ECV = .66), suggesting that the three aspects of mental health (emotional, social, and psychological well-being) can be treated as a single dimension of well-being.

Conclusion The metric level of invariance offers the possibility of comparing correlates and predictors of positive mental functioning across countries; however, the comparison of the levels of mental health across countries is not possible due to lack of scalar invariance. Our study has preliminary character and could serve as an initial assessment of the structure of the MHC-SF across different cultural settings. Further studies on general populations are required for extending our findings.

KEYWORDS
cross-cultural study, measurement invariance, Mental Health Continuum-Short Form

1 | INTRODUCTION

Emerging adults are frequently exposed to the challenges of transition into adulthood (low personal finances, entering the workplace, changes in personal relationships; Arnett, 2000; Roberts, Golding, Towell, Reid, & Woodford, 2000) and are at risk for various mental health problems (Eisenberg, Gollust, Golberstein, & Hefner, 2007). The mental health of university students is often found to be worse than that of the general population (Mikolajczyk et al., 2008; Stock et al., 2008; Vaez, Kristenson, & Laflamme, 2004), perhaps due to the additional challenges and risks facing them, such as increased financial worries, costs, and debt associated with university, academic pressure, moving away from home, changes in sources of emotional support, dealing with new environments, and increased exposure to drinking and drug-taking culture. However, despite the variety of studies reporting low mental health scores among students around the globe (see Boot, Donders, Vonk, & Meijman, 2009; Kurré, Scholl, Bullinger, & Petersen-Ewert, 2011; Stewart-Brown et al., 2000; Vaez et al., 2004), still little research has been done regarding comparisons between different countries. This may be due to the lack of a valid standardized measure which could be used in cross-cultural studies aimed at assessing well-being and could consequently guide the assessment and mental health promotion of university students across the globe.

The problem of mental health and its assessment is one of the broadest and most complex issues in psychology (see Sirgy, 2012), given the many different sources and factors contributing to well-being (Keyes, 1998, 2002; Keyes, Ryff, & Shmotkin, 2002; Ryff, 1989). Therefore, it seems crucial to develop a single valid and reliable instrument that could be used to study, assess, and promote the students' mental health. To achieve these goals we focus on exploring the cross-cultural utility of a measure that assesses several theoretical domains of well-being: the Mental Health Continuum-Short Form (MHC-SF; Keyes, 1998).

The MHC-SF is based on the concept of positive mental health proposed by Keyes (2002) and is an abbreviated form of the 40-item MHC-LF (Keyes, 2002). It is an effort to integrate hedonic and eudaimonic aspects of well-being. Specifically, the Mental Health Continuum is regarded as a syndrome encompassing three broad aspects: emotional well-being (EWB, comprising positive emotions along with life satisfaction); social well-being (SWB, based on the definition offered by Keyes, 1998, comprising social coherence, social acceptance, social actualization, social contribution, and social integration); and psychological well-being (PWB, based on a model by Ryff, 1989, comprising self-acceptance, positive relationships with others, autonomy, purpose in life, environmental mastery, and personal growth).

Comprising 14 items, the MHC-SF (Keyes et al., 2008) measures all three dimensions: emotional, psychological, and social well-being. It can be used both for research purposes (as an indicator of positive functioning of individual) and for diagnosis of the levels of positive functioning (Keyes, 2002). The MHC-SF captures three categorical diagnoses: flourishing, languishing, and moderate mental health. Flourishing is diagnosed when someone reports having experienced at least one of the three hedonic well-being symptoms (items 1–3) and at least 6 of the 11 positive functioning symptoms (items 4–14) "every day" or "almost every day" within the past month. Languishing is diagnosed when someone reports having experienced at least one of the three hedonic well-being symptoms and at least 6 of the 11 positive functioning symptoms "never" or "once or twice" in the past month. Individuals who are neither "languishing" nor "flourishing" are considered "moderately mentally healthy" (Keyes, 2002).

Currently, there are several versions of this scale in different languages (see Keyes, 1998, and Karaś, Cieciuch, & Keyes, 2014, for review), including Korean (Young-Jin, 2014), Serbian (Jovanović, 2015), Italian (Petrillo, Capone, Caso, & Keyes, 2015), and Polish (Karaś et al., 2014). Therefore MHC-SF seems to be a perfect tool for cross-cultural research studies of well-being among university students. These studies could focus on searching for the risk factors and factors important for increasing the mental health, as well as on diagnosing the number of languishing, flourishing, and moderately healthy individuals within different populations.

As the MHC-SF has been used in a number of countries, one could expect that this tool is well-validated in different cultural contexts and there are no controversies and/or obstacles to implementing it in the cross-cultural surveys. Nevertheless, despite the work of Keyes (1998) that assumes a three-factor structure of the MHC-SF, some researchers suggest that the proposed three-factor structure of the MHC-SF scale is problematic, both theoretically, because it fails to provide the information needed to justify the calculation of a general well-being index (de Bruin, & du Plessis, 2015), and empirically, because it often produces only marginally acceptable fit indices (Jovanović, 2015).

To address these issues, two alternative, more flexible models were proposed for MHC-SF: first, a bifactor model, in which each item loads on a general factor (reflecting a common construct underlying all the items) and on one of the uncorrelated specific or "group" factors (which capture the content similarity of homogeneous groups of items forming the subscales). This approach was suggested as particularly useful for composite models of subjective well-being because it allows to separate the general well-being dimension from specific factors related to particular life domains or aspects of human functioning (Sirgy, 2012) and it has been applied successfully with MHC-SF (Jovanović, 2015). The advantage of this model is that it allows the separation of the item variance associated with the general factor and specific factors and evaluates the reliability of a general score and the discriminant validity of subscale scores using a range of indices discussed below (Reise, 2012; see also Chen, West, & Sousa, 2006, for a comparison of correlated-factor and bifactor models).

Another analytic solution recently proposed for MHC-SF is exploratory structural equation modeling (ESEM). Unlike conventional confirmatory factor analysis models (termed "independent-cluster model," or ICM-CFA), ESEM models allow for nonzero cross-loadings, addressing the issue of imperfect indicators. Advocates of this approach have argued that the assumption of the absence of cross-loadings in the conventional CFA may be unrealistic, resulting in overestimation of factor covariances; some of the MHC-SF items were indeed found to show statistically significant cross-loadings (Joshanloo & Jovanović, 2017).

However, ESEM models have some important drawbacks. First, they can be viewed as more data-driven because strong item loadings on nontarget factors may affect the theoretical interpretation of factors. Second, they include a much larger number of free parameters (i.e., loadings), compared to conventional CFA models, which may result in increased sample size requirements and entail convergence difficulties. Finally, a ESEM model with three correlated factors, which was previously applied for MHC-SF (Joshanloo & Jovanović, 2017), may not be optimal because of the presence of a common factor, whose variance could contribute to item cross-loadings, resulting in overestimation of factor overlap. Theoretically, a bifactor ESEM model (Morin, Arens, & Marsh, 2016) should be more relevant for MHC-SF, but it may entail identification or convergence problems due to its complexity.

Past cross-cultural studies using MHC-SF did not take advantage of bifactor or ESEM approaches and have compared only a limited number of national samples (Joshanloo, Wissing, Khumalo, & Lamers, 2013). Because the MHC-SF seems a promising brief measure of positive mental functioning, the issues of structure and measurement invariance of MHC-SF need to be studied in diverse cultural contexts to reveal the possibilities and limitations of this measure for multicultural projects.

1.1 | The present study

The aim of the present study is to examine, first, the structure of MHC-SF in different cultural contexts, comparing the fit of the bifactor model and that of the three-factor solution and, second, the applicability of MHC-SF to cross-cultural studies using multigroup CFA analyses with data from 38 countries. This aim is not only theoretical, but it could also allow to address applied issues (i.e., the comparability of findings on mental health, in terms of conceptual invariance, findings on predictors and correlates, and, finally, mean scores) among youth populations from different cultures.

2 | METHOD

2.1 | Sample and procedure

The sample included 8,066 university students (61.7 % women), ranging in age from 16 to 50 (mean [*M*] = 21.55, standard deviation [*SD*] = 4.37), originating from 38 countries (see Table 1 for details). The students filled out the MHC-SF as part of a broader research project on entitlement and well-being. In addition to the MHC-SF, the study included other measures of subjective well-being: Personal Well-being Index (Cummins, Eckersley, Pallant, Van Vugt, & Misajon, 2003); Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985); Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988); and two scales measuring attitudes–Entitlement Attitudes Scale (Żemojtel-Piotrowska et al., 2015) and belief in life as zero-sum game (Różycka-Tran, Boski, & Wojciszke, 2015).

Data were collected in paper-pencil form and also online (presented in Table 1) between March 2015 and March 2016. The students were recruited at universities during their classes by the members of the research team and did not receive any financial remuneration for participation. The students participated in the study voluntarily and informed consent was obtained from all study participants. The registered data were alphanumerically coded, ensuring anonymity. The study has been conducted according to the principles expressed in the Declaration of Helsinki. All procedures were approved by each participating University Ethics Committee.

The selection of participating countries aimed to reflect cultural diversity in the most comprehensive way possible. In terms of cultural regions, we included countries representative of all Huntington (1996) cultural groups (i.e., Western, Orthodox, Confucian, Japanese, Latin American, Hindu, Buddhist, Islamic, African, and Sinic), and, in terms of religion, we had countries representing all main world religions. In the current study, we included data from Europe (16), Asia (13), Africa (3), and Latin America (6). Former studies indicate the importance of cultural, political, and economic factors related to subjective well-being. For instance, subjective well-being is related to income inequalities (Berg & Veenhoven, 2010), values (Sagiv & Schwartz, 2000; Welzel & Inglehart, 2010), and religion (Donahy et al., 1998). Therefore, our aim was to include countries with different levels of affluence, cultural values, and religion to indicate usefulness of the MHC-SF in measuring mental health as a multidimensional construct.

2.2 | Measures

The MHC-SF (Keyes, 2013) comprises 14 items that represent various aspects of well-being (the items were chosen from the longer version of this tool, as the most prototypical for each aspect of well-being). The response scale consists of 6 points, which describe the frequency of experiencing various well-being symptoms during the past month, ranging from 1 (*never*) to 6 (*every day*). The MHC-SF allows two kinds of assessments: level of well-being, with its three dimensions (social, psychological, and emotional), and a categorical assessment of mental health status, with three categories (flourishing, i.e., high levels of well-being; languishing, i.e., the absence of mental health; and moderate mental health, located between these two extremes).

TABLE 1 Descriptive statistics for the 38 countries

Country	N	Female%	Age*M*(*SD*)	SES*M*(*SD*)	Language	procedure	MHC-SF*M*(*SD*)	*α*
Algeria	240	61.25	19.54 (1.58)	4.13 (1.30)	Arabic	Paper-pencil	51.04(11.39)	.79
Armenia	223	47.98	19.00 (1.17)	4.98 (1.20)	Armenian	Paper-pencil	55.74(10.31)	.81
Azerbaijan	120	60.83	20.83 (1.95)	3.38 (0.99)	Russian	Online	51.55(12.19)	.88
Belgium	232	74.14	19.74 (3.95)	4.63 (1.09)	Flemish	Online	53.99(10.27)	.87
Brazil	223	63.68	20.94 (5.21)	4.38 (0.99)	Portuguese	Paper-pencil	51.44(12.25)	.89
Bulgaria	200	66.00	23.59 (5.25)	4.66 (1.16)	Bulgarian	Paper-pencil	53.16(11.34)	.87
Chile	241	52.28	22.00 (2.10)	4.34 (1.03)	Spanish	Paper-pencil	56.07(11.63)	.90
Colombia	138	50.00	18.82 (1.72)	5.74 (0.90)	Spanish	Online	58.09(12.38)	.92
Czech Republic	223	74.89	24.52 (7.75)	4.37 (1.23)	Czech	Paper-pencil	50.50(12.08)	.89
Estonia	301	69.10	23.11 (6.05)	4.41 (1.23)	Esti	Online	53.84(11.24)	.89
Germany	233	82.83	24.99 (6.53)	4.56 (1.29)	German	Online	54.51(13.02)	.91
Hong Kong	172	68.02	18.82 (1.16)	4.31 (1.39)	English	Paper-pencil	53.17(12.03)	.94
Hungary	206	68.93	21.00 (1.68)		Hungarian	Paper-pencil	56.92(10.46)	.88
India	200	68.50	22.59 (1.45)	4.32 (1.07)	English	Paper-pencil	63.41(10.40)	.86
Indonesia	200	50.00	21.38 (1.65)	4.70 (1.02)	Bahasa	Online	58.98(11.90)	.90
Iran	201	50.25	21.28 (1.53)	4.46 (1.41)	English	Paper-pencil	49.25(11.80)	.86
Japan	195	26.15	18.96 (1.13)	4.11 (1.33)	Japanese	Paper-pencil	42.55(12.81)	.89
Kazakhstan	285	74.74	20.12 (2.32)	3.43 (0.89)	Russian	Online	58.02(13.95)	.92
Kenya	162	53.09	23.49 (4.54)	4.07 (0.92)	English	Paper-pencil	58.09(9.47)	.80
Korea (S)	212	54.72	22.20 (1.91)	3.90 (1.24)	Korean	Paper-pencil	45.81(10.97)	.92
Latvia	221	72.40	27.80 (7.91)	2.97 (0.79)	Russian	Online	53.40(9.86)	.90
Malaysia	199	50.25	21.96 (1.22)	4.02 (1.20)	Malay	Paper-pencil	55.91(11.30)	.93
Nepal	203	49.75	22.70 (4.44)	4.08 (0.93)	English	Paper-pencil	55.34(10.22)	.82
Panama	170	33.53	21.41 (5.08)	4.13 (1.00)	Spanish	Online	56.83(12.89)	.90
Pakistan	200	49.00	21.50 (1.59)	4.97 (1.05)	English	Paper-pencil	54.36(10.13)	.82
Poland	227	60.79	22.31 (4.14)	4.69(1.15)	Polish	Paper-pencil	49.83(13.11)	.92
Portugal	193	77.20	22.18 (5.73)	4.11 (1.08)	Portuguese	Online	54.52(11.50)	.90
Puerto Rico	300	42.67	20.26 (2.23)	4.14(1.24)	Spanish	Paper-pencil	55.67(12.76)	.91
Romania	206	48.54	21.33 (3.47)	4.72 (1.13)	Romanian	Paper-pencil	58.45(11.68)	.90
Russia	229	79.48	21.64 (4.13)	3.11 (1.04)	Russian	Online	49.90(13.54)	.90
Serbia	205	60.98	22.46 (5.75)	3.77 (1.10)	Serbian	Paper-pencil	53.06(12.08)	.90
Slovakia	202	71.78	21.13 (1.26)	4.76 (1.00)	Slovak	Paper-pencil	53.03(11.78)	.90
Spain	196	50.51	21.02 (4.66)	4.01 (1.05)	Spanish (Catalan)	Online	56.29(11.57)	.89
South Africa	186	67.20	20.17 (1.86)	4.49(1.25)	English	Paper-pencil	57.58(11.06)	.86
Ukraine	171	80.70	19.86 (2.66)	3.21 (1.06)	Russian	online	53.00(12.28)	.88
United Kingdom	303	80.86	19.53 (2.80)	4.21 (1.33)	English	online	54.50(13.40)	.92
Uruguay	197	80.71	23.51 (6.14)	5.02 (1.00)	Spanish	Paper-pencil	56.81(10.17)	.87
Vietnam	251	52.19	20.51 (2.68)	4.25(1.01)	Vietnamese	Paper-pencil	53.26(14.17)	.92
Overall	8066	61.73	21.55 (4.37)	4.27 (1.25)			54.12(12.37)	.89

Note. MHC-SF = Mental Health Continuum-Short Form; M = mean; SD = standard deviation; SES = subjective economic status of family (range 1–7).

We used translation and back-translation procedure to obtain versions of the scale in different languages. The resulting back-translated versions were discussed with the author of the MHC-SF, Corey Keyes. We do not report the results of validation of the MHC-SF because they would go beyond the scope of the present paper. However, in different countries we have found a consistent pattern of negative correlations of MHC-SF with revengefulness and belief in life as zero-sum game, as well as positive correlations of MHC-SF with other scales measuring subjective well-being.

2.3 | Data analysis

The analyses were conducted using SPSS (version 20) and Mplus (version 7.4). The robust mixed linear model (MLM) Maximum Likelihood with Satorra-Bentler-scaled chi-square resulted in fewer convergence problems and inadmissible solutions for the bifactor model, compared to the (ML and MLR) Robust Maximum Likelihood estimators. Unfortunately, the MLM estimator in Mplus currently does not handle missing data. Because the percentage of missing responses was quite small (0.28%) and the data were missing at random, we used expectation maximization (EM) imputation in SPSS to impute the missing values, to take advantage of the MLM estimator.

2.3.1 | Preliminary CFA

First, we performed a CFA in each sample separately. These analyses were aimed at finding the best measurement model of the MHC-SF to be used as a basis for cross-cultural comparison.

We identified the models by fixing the latent factor variances to 1 and freely estimating the factor loadings. Given the known limitations of the chi-square test of overall model fit (dependence on sample size making the results incomparable across samples and reliance on the null hypothesis of exact overall fit which is too stringent to be informative in evaluating the usefulness of a model: see West, Taylor, & Wu, 2012), we used practical fit indices (comparative fit index [CFI], root mean square error of approximation [RMSEA], and standardized root mean square residual [SRMR]) to assess model fit. We followed the guidelines proposed by Hu and Bentler (1999; i.e., the values of CFI close to .95 or above, RMSEA close to .06 or below, SRMR close to .08 or below as indications of good fit, using these indices in combination; Brown, 2015). To compare the fit of nested models in individual samples, we relied on the scaled chi-square difference test (Satorra & Bentler, 2001).

Based on theory and previous findings (Jovanović, 2015; Karaś et al., 2014), we tested four different CFA models of the MHC-SF: (a) a single-factor model, in which all 14 items load on one underlying dimension of well-being; (b) a two-factor model with two correlated dimensions of well-being–hedonic well-being (comprising EWB; items 1 through 3) and eudaimonic well-being (comprising both SWB and PWB; items 4 through 14); (c) a three-factor model with three correlated dimensions of well-being–hedonic well-being (items 1 to 3), eudaimonic social well-being (items 4 to 8), and eudaimonic psychological well-being (items 9 to 14); and (d) a bifactor model (Reise, 2012), with a general factor and three uncorrelated "group factors" capturing specific variance or hedonic, social, and psychological well-being. We did not test the hierarchical model with a single second-order factor separately because a hierarchical solution with three first-order factors is mathematically equivalent to the correlated-factor model.

The advantage of the bifactor model is that it makes it possible to separate the general and specific variance. To evaluate the reliability of the general dimension and the subscales, we calculated the omega coefficient (Reise, 2012), which is similar to the alpha because it reflects the proportion of total item variance explained by the model, with joint contribution of the general well-being factor and group factors. To separate the effects of the general well-being factor and those of the group factors, we calculated coefficients ω_H and ω_S (Reise, 2012), the former reflecting the share of total variance explained by the general factor and the latter reflecting the unique share of variance explained by each group factor (excluding the contribution of the general factor). We also calculated the explained common variance (ECV) coefficient (Reise, Scheines, Widaman, & Haviland, 2013), which measures the relative strength of the general factor to the group dimensions.

2.3.2 | Measurement invariance analyses

The second aim of study was to evaluate the measurement invariance of the MHC-SF and establish nonequivalent parameters using a multigroup bifactor CFA model. To evaluate the absolute model fit, we used the same criteria for practical fit indices as described above. Because the chi-square difference test is known to be overly sensitive in large samples (Chen, 2007), we relied on practical fit indices to compare the nested models, using the ΔCFI and ΔRMSEA cutoff values of .010 and .015, respectively, as indicators of pronounced difference in fit between the nested models (Chen, 2007; Cheung & Rensvold, 2002). Because of the large number of parameter constraints tested, we relied only on modification indices significant at $p < .05$ with Bonferroni correction to prevent false positives. We relaxed the parameter constraints sequentially (Yoon & Kim, 2014), one at a time, after which the model was re-estimated.

There are three levels of measurement invariance that are most commonly used to establish whether a measure is equivalent. Configural invariance indicates that the general factor structure of the measure is the same across different groups. At this level, the construct is measured by the same set of indicators in different samples. Metric invariance implies that the factor loadings of items are similar across groups. At this level, the effects of correlates and/or predictors of the measure may be compared across samples. Scalar invariance indicates that item intercepts are equal across groups. At this level, mean scores may be compared between samples (Davidov, Meuleman, Cieciuch, Schmidt, & Billiet, 2014). Scalar invariance is rarely found in large cross-cultural comparisons (see Davidov et al., 2014), so we expected to find metric invariance of the MHC-SF. To examine the structure of the scale and its cross-cultural replicability, however, only configural invariance is required. Because most cross-cultural studies focus on examining predictors and correlates of subjective well-being, the metric level of invariance is sufficient.

2.3.3 | ESEM analyses

We also performed single-group ESEM analyses based on a model with three correlated factors and a bifactor model. However, because of complexity of this model, which resulted in convergence issues, we could not use the ESEM model as a basis for multigroup comparison and we present these results as supplementary findings.

3 | RESULTS

3.1 | Preliminary CFA

3.1.1 | Single-group analyses

The single-factor model did not fit the data well, with at least two out of three fit indices lying outside the acceptable ranges for all samples. Across the 38 countries, the CFI ranged from .508 to .868 ($M = .791$, $SD = .066$), the RMSEA ranged from .079 to .144 ($M = .112$, $SD = .015$), and the SRMR ranged from .058 to .134 ($M = .079$, $SD = .013$). The two-factor model (i.e., factors representing hedonic and eudaimonic well-being) showed a better fit, with CFI ranging from .587 to .926 ($M = .848$, $SD = .060$), RMSEA ranging from .067 to .133 ($M = .095$, $SD = .014$), and SRMR ranging from .053 to .132 ($M = .072$, $SD = .013$). However, based on the combination of fit indices, the fit was still unacceptable in all countries but one (Ukraine).

The fit indices for the three-factor model and the bifactor model are shown in Table 2. Based on the combination of indices, the three-factor model showed a good fit in two countries (Ukraine and Uruguay) and acceptable fit in 14 countries (Azerbaijan, Belgium, Brazil, Czech Republic, Estonia, Hungary, Indonesia, Japan, Kazakhstan, Malaysia, Portugal, Russia, South Africa, Vietnam). In most of the remaining cases, the fit was marginal. The correlations between the factors were moderate to strong in all samples. The mean correlation between the emotional and psychological well-being factors was .75, and social well-being was correlated at .69 and .62 with psychological and emotional well-being, respectively.

The unrestricted bifactor model failed to converge in 10 out of 38 samples (Bulgaria, Czech Republic, Hong Kong, Hungary, India, Latvia, Panama, South Africa, Vietnam, Uruguay) due to a negative error variance. In two samples

TABLE 2 Fit indices for the three-factor and the bifactor model in 38 countries

	Three-Factor Model				Bifactor Model			
Country	$\chi^2(74)$	CFI	RMSEA [90% CI]	SRMR	$\chi^2(63)$	CFI	RMSEA [90% CI]	SRMR
Algeria	135.03***	.899	.059 [.043, .074]	.057	101.40**	.936	.050 [.031, .068]	.046
Armenia	177.02***	.838	.079 [.064, .094]	.067	128.26***	.898	.068 [.051, .085]	.058
Azerbaijan	116.43**	.922	.069 [.044, .092]	.069	77.46	.973	.044 [.000, .074]	.051
Belgium	154.32***	.922	.068 [.053, .084]	.060	103.63**	.960	.053 [.034, .070]	.046
Brazil	156.93***	.929	.071 [.055, .086]	.058	96.90**	.971	.049 [.028, .068]	.038
Bulgaria	165.95***	.887	.079 [.063, .095]	.073	101.80**	.952	.055 [.035, .075]	.043
Chile	208.16***	.889	.087 [.073, .101]	.070	121.09***	.952	.062 [.045, .078]	.044
Colombia	164.27***	.888	.094 [.075, .113]	.066	81.50	.974	.046 [.000, .073]	.041
Czech R.	149.00***	.929	.067 [.052, .083]	.056	98.32**	.966	.050 [.030, .069]	.039
Germany	211.84***	.908	.089 [.075, .104]	.074	114.26***	.925	.059 [.041, .076]	.054
Estonia	156.24***	.941	.061 [.047, .074]	.048	109.46***	.966	.049 [.033, .065]	.037
Hong Kong	160.06***	.940	.082 [.065, .100]	.049	126.53***	.956	.077 [.057, .096]	.039
Hungary	147.03***	.912	.069 [.053, .086]	.064	100.54**	.955	.054 [.033, .073]	.043
India	140.04***	.898	.067 [.050, .084]	.059	115.30***	.920	.064 [.045, .083]	.051
Indonesia	147.04***	.921	.070 [.053, .087]	.061	99.21**	.961	.054 [.032, .073]	.047
Iran	190.15***	.862	.088 [.073, .104]	.066	157.93***	.887	.087 [.070, .104]	.057
Japan	145.21***	.918	.070 [.053, .087]	.066	86.73*	.973	.044 [.016, .065]	.046
Kazakhstan	175.91***	.938	.070 [.056, .083]	.045	125.19***	.965	.059 [.044, .074]	.038
Kenya	191.15***	.777	.099 [.082, .116]	.103	136.36***	.860	.085 [.065, .104]	.086
Korea (S)	209.09***	.902	.093 [.078, .108]	.070	134.50***	.948	.073 [.056, .090]	.043
Latvia	182.14***	.881	.081 [.066, .096]	.081	106.35***	.952	.056 [.037, .074]	.046
Malaysia	142.83***	.946	.068 [.051, .085]	.051	84.20*	.983	.041 [.010, .063]	.047
Nepal	144.92***	.865	.069 [.052, .085]	.071	108.88***	.913	.060 [.040, .079]	.063
Panama	198.87***	.869	.100 [.083, .116]	.089	119.56***	.941	.073 [.053, .092]	.055
Pakistan	165.01***	.843	.078 [.062, .094]	.075	86.10*	.982	.043 [.015, .064]	.047
Poland	192.74***	.915	.084 [.070, .099]	.061	142.04***	.944	.074 [.058, .091]	.043
Portugal	139.51***	.935	.068 [.050, .085]	.058	92.01**	.971	.049 [.025, .069]	.039
Puerto R.	253.88***	.892	.090 [.078, .102]	.067	132.31***	.958	.061 [.046, .075]	.039
Romania	170.37***	.911	.080 [.064, .095]	.063	133.72***	.935	.074 [.056, .091]	.057
Russia	174.29***	.921	.077 [.062, .092]	.065	120.41***	.955	.063 [.046, .080]	.048
Serbia	223.02***	.868	.099 [.084, .114]	.074	111.47***	.957	.061 [.042, .080]	.042
Slovakia	221.42***	.873	.099 [.084, .114]	.095	112.10***	.958	.062 [.043, .081]	.042
S. Africa	135.29***	.905	.067 [.049, .084]	.061	95.32**	.950	.053 [.029, .073]	.048
Spain	177.14***	.886	.084 [.068, .100]	.077	98.31**	.961	.053 [.032, .073]	.047
Ukraine	109.55**	.956	.053 [.030, .073]	.056	74.20	.986	.032 [.000, .059]	.038
United Kingdom	247.14***	.915	.088 [.076, .100]	.066	155.31***	.955	.070 [.056, .083]	.038
Uruguay	118.98***	.950	.056 [.036, .074]	.063	82.37	.978	.040 [.000, .062]	.041
Vietnam	187.83***	.923	.078 [.064, .092]	.057	124.38***	.959	.062 [.046, .078]	.041

Note. CFI = comparative fit index; RMSEA = root mean square of approximation; SRMR = standardized root mean square residual; CI = confiedence interval. The comparisons between models are impossible as they are not nested models. Satorra-Bentler χ^2, ***$p < .001$. **$p < .01$. *$p < .05$.

(Vietnam, South Africa), a proper solution could be reached by adjusting the default starting values for group factor loadings. In five other samples in which the model converged, one of the estimates of residual variances was negative, but not significantly different from zero, suggesting normal sampling variation. An investigation of parameter estimates has shown that in many samples the loadings on one of the group factors (typically, PWB) were generally quite low. Negative error variances are often found when models with a relatively large number of free parameters and low empirical factor loadings are tested in samples of modest size (Chen, Bollen, Paxton, Curran, & Kirby, 2001).

We explored two approaches to resolve this issue: first, to improve the model identification by ruling out inadmissible solutions, we introduced inequality constraints, restricting the estimates of residual variances of observed variables to values above 0. As a result, model convergence was obtained in all samples. The constrained bifactor model showed good fit in 16 countries (Azerbaijan, Belgium, Brazil, Bulgaria, Czech Republic, Estonia, Hungary, Indonesia, Japan, Latvia, Malaysia, Portugal, South Africa, Spain, Ukraine, and Uruguay) and acceptable fit in all others, except Kenya and Iran, where the fit was marginal. The investigation of modification indices revealed an unexplained error covariance of items 9 and 10 in the Kenyan sample. In Iran, we found no pronounced modification indices, but exploratory factor analyses showed that items 4 to 8 failed to form a single dimension. To avoid the necessity for model modifications, we opted to exclude these two samples from the multigroup model.

Second, to investigate the potential bias introduced by constraints, we simplified the model by omitting the PWB group factor (Eid, 2000; Reise, 2012). The resulting incomplete bifactor model (unconstrained) converged successfully in all samples. Predictably, the fit of the incomplete bifactor model was significantly worse, compared to that of the full bifactor model (scaled chi-square difference test, $p < .05$) in all but three samples (Hong Kong, Kazakhstan, and Malaysia). The differences in practical fit indices were quite small (average ΔCFI = −.025, ΔRMSEA = .010, ΔSRMR = .008), generally favoring the full bifactor model. The fit indices for this model are given in Supporting Information.

3.1.2 | Bifactor structure analyses

Based on the bifactor models for each sample, we calculated a set of indices to evaluate the reliability and dimensionality of the MHC-SF in each sample. The results are shown in Table 3. The ω reliability coefficients, reflecting the proportion of true score variance (with contribution of both the general factor and group factors), ranged from .82 to .95 for the general well-being index and from .57 to .92 for the subscales, indicating good reliability. The social well-being subscale showed somewhat lower reliability, compared to the emotional and psychological well-being subscales.

The ω_H coefficient, reflecting the proportion of total variance explained by the general factor, ranged from .56 to .87, indicating a substantial contribution of the general factor. The ECV index, reflecting the share of the general factor in the true score variance, ranged from .40 to .76 ($M = .66$). This suggests that, on average, two thirds of the variance captured by the MHC-SF is shared by the three scales, and only one third is related specifically to emotional, psychological, or social well-being. According to O'Connor's (2014) recommendations, ECV values above .70 indicate unidimensionality of scales. ECV exceeded this value in Brazil, Colombia, Germany, Estonia, Hong Kong, Iran, Kazakhstan, Malaysia, Poland, Russia, United Kingdom, and Vietnam, suggesting that the general dimension of MHC-SF may be most relevant in these countries as an indicator of overall mental health.

The residual reliability coefficients (ω_S) reflect the proportion of true score variance of each subscale excluding the contribution of the general factor. The psychological well-being subscale reveals a comparatively small amount of unique variance ($M = .12$), indicating that the variance it captures is mainly shared by all three subscales of MHC-SF. The emotional and social well-being subscales emerge as more distinct ($M = .29$ and .31, respectively), suggesting that their associations with other variables may be different from those exhibited by the MHC-SF as a whole.

The structural coefficients based on the incomplete bifactor model (provided in Supporting Information) were convergent with those based on the full bifactor model, with ω reliability in the .81–.94 range and ω_H in the .55–.88 range. The ECV (ranging from .44 to .82, $M = .71$) and the residual reliability coefficients for the two subscales modelled were marginally higher ($M = .33$ and .36 for EWB and SWB, respectively), but led to the same substantive conclusions.

TABLE 3 Reliability and dimensionality indices for the MHC-SF in 38 countries

	Reliability, ω				Variance Explained				
Country	Gen	EWB	SWB	PWB	ω_H	ω_S EWB	ω_S SWB	ω_S PWB	ECV
Algeria	.82	.73	.63	.67	.68	.17	.30	.24	.61
Armenia	.83	.74	.57	.82	.68	.42	.40	.07	.54
Azerbaijan	.91	.81	.78	.82	.83	.33	.28	.02	.69
Belgium	.90	.87	.74	.83	.78	.39	.35	.09	.62
Brazil	.92	.85	.74	.86	.83	.31	.27	.10	.71
Bulgaria	.90	.84	.73	.81	.80	.18	.27	.14	.66
Chile	.91	.83	.80	.85	.79	.18	.32	.22	.65
Colombia	.94	.89	.84	.87	.86	.12	.28	.09	.71
Czech Rep.	.92	.85	.80	.83	.83	.23	.36	.06	.69
Estonia	.91	.85	.76	.84	.83	.32	.31	.03	.70
Germany	.94	.88	.81	.89	.87	.18	.33	.00	.72
Hong Kong	.95	.90	.87	.92	.87	.27	.27	.11	.74
Hungary	.90	.79	.78	.84	.80	.27	.30	.14	.62
India	.89	.76	.79	.79	.76	.32	.38	.07	.57
Indonesia	.92	.82	.81	.87	.83	.31	.30	.07	.69
Iran	.88	.83	.67	.78	.78	.09	.15	.31	.71
Japan	.92	.83	.81	.86	.84	.46	.17	.11	.66
Kazakhstan	.94	.85	.83	.89	.87	.23	.28	.00	.76
Kenya	.86	.75	.82	.78	.56	.50	.20	.68	.40
Korea (S)	.94	.91	.83	.89	.85	.36	.35	.03	.69
Latvia	.93	.88	.83	.87	.76	.18	.47	.28	.60
Malaysia	.95	.87	.88	.89	.86	.40	.00	.28	.71
Nepal	.86	.67	.75	.77	.71	.40	.43	.07	.52
Pakistan	.86	.69	.75	.80	.71	.44	.50	.00	.51
Panama	.93	.88	.83	.88	.83	.47	.35	.00	.64
Poland	.94	.88	.83	.89	.85	.32	.35	.03	.72
Portugal	.92	.86	.79	.87	.82	.29	.29	.14	.69
Puerto Rico	.93	.87	.82	.86	.81	.15	.32	.20	.69
Romania	.92	.87	.77	.89	.80	.49	.38	.06	.65
Russia	.92	.86	.76	.88	.84	.30	.34	.04	.71
Serbia	.92	.88	.77	.87	.82	.11	.35	.20	.67
Slovakia	.93	.78	.81	.89	.82	.17	.36	.15	.66
Spain	.92	.85	.77	.87	.83	.36	.34	.01	.66
South Africa	.89	.76	.77	.79	.77	.27	.40	.01	.62
Ukraine	.91	.84	.75	.84	.80	.35	.27	.15	.67
United Kingdom	.94	.90	.86	.85	.85	.18	.29	.13	.72
Uruguay	.90	.87	.72	.85	.78	.31	.38	.08	.63
Vietnam	.94	.84	.86	.89	.86	.37	.11	.18	.71
M	.91	.83	.78	.85	.80	.29	.31	.12	.66
SD	.03	.06	.06	.05	.06	.11	.09	.13	.07

Note. MHC-SF = Mental Health Continuum-Short Form; Gen = general score; EWB = emotional well-being; SWB = social well-being; PWB = psychological well-being; ECV = explained common variance; M = mean; SD = standard deviation.

TABLE 4 Fit indices for the multigroup models

Model	S-B χ^2	*df*	CFI	RMSEA [90 % CI]	SRMR
ICM-CFA models (36 countries)					
Configural invariance	3990.57*	2268	.955	.060 [.057, .063]	.045
Metric invariance	5875.03*	3108	.928	.065 [.062, .067]	.081
Scalar invariance	9162.20*	3458	.851	.088 [.086, .090]	.102
Partial metric invariance	5834.54*	3106	.929	.064 [.062, .067]	.080
Partial scalar invariance	7047.52*	3402	.905	.071 [.068, .073]	.084

Note. df = degree of freedom; CFA = confirmatory factor analysis; CI = confiedence interval; ICM = independent-cluster model; CFI = comparative fit index; RMSEA = root mean square of approximation; SRMR = standardized root mean square residual. Satorra-Bentler χ^2, *$p < .001$.

These findings support the validity of the general index of the MHC-SF and the discriminant validity of its individual subscales, particularly, EWB and SWB.

3.2 | Measurement invariance analyses

We proceeded by investigating the measurement invariance of the MHC-SF based on the full bifactor model for 36 countries (excluding Kenya and Iran). We failed to achieve convergence of an unrestricted configural invariance model. To rule out inadmissible solutions, we introduced 20 inequality constraints restricting residual variances of observed variables to positive values. This allowed to obtain convergence of the configural invariance model, which showed good fit. The metric and scalar invariance models converged successfully without any constraints, suggesting that the model identification issues were due to a combination of model complexity and modest size of individual samples.

The fit of the metric invariance model was acceptable. Using Bonferroni correction, we established critical chi-square values to detect loading and intercept noninvariance ($\Delta\chi^2$ = 16.46 based on *N* = 1008 for loadings and $\Delta\chi^2$ = 15.15 based on *N* = 504 for intercepts). We proceeded by searching for noninvariant loadings based on the metric invariance model. The complete list of noninvariant parameters is given in Supplementary Information 1.

Only one loading revealed a strong noninvariance ($\Delta\chi^2$ = 39.57), the loading of item 12 on the general well-being factor in Algeria, which was negative (λ = −.22). To prevent a negative group factor variance in Algeria, we also relaxed the constraint for this loading on the psychological well-being factor. The remaining four modification indices were marginal ($\Delta\chi^2$ = 19 or below) and, when they were all addressed, the fit of the model did not change substantially (ΔCFI ≤ .001, ΔRMSEA ≤ .001), so we opted against including them into the model for the sake of theoretical parsimony.

The noninvariance of intercepts was more pronounced. The fit indices of the model with full scalar invariance were well outside the acceptable range. Based on modification indices, we relaxed the equality constraints for 54 noninvariant intercepts (listed in Supplementary Information 1). Although the target difference in practical fit indices was only reached for the RMSEA, but not for the CFI (ΔCFI = .024, ΔRMSEA = .007), the remaining modification indices were all below the cutoff and exhibited no pronounced outliers.

The items tapping into social well-being turned out to be the most problematic, with 30 noninvariant intercepts (55.6%). The psychological well-being items were less biased, with 15 noninvariant intercepts (27.8%). Finally, emotional well-being items revealed only nine instances of intercept bias (16.7%), mainly confined to item 3 ("satisfied with life," *N* = 6). After all the relevant constraints were relaxed in the model, the resulting partial scalar invariance model showed acceptable fit. There were no noninvariant intercepts in 10 countries (Chile, Colombia, India, Kazakhstan, Nepal, Pakistan, Portugal, South Africa, United Kingdom, and Vietnam), suggesting that MHC-SF data across these countries can be considered scalar-invariant.

The parameters of the model were within acceptable ranges in all groups. The model-based estimates of ω ranged from .84 to .96 (*M* = .93, *SD* = .02) for the general index and from .67 to .93 (*M* = .84, *SD* = .05) for the subscales. The

ω_H values ranged from .73 to .85 ($M = .82$, $SD = .02$). The ω_S values ranged from .11 to .15 ($M = .14$, $SD = .01$) for hedonic well-being, from .28 to .35 ($M = .32$, $SD = .01$) for social well-being, and from .14 to .19 ($M = .15$, $SD = .01$) for psychological well-being. The ECV based on the multigroup model was .72. These bifactor structure estimates based on the multigroup model were consistent with the results of single-group analyses.

In our sample of countries, metric and scalar invariance were partially supported. The comparison of practical fit indices between the nested models indicates that the noninvariance of loadings is much less pronounced, compared to the noninvariance of intercepts. We found only one strongly noninvariant factor loading (i.e., item 12 in Algeria), suggesting that metric invariance can be assumed for all the other countries. These findings indicate that the effects found using the MHC-SF can be safely compared across countries, but the comparison of mean individual and group scores necessitates using latent factor scores based on the partial invariance model.

We used the final partial invariance model to investigate the mean scores across countries. We chose the Armenian group, whose scores were the closest to the grand mean, as the reference group, setting its latent factor variances to 1 and latent means to 0. The results are shown in Supplementary Information 2. We used a basic multilevel model to investigate the associations of observed scores with latent score estimates based on the multigroup bifactor model at the individual and group level. For the general factor, this association was very strong at the individual level ($r = .98$), but moderate at the group level ($r = .31$). Similarly, the correlations of subscale scores with estimates of group factors were moderate to strong at the individual level (.71, .68, and .62 for hedonic, social, and psychological well-being, respectively), but weak at the group level (.26, .61, and .28). These findings suggest that observed scores provide fairly good estimates of the general factor and group factors for individual-level analyses, but country-level analyses may be biased, unless the noninvariance of intercepts is accounted for.

To find out the possibility that the mode of administration could contribute to measurement noninvariance, we conducted measurement invariance analyses across the mode of administration. The differences in practical fit indices were well below the thresholds (ΔCFI < .003, ΔRMSEA < .002), supporting scalar invariance. This suggests the absence of effects of mode of administration independent of those of culture and language.

3.3 | ESEM analyses

The results of single-group ESEM analyses are given in Supplementary Information. We failed to obtain convergence of the three-factor ESEM model in two samples, and the bifactor ESEM model failed to converge in five other samples. Predictably, the fit of the 3-factor ESEM model was generally better, compared to that of the ICM-CFA model (scaled chi-square difference test significant at $p < .05$ in 34 out of 36 samples). However, the difference in the change in practical fit indices showed a great variability, with ΔCFI ranging from −.090 to .107 ($M = .032$, $SD = .038$), and ΔRMSEA ranging from −.033 to .042 ($M = -.006$ to $SD = .016$) across the samples.

Out of the 33 samples where the bifactor ESEM model converged, its fit, compared to that of the bifactor model, was only significantly better in 21 samples, based on the scaled chi-square difference test ($p < .05$). The change in practical fit indices was quite marginal, with ΔCFI ranging from −.097 to .087 ($M = .015$, $SD = .031$), ΔRMSEA ranging from −.060 to .054 ($M = -.004$, $SD = .022$). These findings suggest instability of the ESEM model.

We failed to obtain convergence of the multigroup bifactor ESEM model in the 36-country sample (excluding Kenya and Iran). The three-factor ESEM model converged only after five countries were removed, which contributed to negative residual variances (Hungary, India, Colombia, Hong Kong, and Pakistan). The fit indices of the three-factor metric invariance model were comparable to those of the bifactor metric invariance model, $\chi^2(2602) = 4896.49$, CFI = .928, RMSEA = .063 (90% confidence interval [CI] [.061, .066]), SRMR = .073. However, most of the cross-loadings were weak (below .20), the only exception being the cross-loading of item 4 ("that you had something important to contribute to society") on the psychological well-being factor (in the .30–.40 range). The factor intercorrelations remained strong, ranging from .44 to .86 across the samples, with mean correlation of emotional and psychological well-being $r = .69$, and those of social well-being .57 and .62 with emotional and psychological well-being, respectively, suggesting a strong common construct.

4 | DISCUSSION

The current study aimed to examine the measurement invariance of the MHC-SF across 38 countries. It is the first attempt to establish metric invariance for the MHC-SF in a broad group of countries. Additionally, we examine whether the proposed bifactor structure of the MHC-SF is cross-culturally replicable and whether it better represents the factor structure of this scale compared to other competitive models, especially the three-factor model proposed by Keyes (1998) and replicated in some cultural context (e.g., Joshanloo et al., 2013; Jovanović, 2015; Young-Jin, 2014), as well as three-factor and bifactor ESEM models.

We believe the findings reported in this study to have both theoretical and applied significance. From the theoretical perspective, a bifactor model allows the examination of the extent to which specific (group) factors are independent from the general factor and therefore may have a differential association with other mental health predictors, correlates or outcomes. At the same time, the bifactor approach also supports the validity of Keyes's (1998) broad model of mental health as comprised by three strictly related components (i.e., emotional, psychological, and social). Regarding the applied perspective, it is useful to know whether the MHC-SF could be used as a screening test measuring mental health in different cultural contexts. Given that nowadays many young people study and work in different countries, it is necessary to have a valid instrument to assess their mental health across countries. Finally, in cross-cultural studies, the issue of measurement invariance is crucial for evaluating the possibility of generalizing findings across cultural contexts and comparing the levels of mental health across populations.

Our study has shown that a bifactor model provides a better approximation of the factor structure of the MHC-SF than do alternative models, including a three-factor solution. The model identification difficulties we encountered are to be expected in small samples, given that bifactor models include a large number of parameters and some factor loadings are expected to be low (due to partitioning of the variance between two sets of factors). The correspondence between findings obtained using constrained full bifactor model and unconstrained incomplete bifactor model indicates that inequality constraints, although theoretically debatable, turn out to be viable approach in these situations.

We have found that the bifactor model showed a good fit to the data in nearly all countries (with the exception of Kenya and Iran). More in-depth analyses revealed substantial differences in terms of common and specific variance captured by different MHC-SF subscales. In short, emotional and social well-being subscales capture a more substantial proportion of specific variance, whereas the variance captured by the psychological well-being factor largely overlaps with that of the general factor. These findings indicate that the effects obtained for the psychological well-being subscale are most likely to be very similar to those obtained for the total score and using this subscale on its own may be the best choice when a shorter instrument is needed.

We also found differences across countries in the extent of common variance captured by the general factor. Although the countries comprising our sample are mostly collectivist (with the exception of Germany and the United Kingdom: see Hofstede, Hofstede, & Minkov, 2010), we found that in countries with higher collectivism the general factor tended to be stronger, to the point of making subscale scores redundant. There are two potential explanations. First, because of interdependent self-construal present in collectivistic societies, individual and social well-being could be less distinct domains of subjective well-being (Cross, Bacon, & Morris, 2000; Singelis, 1994). Second, because there are no reverse-scored items, the effects of acquiescence, which are stronger in collectivistic contexts (Harzing, 2006), may have contributed to the common factor variance.

We also investigated differences in the invariance of items belonging to different well-being domains. The data supported metric invariance of the MHC-SF in all countries except Algeria, where partial metric invariance was found, as well as full scalar invariance in 10 countries and partial scalar invariance in 26 countries. Although the target ΔCFI was not reached, recent studies suggest that more lenient cutoff criteria are optimal when the number of groups is large (Rutkowski, & Svetina, 2014) and models based on a more realistic approximate invariance assumption may perform better in these conditions (Kim, Cao, Wang, & Nguyen, 2017).

The emotional well-being subscale emerged as the most universal in terms of item invariance, whereas the social well-being subscale turned out to be the most problematic. There are two possible explanations for these findings. First, the items measuring emotional well-being have simple content and their translations are less likely to be biased, compared to those of more complex social and psychological well-being items. Second, cultural differences may play a role: While emotions appear to be universal (Frijda, 2016), social context is strongly culturally diverse because it is conditioned by the type of interpersonal relations in society (e.g., collectivism-individualism, power distance), quality of social environment (as measured by functioning of democracy or number of crimes), and social beliefs (such as interpersonal trust or societal cynicism).

We found the application of ESEM models to MHC-SF to be problematic, for several reasons. The instability of ESEM models we observed can be explained by their complexity and by the presence of an underlying common construct. Given the presence of a common construct, a bifactor model ESEM model would be more appropriate. However, its nonconvergence is not surprising given the similarity of bifactor ESEM models to multitrait–multimethod models, in which this issue is well-known (Marsh & Bailey, 1991).

Also, in our analyses, we found that the factor correlations based on correlated-factor ESEM models were not much lower than those obtained using a conventional ICM-CFA model, the difference in the fit indices between the correlated-factor ICM-CFA and ESEM models was minor, and most cross-loadings (except for the cross-loading of item 4) were quite weak. Taken together, these findings suggest that indicator cross-loadings do not pose a serious issue in the case of MHC-SF and that bifactor CFA model is a more optimal choice for this instrument.

In general, our analyses indicate that the MHC-SF is invariant at the metric level across university student samples from most countries and is partially invariant at the scalar level. Therefore, research findings on correlates, predictors, and consequences of mental health measured by MHC-SF could be regarded as cross-culturally comparable among university students, but the bias needs to be addressed whenever a comparison of mean scores is to be performed. Because the participants in our study were sampled from 38 different countries with different cultural traditions and sociopolitical situations, our findings concerning the metric invariance suggest that MHC-SF can be used with confidence for the assessment and promotion of mental health in university student samples around the globe. This finding has applied importance, given the internationalization of universities at both undergraduate and postgraduate level because it suggests that health promotion campaigns encompassing emotional, psychological and social well-being developed for home students may translate well for international students.

4.1 | Limitations

An important limitation of our findings is the inclusion of convenience samples, made up of students, which reduces the level of representativeness. Therefore, future research should prioritize the study of the validity of the MHC-SF in larger and more heterogeneous samples, accounting for individual differences in age, sex, socioeconomic status, educational level, and exposure to stressful life events. Our study focuses only on measurement issues and does not include the validity criteria (i.e., correlations of three MHC-SF factors with objective or observational data or other established indicators of good vs. poor emotional and psycho-social functioning), which could be a next step.

Other limitations of the present study include the potential systematic effects of language (some languages are represented by samples from more than one country, i.e., English, Portuguese, Russian, and Spanish) and data collection method (paper-based or online survey). Although we found no uniform effects of the mode of administration on invariance across countries, the effects of language and mode of administration could potentially interact with those of culture. Specially designed future studies using parallel samples of respondents from the same cultures filling out the questionnaire in different languages and using different modes of administration could separate these effects reliably.

5 | CONCLUSION

The MHC-SF scores were found to be reliable and valid for comparative cross-cultural research. Our results are congruent with those obtained in earlier studies (de Bruin & du Plessis, 2015; Joshanloo et al., 2013; Jovanović, 2015; Keyes, 1998; Young-Jin, 2014) and support the cross-cultural utility of the MHC-SF and its scoring procedures. Moreover, our project extends the previous findings to countries from different cultural regions (like Asia, e.g., Nepal, Vietnam, and Korea; Africa, e.g., RSA or Kenya; and Latin America, e.g., Brazil, Chile, and Puerto Rico).

Despite the fact that the findings of our study suggest that the bifactor model is optimal across different countries, we recommend researchers investigate the internal structure of the MHC-SF in each new country and determine the best-fitting solution. More specifically, in collectivist countries the general score of the MHC may be the most informative and bifactor structure may be unstable due to low variance of domain-specific factors. Therefore, using the MHC-SF for categorical diagnosis in collectivistic countries should be done with caution because this diagnosis is based on the distinction between hedonic (emotional) and eudaimonic (psychosocial) functioning.

Our findings suggest that the differentiation between the indicators of emotional, social, and psychological well-being is not very strongly pronounced. The emotional and social well-being scales capture somewhat more specific phenomena, whereas the psychological well-being subscale is the closest to the general construct (i.e., it has little to no unique contribution to the general index of mental health in many countries). The psychological well-being subscale, which has sufficient reliability for research purposes, could be used on its own when a brief indicator of general positive functioning is required.

The results concerning the measurement invariance indicate that although effects can be safely compared across cultures, the comparison of mean scores between countries may be biased by noninvariant intercepts. This bias is more likely to be associated with the social well-being items and is less likely to be pertinent to the emotional well-being items. In cases when a cross-cultural comparison of raw scores is necessary, we recommend researchers to investigate the level of score comparability in their samples.

We have found scalar invariance only for limited selection of countries, namely, Chile, Colombia, India, Kazakhstan, Nepal, Pakistan, Portugal, South Africa, United Kingdom, and Vietnam. It means that cross-cultural comparisons in the levels of scores could be made only within this group. Most of these countries are collectivistic. Therefore it is possible to examine a distribution of different categories of mental health and search for cultural factors associated to the differences in the levels of mental health in terms of risk factors and protective factors.

In sum, we believe that our study provides initial evidence showing that the MHC-SF demonstrates good psychometric properties in student samples from 38 different countries. This empirical evidence of its structural validity and reliability can contribute to the progress in the study of mental health in cross-cultural perspective.

ORCID

Magdalena Żemojtel-Piotrowska http://orcid.org/0000-0002-8017-8014

REFERENCES

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychology*, *55*, 469–480.

Berg, M., & Venhoveen, R. (2010). Income inequality and happiness in 119 nations. In search for an optimum that does not appear to exist. In BentGreve (Ed.), '*Social policy and happiness in Europe*' (pp, 174–194). Cheltenham, UK: Edgar Elgar.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*(2nd ed). New York: Guilford Press.

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., & Kirby, J. B. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods & Research*, *29*(4), 468–508.

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

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*, 189–225.

Chetveryk-Burchak, A., & Nosenko, E. (2014). On assessing emotional intelligence as a precursor of positive personality functioning and mental health. *Advances in Social Science Research Journal*, *1*, 33–42. https://doi.org/10.14738/assrj.14.248

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, *78*, 791–808. https://doi.org/10.1037/022-3514.78.4.791

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. https://doi.org/10.1023/A:1024704320683

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55–75.

de Bruin, G. P., & du Plessis, G. A. (2015). Bifactor analysis of the mental health continuum-short form (MHC-SF). *Psychological Report*, *116*, 438–446. https://doi.org/10.2466/03.02.PR0.116k20w6

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*, 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Donahy, M. J., Lewis, C. A., Schumaker, J. F., Akuomah-Boateng, R., Duze, M. C., & Sibiya, T. E. (1998). A cross-cultural analysis of religion and life satisfaction. *Mental Health, Religion, & Culture*, *1*, 37–43.

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, *65*(2), 241–261.

Eisenberg, D., Gollust, S. E., Golberstein, E., & Hefner, J. L. (2007). Prevalence and correlates of depression, anxiety, and suicidality among university students. *American Journal of Orthopsychiatry*, *77*, 534–542.

Frijda, N. (2016). The evolutionary emergence of what we call 'emotions'. *Cognition and Emotion*, *30*, 609–620.

Harzing, A. W. (2006). Response Styles in Cross-national Survey Research A 26-country Study. *International Journal of Cross Cultural Management*, *6*, 243–266.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind*(rev 3rd ed). New York: McGraw-Hill.

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.

Joshanloo, M., & Nostrabadi, M. (2009). Levels of mental health continuum and personality traits. *Social Indicators Research*, *90*, 211–224.

Joshanloo, M., Wissing, M. P., Khumalo, I. P., & Lamers, S. M. A. (2013). Measurement invariance of the mental health continuum-short form (MHC-SF) across three cultural groups. *Personality and Individual Differences*, *55*, 755–759. https://doi.org/10.1016/j.paid.2013.06.002.

Joshanloo, M., & Jovanović, V. (2017). The factor structure of the mental health continuum-short form (MHC-SF) in Serbia: an evaluation using exploratory structural equation modeling. *Journal of Mental Health*, *26*, 510–515.

Jovanović, V. (2015). Structural validity of Mental Health Continuum – Short Form: The bifactor model of emotional, social and psychological well-being. *Journal of Personality and Individual Differences*, *75*, 154–159.

Karaś, D., Cieciuch, J., & Keyes, C. L. M. (2014). The Polish Adaptation of the Mental Health Continuum–Short Form (MHC–SF). *Personality and Individual Differences*, *69*, 104–109.

Keyes, C. L. M. (1998). Social well-being. *Social Psychology Quarterly*, *61*, 121–140.

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, *43*, 207–222.

Keyes, C. L. M. (2013). Atlanta: Brief description of the mental health continuum short form (MHC–SF). Retrieved from https://www.sociology.emory.edu/ckeyes/

Keyes, C. L. M., Wissing, M., Potgieter, J., Temane, M., Kruger, A., & van Rooy, S. (2008). Evaluation of the mental health continuum-short form (MHC-SF) in Setswana-speaking South Africans. *Clinical Psychology and Psychotherapy*, *15*, 181–192. https://doi.org/10.1002/cpp.572

Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: A comparison of five approaches. Structural Equation Modeling: A Multidisciplinary Journal. https://doi.org/10.1080/10705511.2017.1304822

Kurré, J., Scholl, J., Bullinger, M., & Petersen-Ewert, C. (2011). Integration and health-related quality of life of undergraduate medical students with migration backgrounds–Results of a survey. *Psycho-Social-Medicine*, *8*, 7.

Marsh, H. W., & Bailey, M. (1991). Confirmatory factor analyses of multitrait-multimethod data: A comparison of alternative models. *Applied Psychological Measurement*, *15*(1), 47–70.

Mikolajczyk, R. T., Maxwell, A. E., El Ansari, W., Naydenova, V., Stock, C., Ilieva, S., … Nagyova, I. (2008). Prevalence of depressive symptoms in university students from Germany, Denmark, Poland and Bulgaria. *Social Psychiatry and Psychiatric Epidemiology*, *43*, 105–112.

Morin, A., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*, 116–139.

O'Connor Quinn, H. (2014). Bifactor models, explained common variance (ECV) and the usefulness of scores from unidimensional item response theory analyses (Unpublished master's thesis). University of North Carolina at Chapel Hill, NC.

Petrillo, G., Capone, V., Caso, D., & Keyes, C. L. M. (2015). The Mental Health Continuum-Short Form (MHC-SF) as a measure of well-being in the Italian context. *Social Indicators Research*, *121*, 291–312. https://doi.org/10.1007/s11205-014-0629-3

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

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013). Multidimensionality and structural coefficient bias in structural equation modelling a bifactor perspective. *Educational and Psychological Measurement*, *73*, 5–26.

Roberts, S., Golding, J., Towell, T., Reid, S., & Woodford, S. (2000). Mental and physical health in students: The role of economic circumstances. *British Journal of Health Psychology*, *5*, 289–297.

Różycka-Tran, J., Boski, P., & Wojciszke, B. (2015). Belief in a zero-sum game as a social axiom: A 37-nation study. *Journal of Cross-Cultural Psychology*, *46*, 525–548. https://doi.org/10.1177/0220221557226

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, *74*, 31–57. https://doi.org/10.1177/0013164413498257

Ryff, C. D. (1989). Happiness is everything, or is It? Explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57*, 1069–1081.

Sagiv, L., & Schwartz, S. H. (2000). Value priorities and subjective well-being: Direct relations and congruity effects. *European Journal of Social Psychology*, *30*, 177–198.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*, 580–591. https://doi.org/10.1177/0146167294205014

Sirgy, M. J. (2012). The psychology of quality of life. *Hedonic well-being, life satisfaction, and eudaimonia*. New York: Springer.

Stock, C., Kücük, N., Miseviciene, I., Guillén-Grima, F., Petkeviciene, J., Aguinaga-Ontoso, I., & Kramer, A. (2003). Differences in health complaints between university students from three European countries. *Preventive Medicine*, *37*, 535–543.

Vaez, M., Kristenson, M., & Laflamme, L. (2004). Perceived quality of life and self-rated health among first-year university students: A comparison with their working peers. *Social Indicators Research*, *68*, 221–234.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Welzel, C., & Inglehart, R. (2010). Agency, values and well-being: A human development model. *Social Indicators Research*, *97*, 43–63. https://doi.org/10.1007/s11205-009-9557-z

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York: Guilford Press.

Yoon, M., & Kim, E. S. (2014). A comparison of sequential and nonsequential specification searches in testing factorial invariance. *Behavior Research Methods*, *46*(4), 1199–1206.

Young-Jin, L. (2014). Psychometric characteristics of the Korean Mental Health Continuum–Short Form in an adolescent sample. *Journal of Psychoeducational Assessment*, *32*, *4*, 356–364. https://doi.org/10.1177/0734282913511431.

Żemojtel-Piotrowska, M., Piotrowski, J., Cieciuch, J., Calogero, R. M., Van Hiel, A., Argentero, P., … Wills-Herrera, E. (2015). Measurement of psychological entitlement in 28 countries. European Journal of Psychological Assessment. https://doi.org/10.1027/1015-5759/a000286

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Żemojtel-Piotrowska M, Piotrowski JP, Osin EN, et al. The mental health continuum-short form: The structure and application for cross-cultural studies–A 38 nation study. *J Clin Psychol*. 2018;74:1034–1052. https://doi.org/10.1002/jclp.22570

Appendix

MHC-SF

Please answer the following questions are about how you have been feeling during the past month. Place a check mark in the box that best represents how often you have experienced or felt the following:

During the past month, how often did you feel …	never	once or twice	about once a week	about 2 or 3 times a week	almost every day	every day
1. happy	1	2	3	4	5	6
2. interested in life	1	2	3	4	5	6
3. satisfied with life	1	2	3	4	5	6
4. that you had something important to contribute to society	1	2	3	4	5	6
5. that you belonged to a community (like a social group, or your neighborhood)	1	2	3	4	5	6
6. that our society is a good place, or is becoming a better place, for all people	1	2	3	4	5	6
7. that people are basically good	1	2	3	4	5	6
8. that the way our society works makes sense to you	1	2	3	4	5	6
9. that you liked most parts of your personality	1	2	3	4	5	6
10. good at managing the responsibilities of your daily life	1	2	3	4	5	6
11. that you had warm and trusting relationships with others	1	2	3	4	5	6
12. that you had experiences that challenged you to grow and become a better person	1	2	3	4	5	6
13. confident to think or express your own ideas and opinions	1	2	3	4	5	6
14. that your life has a sense of direction or meaning to it	1	2	3	4	5	6