

Measurement Invariance of the Brief Multidimensional Student's Life Satisfaction Scale Among Adolescents and Emerging Adults Across 23 Cultural Contexts

Amina Abubakar^{1,2}, Fons van de Vijver¹, Itziar Alonso-Arbiol³, Jia He¹, Byron Adams¹, Said Aldhafri⁴, Arzu Aydinli-Karakulak⁵, Josephine Arasa⁶, Diana Boer⁷, Ozgur Celenk¹, Radosveta Dimitrova⁸, Maria Cristina Ferreira⁹, Ronald Fischer^{10,11}, Fomba Emmanuel Mbebeb¹², María Teresa Frías¹³, Andrés Fresno¹⁴, Omri Gillath¹⁵, Charles Harb¹⁶, Penny Handani¹⁷, Given Hapunda¹⁸, Shanmukh Kamble¹⁹, Marianna Kosic²⁰, Joseph Lah Looh²¹, Lubna Mazrui²², Rafael Emilio Mendia²³, Margaret Murugami²², Mei Mason-Li²⁴, Weny Savitry Pandia¹⁷, Cristina Perdomo¹, Maja Schachner²⁵, Samantha Sim²⁶, Rosario Spencer¹⁴, Angela Suryani^{1,17}, and Ergyul Tair²⁷

Abstract

There is hardly any cross-cultural research on the measurement invariance of the Brief Multidimensional Students' Life Satisfaction Scales (BMSLSS). The current article evaluates the measurement invariance of the BMSLSS across cultural contexts. This cross-sectional study sampled 7,739 adolescents and emerging adults in 23 countries. A multi-group confirmatory factor analysis showed a good fit of configural and partial measurement weights invariance models, indicating similar patterns and strengths in factor loading for both adolescents and emerging adults across various countries. We found insufficient evidence for scalar invariance in both the adolescents' and the emerging adults' samples. A multi-level confirmatory factor analysis indicated configural invariance of the structure at country and individual level. Internal consistency, evaluated by alpha and omega coefficients per country, yielded acceptable results. The translated BMSLSS across different cultural contexts presents good psychometric characteristics similar to what has been reported in the original scale, though scalar invariance remains problematic. Our results indicate that the BMSLSS forms a brief measure of life satisfaction, which has accrued substantial evidence of construct validity, thus suitable for use in cross-cultural surveys with adolescents and emerging adults, although evaluation of degree of invariance must be carried out to ensure its suitability for mean comparisons.

Keywords

confirmatory factor analysis, factor structure, BMLSS, cross-cultural, invariance

Life satisfaction forms an important component of subjective well-being among adolescents and emerging adults, as it has been associated with a host of health, educational, and behavioral outcomes (Bussing et al., 2009; Haranin, Huebner, & Suldo, 2007; Proctor, Linley, & Maltby, 2009). In assessing life satisfaction among students, there are various measures currently available, one of the most popular being Huebner's Multidimensional Life Satisfaction Scale (MSLSS; Huebner & Gilman, 2002). This scale has 40 items assessing satisfaction with life in general, and satisfaction in five domains perceived to be salient for adolescents: self, family, friends, living environment, and school. Since its introduction, the scale has been used in various contexts to assess outcome in adolescents and emerging adults (Galindez & Casas, 2011; Greenspoon & Saklofske, 1997; Huebner, Laughlin, Ash, & Gilman, 1998; Weber, Ruch, & Huebner, 2013).

To deal with the often considerable time and resource constraints in large-scale surveys, Huebner has developed a brief version of the MSLSS. This BMSLSS contains six items (BMSLSS; Huebner, Seligson, Valois, & Suldo, 2006). Five items assess satisfaction in each of the domains previously mentioned, while the sixth evaluates global life satisfaction (Seligson, Huebner, & Valois, 2003). The scarce data available indicate that the BMSLSS has good psychometric properties (Huebner, Suldo, Valois, Drane, & Zullig, 2004). For instance, Zullig, Huebner, Gilman, Patton, and Murray (2005) evaluated its psychometric properties in the United States and reported that the measure has good internal consistency, construct, and criterion validity, and shows adequate discriminant validity. Similar results in studies involving both adolescents and emerging adults have been reported from other parts of the world, such as Serbia (Jovanovic & Zuljevic, 2013) and Turkey (Civitci, 2007; Siyez & Kaya, 2008).

¹Tilburg University, Tilburg, The Netherlands

²Utrecht University, Utrecht, The Netherlands

³University of the Basque Country UPV/EHU, Spain

⁴Sultan Qaboot University, Muscat, Oman

⁵Bahcesehir University, Istanbul, Turkey

⁶United States International University, Nairobi, Kenya

⁷University Koblenz-Landau, Koblenz, Germany

⁸Stockholm University, Stockholm, Sweden

⁹Salgado de Oliveira University, Rio de Janeiro, Brazil

¹⁰Victoria University of Wellington, Wellington, New Zealand

¹¹Aarhus Institute of Advanced Studies, Aarhus University

¹²University of Dschang, Dschang, Cameroon

¹³Universidad Nacional Autónoma de Mexico, Mexico

¹⁴University of Talca, Talca, Chile

¹⁵University of Kansas, Lawrence, USA

¹⁶American University of Beirut, Beirut, Lebanon

¹⁷Atma Jaya Catholic University, Jakarta, Indonesia

¹⁸University of Zambia, Lusaka, Zambia

¹⁹Karnatak University, Dharwad, India

²⁰Scientific-Cultural Institute Mandala, Slovene Research Institute, Italy

²¹University of Buea, Buea, Cameroon

²²Kenyatta University, Nairobi, Kenya

²³Universidad Rafael Landívar, Guatemala City, Guatemala

²⁴Southampton Solent University, Hampshire, UK

²⁵Friedrich Schiller University Jena, Germany

²⁶Singapore Management University, Singapore

²⁷Bulgarian Academy of Sciences, Bulgaria

Corresponding Author:

Amina Abubakar, Department of Culture Studies, Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands.

Email: A.AbubakarAli@uvt.nl

Despite its potential usefulness as a brief and valid measure of salient aspects of the psychosocial functioning of adolescents and emerging adults, there seems to be a dearth of research evaluating the cross-cultural validity and invariance of this measure (the terms *invariance* and *equivalence* are used interchangeably). Two types of invariance are addressed here. The first and most commonly studied invariance examines whether the construct underlying the instrument, namely, life satisfaction, is measured in each cultural context and, if so, whether scores can be compared across cultures (Van de Vijver & Leung, 2000). The second type of invariance compares life satisfaction at individual and country levels. It is referred to as multi-level equivalence (Van de Vijver & Poortinga, 2002) or isomorphism (Chan, 1998; Fischer, 2009). In multi-level equivalence, or isomorphism, the key questions are as follows: Does life satisfaction have the same meaning at individual and country level, or does score aggregation of all individuals pertaining to a cultural group lead to incomparabilities? In both types of analysis, life satisfaction is the latent variable and the items are the indicators. For both types of invariance analysis, three levels of invariance are examined: configural invariance (all items are associated with life satisfaction in each country), metric invariance (all items are associated with life satisfaction in the same way across countries), and scalar invariance (the regression function linking the scores on an item to satisfaction scores has the same intercept in all cultures). In the individual-level analysis, all cultural groups are compared, whereas in the multi-level analysis, the first “group” is formed by the pooled individual data of all cultures and the second “group” is formed by the cultures (each country constitutes one observation; data in cells are average scores obtained in the various cultures; see Fontaine & Fischer, 2010). We examined both individual- and country-level invariance to be able to evaluate the extent to which BMSLSS can be used for cross-cultural comparisons.

To the best of our knowledge, the current study presents the first effort to investigate this measure in a cross-cultural context. We set out to examine the following:

1. Whether the BMSLSS factorial structure is invariant across contexts;
2. Whether the BMSLSS shows multi-level invariance at individual and country level.

Method

Sample and Procedures

The study was carried out in 23 countries across the world as part of a larger study on mental health and well-being in these countries. The data were collected among adolescents and emerging adults (see Table 1 for sample descriptives). In this article, the term *adolescents* is used to refer to samples recruited from high schools, whereas the term *emerging adults* refers to samples of undergraduate students recruited from universities. The emerging adult sample is analyzed on the basis of countries, whereas the adolescent sample is analyzed based on ethnic groups within countries. This was motivated by the fact that in several countries, the high school samples also had large numbers of adolescents of minority background (e.g., Moroccan-Dutch in the Netherlands). As there is evidence that scoring patterns and mean scores on scales such as those of life satisfaction may differ significantly between adolescents of majority and minority background (e.g., Alonso-Arbiol, Abubakar, & Van de Vijver, 2014), we decided to separate the samples accordingly. In most of the countries, university samples were recruited from a single university, with the exception of a few countries such as Kenya, Spain, and Cameroon. Adolescent data were collected from multiple schools ranging per country from 1 to 10 schools. However, no attempts were made to get a nationally representative sample. In each of these countries, ethical approval and informed consent were attained based on the requirements of the local institutional review boards (IRBs). Translations and back-translation approaches were used to develop the non-English versions of the questionnaire as needed.

Table 1. Sample Descriptives.

Country	Language	Sample size	Omega coefficients	Alpha
Emerging adults				
1. Bulgaria	Bulgarian	208	0.78 [0.73, 0.84]	.791
2. Brazil	Portuguese	170	0.81 [0.75, 0.85]	.807
3. Cameroon	English	494	0.72 [0.67, 0.77]	.724
4. Chile	Spanish	150	0.86 [0.81, 0.91]	.867
5. China	Chinese	139	0.87 [0.83, 0.91]	.878
6. Guatemala	Spanish	128	0.85 [0.81, 0.88]	.847
7. India	English	217	0.81 [0.74, 0.86]	.815
8. Indonesia	Bahasa	197	0.83 [0.78, 0.86]	.832
9. Kenya	English	174	0.78 [0.71, 0.83]	.782
10. Lebanon	English	170	0.83 [0.78, 0.86]	.830
11. Mexico	Spanish	135	0.77 [0.67, 0.85]	.769
12. The Netherlands	Dutch	218	0.80 [0.75, 0.86]	.791
13. New Zealand	English	217	0.83 [0.79, 0.87]	.830
14. Oman	Arabic	226	0.84 [0.80, 0.88]	.849
15. South Africa	English	475	0.76 [0.71, 0.80]	.764
16. Spain	Spanish	231	0.80 [0.73, 0.85]	.807
17. Taiwan	Chinese	214	0.81 [0.75, 0.86]	.811
18. Turkey	Turkish	118	0.87 [0.80, 0.91]	.872
19. The United Kingdom	English	144	0.77 [0.72, 0.83]	.775
20. The United States	English	166	0.87 [0.83, 0.89]	.866
21. Zambia	English	103	0.71 [0.62, 0.79]	.697
Adolescents				
22. Bulgaria Mainstream	Bulgarian	188	0.82 [0.77, 0.86]	.824
23. Bulgaria Turkish	Bulgarian	110	0.72 [0.59, 0.80]	.715
24. Bulgaria Roma	Bulgarian	104	0.94 [0.92, 0.96]	.946
25. Chile	Spanish	143	0.85 [0.79, 0.88]	.845
26. China	Chinese	142	0.90 [0.86, 0.93]	.901
27. Germany Mainstream	German	112	0.85 [0.79, 0.90]	.847
28. Germany Turkish	German	114	0.79 [0.72, 0.84]	.743
29. India	English	273	0.78 [0.74, 0.82]	.787
30. Indonesia	Bahasa	298	0.73 [0.67, 0.79]	.735
31. Italy Mainstream	Italian	134	0.85 [0.79, 0.88]	.849
32. Italy Slovenes	Italian	118	0.75 [0.64, 0.82]	.743
33. Kenya	English	172	0.72 [0.65, 0.79]	.719
34. Dutch-Moroccan	Dutch	89	0.82 [0.70, 0.89]	.821
35. Mainstream Dutch	Dutch	178	0.78 [0.70, 0.89]	.771
36. Oman	Arabic	332	0.82 [0.77, 0.84]	.820
37. South Africa	English	70	0.71 [0.55, 0.82]	.706
38. Spain	Spanish	580	0.78 [0.75, 0.81]	.787
39. Zambia	English	246	0.74 [0.69, 0.79]	.739

Measure

The BMSLSS (Huebner et al., 2006) was administered. The measure includes six items, five of which focus on specific domains (family, friends, school, self, and living environment), and one concerning global well-being. The sixth item was initially included as a validity check (i.e., to check the extent to which the total score from the five items correlates to a general well-being

question). In our analysis, this sixth item is also included in the total score, as we have observed that its inclusion enhances reliability. A sample item includes “I would describe my satisfaction with my family life as,” scored on a 7-point Likert-type scale ranging from 1 (*terrible*) to 7 (*delighted*).

Analysis

Our analysis comprised three main steps. First, we estimated a multi-group confirmatory factor analysis (MGCFA) model using Amos 18 (Arbuckle, 2009). We had very limited missing data in this scale. For adolescents, the percentage of missing data was between 0.2% and 0.8% per item, while for young adults, the percentage was between 0.0% and 0.2%. Given this very low rate of missing data, we used mean replacement based on data split by country/group. A unidimensional model including all six items was estimated. We assessed the goodness of fit for each model using various parameters, including Chi-square statistics, the Tucker–Lewis index (TLI), and the comparative fit index (CFI). The general guideline is that a non-significant chi-square reflects an acceptable fit to the data (Hu & Bentler, 1999). However, given the sensitivity of the chi-square statistic to sample size, we did not consider this in the current study. For TLI and CFI, values greater than .95 are considered to reflect an excellent fit, while values between .95 and .90 are considered indicative of an acceptable fit. The root mean square of approximation (RMSEA) is also reported, as it has been shown to be sensitive to model misspecification. Values of less than .06 are considered indicative of a good fit, while those between .06 and .08 are considered indicative of an acceptable model.

In a multi-group analysis, the change in CFI is an important indicator for evaluating the suitability of hierarchically nested models: A CFI change of less than .010 is taken to be supportive of the more restrictive model. Three levels of statistical equivalence are important (van de Schoot, Lugtig, & Hox, 2012). The first is configural equivalence, which is achieved when items in the measuring instrument show the same pattern of factor loadings within each group. The second level is metric equivalence, which indicates whether or not respondents from different groups answer to the questions in a similar manner. It requires that the factor loadings linking items and constructs are equal, which is an indicator of similarity of measurement unit (the metric of the response scale). The third level is scalar invariance, which requires equality in both factor loadings and intercepts across groups. It has been recommended that in the absence of either metric or scalar invariance, one may release invariance constraints on some of the factor loadings or intercepts to evaluate whether they have partial invariance at the respective level (Meredith, 1993). Mean score comparisons are only permissible when one achieves scalar (full or partial) invariance; when one achieves metric (full or partial) invariance, then it is only permissible to compare the relationship between variables across groups (Milfont & Fischer, 2010).

Following multi-group analysis and the observation that some items are not invariant, we carried out a next level of analysis to examine the impact of non-invariant items on country-level means and rankings. In addition, we checked for multi-level invariance or isomorphism at individual and country level. Multi-level invariance was evaluated on MPlus version 7.1 (Muthén & Muthén, 2010).

Finally, to evaluate internal consistency of the scale per country, both Cronbach’s alpha and omega coefficients with a 95% confidence interval were computed. The omega coefficients were computed using the MBESS package in R. Cronbach’s alpha has previously come under criticism as being an inadequate measure of reliability of psychological scales for various reasons (for details, see Schmitt, 1996; Starkweather, 2012). We therefore computed the omega coefficient so as to be able to estimate reliability in an alternative manner. Values of above .70 are considered acceptable when examining internal consistency (Cicchetti, 1994).

Table 2. Invariance Models and Goodness-of-Fit Indexes of the Multi-Group Analysis for Adolescents and Emerging Adults.

	χ^2 (df)	$\Delta\chi^2$ (Δ df)	RMSEA	TLI	CFI	Δ CFI
Adolescent model						
Configural	416.39 (144)	—	0.024	0.917	0.956	—
Metric	625.48 (229)	209.08 (85)	0.023	0.924	0.935	.021
Partial metric	500.04 (195)	83.03 (51)	0.021	0.931	0.950	.006
Scalar	1,481.21 (331)	865.73 (102)	0.032	0.847	0.812	.123
Partial scalar	914.24 (246)	414.20 (85)	0.028	0.880	0.891	.059
Emerging adults model						
Configural	418.36 (168)	—	0.019	0.942	0.969	—
Metric	647.66 (268)	256.30 (100)	0.018	0.945	0.953	.016
Partial metric	539.62 (228)	110.26 (60)	0.018	0.947	0.962	.007
Scalar	1,735.03 (388)	1,060.37 (120)	0.028	0.886	0.835	.118
Partial scalar	1,074.71 (328)	535.09 (100)	0.023	0.912	0.909	.053

Note. RMSEA = root mean square error of approximation; TLI = Tucker–Lewis index; CFI = comparative fit index; Δ = change in the model.

Results

Factorial Structure Among Adolescents

We tested a single factor model as originally conceptualized by the test developer. The data indicated that the model did not have a good fit at the configural level, with some of the fit indices being below acceptable standards, $\chi^2(162, N = 3043) = 547.32, p < .001, TLI = .895, CFI = .937,$ and $RMSEA = .027$ (fit indices prior to adding the correlated error). An examination of modification indices indicated the need to add one correlated error between Items 1 and 2. Following these modifications, fit indices were all within acceptable standards (see Table 2 for the full results). Having achieved configural invariance, we evaluated metric invariance. While most of the indices were acceptable, the difference in CFI was above the recommended cutoff of .01: $\Delta CFI = .021$. We therefore freed the factor loadings of three items based on modification indices (we freely estimated loadings for Items 1, 2, and 3). Having carried out these changes, the fit indices were all within acceptable ranges. Having achieved partial metric invariance, we then tested for scalar invariance. The fit indices at scalar level were all below acceptable standards. Relaxing invariance constraints for the intercepts of three of the six items (Item 2, 3, and 4) improved the fit indices. The fit indices were, however, still below the acceptable standards.

Factorial Structure Among Emerging Adults

We tested a single factor model as originally conceptualized by the test developer. The data indicated that the configural model had a good fit after adding a single correlated error term for Items 2 and 3 (see Table 2 where all the fit indices are presented). Having achieved configural invariance, we evaluated metric invariance. While most of the indices were acceptable, the difference in CFI was above the recommended cutoff: $\Delta CFI = .016$. We therefore freed the factor loadings of three items based on modification indices (we relaxed item loadings for Items 2, 4, and 5). With these changes, the fit indices were all within acceptable standards. Having achieved partial metric invariance, we then tested for scalar invariance. The fit indices for the scalar level were all below acceptable standards. We then released the intercepts of three items that reduced the fit

Table 3. Ranking by Invariant and Variant Items Adolescents.

Country	MLSF	MLSF rank	MINV	MINV rank	Diff rank
Bulgarian Roma	4.64	1	4.61	1	0
China	4.73	2	4.78	3	-1
Bulgaria Turkish	4.76	3	4.71	2	1
Bulgaria Mainstream	5.10	4	5.20	6	-2
Kenya	5.18	5	5.18	4	1
German	5.23	6	5.37	9	-3
Zambia	5.28	7	5.20	5	2
Chile	5.31	8	5.30	7	1
South Africa	5.35	9	5.33	8	1
Turkish German	5.36	10	5.47	10	0
Dutch	5.45	11	5.60	12	-1
Indonesia	5.47	12	5.50	11	1
Italians	5.58	13	5.69	14	-1
Slovene Italians	5.63	14	5.74	15	-1
India	5.67	15	5.67	13	2
Dutch-Moroccan	5.67	16	5.75	16	0
Spain	5.68	17	5.89	18	-1
Oman	5.76	18	5.87	17	1

Note. MLSF = mean life satisfaction full score; MINV = mean life satisfaction using only invariant items; diff rank = difference in ranking.

most (Item 2, 4, and 5). Having carried out these modifications, the fit indices improved; however, they were still below standards of acceptability.

Cross-Context Ranking Based on Observed Full Scalar Invariance or Partial Scalar Invariance

The lack of scalar invariance suggested that scores cannot be compared across cultures. Previous research has noted similar difficulties in attaining full scalar invariance in large-scale surveys (Byrne & Van de Vijver, 2010), and it has been suggested that comparing full and partial invariance provides some insight into the comparability of scores if full scalar invariance has not been obtained. Consequently, we computed the mean scores based on all the six items and mean scores based on only the three items showing weak invariance. At an individual level, these data were strongly and significantly correlated both for adolescents, $r(3403) = .921, p < .001$, and emerging adults, $r(4336) = .935, p < .001$. Similar results were seen at the country/cultural level, where the correlations were $r(18) = .983, p < .001$, and $r(21) = .979, p < .001$, for adolescents and emerging adults, respectively. Moreover, we also examined the extent to which the country and cultural group ranking differ based on whether or not the invariant items were used alone or used for all the items. Our findings indicate that not only are their means closely related, but the ranking is very similar, with most of the groups retaining the same rank, while others change by one or two ranks, and only one group shows a 3-point change (see Tables 3 and 4 where these ranks are presented). We also evaluated the correlations between rankings based on the invariant and non-invariant items. Again, these correlations were strong: adolescents, $r(18) = .967, p < .001$, and emerging adults, $r(21) = .981, p < .001$. Our findings suggest that the bias due to the lack of invariance at scalar level may be too small to make it practically consequential. However, researchers using the scale in large surveys will

Table 4. Ranking by Invariant and Variant Items Emerging Adults.

Country	MLSF	MLSF rank	MINV	MINV rank	Diff rank
Turkey	4.94	1	4.89	1	0
Taiwan	5.07	2	5.07	2	0
China	5.12	3	5.20	4	-1
Cameroon	5.21	4	5.25	6	-2
The United Kingdom	5.27	5	5.23	5	0
Bulgaria	5.27	6	5.17	3	3
Lebanon	5.32	7	5.37	8	-1
New Zealand	5.33	8	5.35	7	1
Brazil	5.34	9	5.42	10	-1
Chile	5.46	10	5.41	9	1
Kenya	5.46	11	5.42	12	-1
Zambia	5.49	12	5.45	13	-1
South Africa	5.50	13	5.42	11	2
The Netherlands	5.51	14	5.50	14	0
India	5.54	15	5.62	16	-1
Oman	5.61	16	5.71	17	-1
Indonesia	5.64	17	5.57	15	2
Taiwan	5.75	18	5.76	18	0
Spain	5.78	19	5.79	19	0
The United States	5.85	20	5.84	20	0
Guatemala	6.14	21	6.11	21	0

Note. MLSF = mean life satisfaction full score; MINV = mean life satisfaction using only invariant items; diff rank = difference in ranking.

need to evaluate the degree of bias, and its impact on ranking in their dataset, before comparing group- or country-level means.

Multi-Level Invariance

We addressed multi-level invariance in both the adolescents' and the emerging adults' samples. In the adolescent sample, we got evidence of configural invariance at individual and country level. The fit indices in these model were all within acceptable standards, $\chi^2(30, N = 18) = 2,161.70, p < .001, TLI = .922, CFI = .953, \text{ and } RMSEA = .040$. A similar pattern of results was observed for the emerging adult sample, where the fit indices showed acceptable values, $\chi^2(30, N = 21) = 2,708.31, p < .001, TLI = .921, CFI = .963, \text{ and } RMSEA = .040$.

Internal Consistency

Internal consistencies were evaluated using both alpha and omega coefficients. We investigated the internal consistency of the BMSLSS. The alpha values ranged from .69 to .94 and the omega values from .71 to .94 (see Table 1 for the results per country).

Discussion

The current study set out to investigate the invariance of the BMSLSS across contexts and age, as well as to evaluate whether its structural patterns at individual level can be replicated at country level. We observed some degree of invariance across contexts (including partial scalar invariance), as well as configural multi-level invariance (isomorphism), indicating that both individual and country differences

refer to life satisfaction. We achieved configural invariance, an indicator that the unidimensional model works relatively well across cultural contexts and age groups. Moreover, in all the countries, the internal consistency values were all above the acceptable cutoff values. These good psychometric results are generally in line with what has been reported in earlier studies investigating the psychometric values of BMSLSS (Funk, Huebner, & Valois, 2006; Man et al., 2014).

However, a more complicated picture arises when one asks the question: Can the BMSLSS be used in cross-country comparisons of the level of life satisfaction? Using an MGCFA model, we could not achieve full scalar invariance. Two points are noteworthy here. First, while there have been no studies of the cross-cultural invariance of BMSLSS, invariance analyses involving other life satisfaction scales such as the MSLSS and the Satisfaction with Life Scale (SWLS) have reported problems in achieving scalar invariance (Tucker, Ozer, Lyubomirsky, & Boehm, 2006; Zanon, Bardagi, Layous, & Hutz, 2013). These problems could arise from methodological issues, such as a lack of semantic equivalence of translated items, differences in response styles across cultural contexts, or differential meaning of life satisfaction items across countries. Our data do not allow for a more fine-grained evaluation of which of these aspects contributes (alone or in combination) to the comparability problem. Future studies in which mixed-method approaches are used to evaluate the BMSLSS may provide a richer understanding of the sources of bias. In addition, although the structural equation modeling (SEM) is commonly used for invariance analysis, it can be problematic, especially when dealing with comparisons involving a large number of cultural groups. As noted by Byrne and Van de Vijver (2010), “when comparisons comprise large-scale cross-cultural studies, the standard SEM strategy can be extremely problematic both statistically and substantively” (p. 107). The authors noted that problems may arise from conceptual misappropriation or from an accumulation of small and inconsequential differences in parameters. It was, therefore, advisable to use other strategies to evaluate the extent of the impact of lack of full scalar invariance. Our analysis indicates that the rank order of scores among countries and cultures does not change with the inclusion or exclusion of non-invariant items. Therefore, it can be concluded that while some items are problematic from an invariance perspective, they do not cause consequential or practical large effects on cross-cultural score comparisons. Future studies intending to use the BMSLSS for cross-cultural comparisons need to carefully evaluate invariance of the measure in their populations; if they face similar problems to ours in terms of scalar invariance, an evaluation of the extent to which these problems influence mean ranking would be advised. If the mean ranking does not differ substantially, the argument can be made that a comparison of observed means would be acceptable. However, if there are substantial differences in mean ranking, then alternative approaches to cross-cultural comparisons would be advisable.

We evaluated the extent to which the factor structure identified at the individual level can be replicated at country level and found that the structure is highly stable. The results indicate that aggregating individual-level data for country-level comparison is permissible, as the scale has the same structure and meaning both at individual and country level. These results provide important additional information on the potential utility of BMSLSS for cross-cultural comparisons.

Our results indicate that the BMSLSS can be used to study life satisfaction across cultural contexts. In the literature, there is a great need to understand the psychosocial adjustment of adolescents and emerging adults in a variety of contexts. The evaluation of the psychometric properties and cross-cultural utility of available scales contributes not only to understanding the theoretical underpinnings of these scales but also to providing researchers with useful information to guide their choice of indicators.

Limitations

Our evaluation of invariance issues is purely statistical. There may be explanations at semantic and functional equivalence level for our findings that we could not address. Future studies, where

mixed-method approaches are used, may go a long way in elucidating the sources of errors that may have contributed to a lack of scalar invariance. In addition, though the study involves large samples from various countries, their national representative cannot be taken for granted.

Conclusion

In conclusion, the BMSLSS forms a brief measure of life satisfaction, which has accrued substantial evidence of construct validity and is suitable for use in cross-cultural surveys with adolescents and emerging adults.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Alonso-Arbiol, I., Abubakar, A., & Van de Vijver, F. J. R. (2014). Parenting styles, attachment and psychological well-being of Dutch and Dutch Moroccan adolescents. In R. Dimitrova, M. Bender, & F. J. R. Van de Vijver (Eds.), *Global perspectives on well-being in immigrant families* (pp. 291-311). New York, NY: Springer.
- Arbuckle, J. (2009). *Amos 18*. Crawfordville, FL: AMOS Development.
- Bussing, A., Fischer, J., Haller, A., Heusser, P., Ostermann, T., & Matthiessen, P. F. (2009). Validation of the Brief Multidimensional Life Satisfaction Scale in patients with chronic diseases. *European Journal of Medical Research, 14*, 171-177. doi:10.1186/2047-783X-14-4-171
- Byrne, B. M., & Van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing, 10*, 107-132. doi:10.1080/15305051003637306
- Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of compositional models. *Journal of Applied Psychology, 83*, 234-246. doi:10.1037/0021-9010.83.2.234
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290. doi:10.1037/1040-3590.6.4.284
- Civitci, A. (2007). The adaptation of Multidimensional Students' Life Satisfaction Scale into Turkish: Validity and reliability studies. *Egitim Arastirmalari-Eurasian Journal of Educational Research, 7*, 51-60.
- Fischer, R. (2009). Where is culture in cross-cultural research? An outline of a multi-level research process for measuring culture as a shared meaning system. *International Journal of Cross-Cultural Management, 9*, 25-49. doi:10.1177/1470595808101154
- Fontaine, J. J. R., & Fischer, R. (2010). Data-analytical approaches for investigating isomorphism between the individual-level and cultural-level internal structure. In D. Matsumoto & F. van de Vijver (Eds.), *Cross-cultural research methods in psychology* (pp. 273-298). New York, NY: Oxford University Press.
- Funk, B., Huebner, A., & Valois, E. (2006). Reliability and validity of a Brief Life Satisfaction Scale with a high school sample. *Journal of Happiness Studies, 7*, 41-54. doi:10.1007/s10902-005-0869-7
- Galindez, E., & Casas, F. (2011). Adaptation and validation of the MSLSS of multidimensional life satisfaction with a sample of adolescents. *Revista de Psicologia Social, 26*, 309-323.
- Greenspoon, P. J., & Saklofske, D. H. (1997). Validity and reliability of the Multidimensional Students' Life Satisfaction Scale with Canadian children. *Journal of Psychoeducational Assessment, 15*, 138-155. doi:10.1177/073428299701500204
- Haranin, E. C., Huebner, E. S., & Suldo, S. M. (2007). Predictive and incremental validity of global and domain-based adolescent life satisfaction reports. *Journal of Psychoeducational Assessment, 25*, 127-138. doi:10.1177/0734282906295620

- Hu, L., & Benter, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling*, *6*, 1-55. doi:10.1080/10705519909540118
- Huebner, E. S., & Gilman, R. (2002). An introduction to the Multidimensional Students' Life Satisfaction Scale. *Social Indicators Research*, *60*, 115-122. doi:10.1023/a:1021252812882
- Huebner, E. S., Laughlin, J. E., Ash, C., & Gilman, R. (1998). Further validation of the Multidimensional Students' Life Satisfaction Scale. *Journal of Psychoeducational Assessment*, *16*, 118-134. doi:10.1177/073428299801600202
- Huebner, E. S., Seligson, J. L., Valois, R. F., & Suldo, S. M. (2006). A review of the Brief Multidimensional Students' Life Satisfaction Scale. *Social Indicators Research*, *79*, 477-484.
- Huebner, E. S., Suldo, S., Valois, R. F., Drane, J. W., & Zullig, K. (2004). Brief Multidimensional Students' Life Satisfaction Scale: Sex, race, and grade effects for a high school sample. *Psychological Reports*, *94*, 351-356.
- Jovanovic, V., & Zuljevic, D. (2013). Psychometric evaluation of the Serbian version of the Multidimensional Students' Life Satisfaction Scale. *Social Indicators Research*, *110*, 55-69. doi:10.1007/s11205-011-9916-4
- Man, Y., Lezhi, L., Yingxia, L., Ruoling, S., Shali, W., & Jingping, Z. (2014). Life satisfaction of adolescents in Hunan, China: Reliability and validity of Chinese Brief Multidimensional Students' Life Satisfaction Scale (BMSLSS). *Social Indicators Research*, *118*, 515-522. doi:10.1007/s11205-013-0438-0
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525-543. doi:10.1007/BF02294825
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3*, 111-130.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Author.
- Proctor, C. L., Linley, P. A., & Maltby, J. (2009). Youth life satisfaction: A review of the literature. *Journal of Happiness Studies*, *10*, 583-630. doi:10.1007/s10902-008-9110-9
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350-353. doi:10.1037/1040-3590.8.4.350
- Seligson, J. L., Huebner, E. S., & Valois, R. F. (2003). Preliminary validation of the Brief Multidimensional Students' Life Satisfaction Scale (BMSLSS). *Social Indicators Research*, *61*, 121-145. doi:10.1023/a:1021326822957
- Siyez, D. M., & Kaya, A. (2008). Validity and reliability of the Brief Multidimensional Students' Life Satisfaction Scale with Turkish children. *Journal of Psychoeducational Assessment*, *26*, 139-147. doi:10.1177/0734282907307802
- Starkweather, J. (2012). *Step out of the past: Stop using coefficient alpha; there are better ways to calculate reliability*. Retrieved from <http://it.unt.edu/benchmarks/issues/2012/06/rss-matters>
- Tucker, K. L., Ozer, D. J., Lyubomirsky, S., & Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research*, *78*, 341-360. doi:10.1007/s11205-006-8330-9
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*, 486-492. doi:10.1080/17405629.2012.686740
- Van de Vijver, F. J. R., & Leung, K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, *31*, 33-51. doi:10.1177/0022022100031001004
- Van de Vijver, F. J. R., & Poortinga, Y. H. (2002). Structural equivalence in multilevel research. *Journal of Cross-Cultural Psychology*, *33*, 141-156. doi:10.1177/0022022102033002002
- Weber, M., Ruch, W., & Huebner, E. S. (2013). Adaptation and initial validation of the German version of the Students' Life Satisfaction Scale (German SLSS). *European Journal of Psychological Assessment*, *29*, 105-112. doi:10.1027/1015-5759/a000133
- Zanon, C., Bardagi, M. P., Layous, K., & Hutz, C. S. (2013). Validation of the Satisfaction With Life Scale to Brazilians: Evidences of measurement noninvariance across Brazil and US. *Social Indicators Research*, *31*, 1-11. doi:10.1007/s11205-013-0478-5
- Zullig, K. J., Huebner, E. S., Gilman, R., Patton, J. M., & Murray, K. A. (2005). Validation of the Brief Multidimensional Students' Life Satisfaction Scale among college students. *American Journal of Health Behavior*, *29*, 206-214. doi:10.5993/AJHB.29.3.2