

Measures of Children’s Subjective Well-Being: Analysis of the Potential for Cross-National Comparisons

Ferran Casas · Gwyther Rees

Accepted: 2 December 2014 / Published online: 19 January 2015
© Springer Science+Business Media Dordrecht 2014

Abstract The potential to make cross-national comparisons is an important aspect of the growing global interest in subjective well-being. Such comparisons offer the prospect of understanding variations in levels of well-being and the factors contributing to it which can be useful for practical and policy initiatives to improve the lives of the population. However, relatively little is known about the extent to which such comparisons are reliable and valid, particularly in relation to children’s well-being. We make use of a large-scale pilot survey undertaken in 11 countries with over 16,000 children around the age of 12 to explore this issue. We examine four different multi-item measures of children’s subjective well-being, three of which have previously been proposed in the literature. We use multi-group confirmatory factor analysis to assess the extent to which it is valid to make cross-national comparisons using these measures. Overall, our results suggest that it should be possible to compare correlations and regressions between most of the countries in our survey using each of these measures. However, cross-national comparisons of mean scores on the measures is generally not supported by our analysis.

Keywords Subjective well-being · Life satisfaction · Happiness · Children · Adolescents · PWI · SLSS · BMSLSS · International comparisons

Electronic supplementary material The online version of this article (doi:10.1007/s12187-014-9293-z) contains supplementary material, which is available to authorized users.

F. Casas
ERIDIQV, Institut de Recerca sobre Qualitat de Vida (IRQV) (Quality of Life Research Institute),
Universitat de Girona, Girona, Spain
e-mail: ferran.casas@udg.edu

G. Rees (✉)
Social Policy Research Unit, University of York, York, UK
e-mail: gwyther.rees@york.ac.uk

1 Introduction

1.1 Background

An important aspect of the growing global interest in subjective well-being is the possibility of making international comparisons of quality of life of general populations. Such comparisons have huge potential to highlight differences in quality of life, and the underlying factors associated with those differences, which will be valuable to national and international policy makers.

A substantial amount of international comparative analysis has now been published in relation to adult populations. For example, in the World Happiness Report, Helliwell et al. (2012) used data from several international surveys of adult populations to compare levels of life satisfaction and happiness (mainly using single-item measures) in a large range of countries around the globe.

However, a key question in conducting such analysis is whether it is possible and appropriate directly to compare subjective well-being ratings in different national, cultural and linguistic contexts. Concerns have been raised that these ratings are influenced by differences of culture, language, interpretation and understanding and this casts doubt on making straightforward comparisons.

For example, Uchida et al. (2004) review cross-cultural evidence published primarily in the 1990s and early 2000s and conclude that there are '*substantial cultural variations in (1) cultural meanings of happiness, (2) motivations underlying happiness, and (3) predictors of happiness*'. Of particular relevance to the current article is the issue of different meanings of happiness. In broad terms, Uchida et al. argue that the individualistic orientation of European-American cultures leads to a greater emphasis on personal achievement and positive self-concept as routes to happiness, while the more collectivist orientation of East Asian cultures puts a stronger focus on inter-relatedness and social harmony.

A second issue for cross-cultural comparative work is the possibility of different response patterns to subjective well-being questions. A number of studies have identified differences between East Asian and European-American respondents. East Asian respondents have been found to have a tendency to give less extreme answers than European-Americans (Boehm et al. 2011; Chong et al. 2011; Hommerich and Klient 2012; Tov and Diener 2007).

This issue also extends to other comparative work. For example, Kim et al. (2012) also found evidence of systematic cultural response differences in a study of subjective well-being of European Canadians and Asian Canadians. European Canadians tended to have a more positive bias in terms of perceptions of self and others and this bias influenced self-ratings of life satisfaction. Vittersø et al. (2005) compare the results of the widely-used Satisfaction with Life Scale in samples of adults in Greenland and Norway. They find evidence of different response patterns in the two samples. In particular, adults in Greenland were more likely to choose extreme response values than those in Norway. The authors suggest that this difference may be due to greater familiarity with self-report surveys in Norway than in Greenland. In any case, this tendency potentially confounds the comparison of mean scores between the two countries.

Some authors (Veenhoven 2009; Bolle and Kemp 2009; Helliwell et al. 2012) have argued that, despite these concerns, mean ratings from population samples are sufficiently robust to support international comparative work, citing for example the evidence of high correlations between mean national scores in subjective well-being and objective indicators (economic and political). Others (e.g. Kapteyn et al. 2009; Angelini et al. 2013) have suggested methodological approaches, such as the use of anchoring vignettes, to correct for cultural and linguistic factors.

Most of this debate has been conducted in relation to the subjective well-being of adults. There has been relatively little discussion about its relevance to children's subjective well-being. Partly, this is because of a lack of large-scale data sets containing subjective well-being ratings from a range of countries, although there are some recent examples using subjective data from the HBSC survey. Adamson (2007) used a single-item measure of children's life satisfaction (Cantril's Ladder), as part of UNICEF Report Card 7, which compared the well-being of children in 21 rich countries. More recently, Bradshaw et al. (2013) utilised eight different subjective questions from the same survey covering relationships, health, education and overall life satisfaction.

Two recent articles have begun to explore the question of cross-cultural comparability of measures of children's subjective well-being. Casas et al. (2011) analysed the performance of several single-item and multi-item measures in four Latin-language speaking countries. In relation to the multi-item measures, using structural equation modelling, they found support for factor invariance across the four countries, meaning that the measures could be used to compare country differences at least in terms of regression coefficients. This study also highlights some of the complexities and challenges involved in translating and representing the same psychometric measures of well-being in different languages. English words such as 'happiness' and 'satisfaction' can not necessarily be directly translated, and the same goes for some of the concepts covered in the scales – for example 'community'.

Casas et al. (2013) found similar results in a comparative analysis of a multi-item measure of children's subjective well-being in Algeria and Spain, supporting the comparisons of correlations and regressions between samples, but not means. They suggest that the differences in mean scores between countries may be partly attributable to different response styles, which in turn may be related to differences in culture and socialisation. The authors conclude that caution is required in making simple cross-national comparisons of mean scores for subjective well-being measures.

There is therefore an extensive literature on international comparisons of adult subjective well-being and a much more limited body of work in relation to children's subjective well-being.

1.2 Aims

The aim of this article is to contribute to this ongoing debate about the potential to compare subjective well-being across nations, cultures and languages. We make use of a newly available data set and use multi-group confirmatory factor analysis to test the extent to which several different multi-item psychometric scales aimed at measuring children's subjective well-being may be suitable for international comparative work.

To check for the convergent validity of the psychometric scales in each different language and cultural context we will use the Overall Life Satisfaction (OLS) single-item scale. According to Cummins et al. (2003) this item represents the most abstract and least deconstructed level of life satisfaction and therefore, any more deconstructed instrument should show high contribution when regressed on the OLS. It is important to acknowledge that taking this approach implies an underlying assumption that people's ratings of domain satisfaction contribute to their rating of overall life satisfaction – that is, at least to some extent, survey respondents make use of their assessments of how they feel about different aspects of their lives in order to come to a conclusion about their assessment of their life as a whole. This approach is, in fact, the subject of considerable debate within the well-being literature with some authors arguing that factors such as personality traits influence ratings of both overall life satisfaction and domain satisfaction. This debate has not been resolved, with competing evidence to support both sides of the argument. It seems likely that both approaches have some merit and that an integrative approach (e.g. Heller et al. 2004), recognising the role of personality factors and situational factors, may be appropriate.

Additionally, because some research results suggest that subjective well-being may differ according to gender in some cultural contexts, we will include gender in the final multi-group models.

2 Data

The data set used in this article is from the pilot wave of the International Survey of Children's Well-Being (ISCWeB). ISCWeB is an important new cross-national survey of children's subjective well-being and daily activities that has been developed by an international group of researchers linked to the International Society of Child Indicators. The aims of the ISCWeB project as a whole are to collect solid and representative data on children's lives and daily activities, their time use and, in particular, their own perceptions and evaluations of their well-being in order to improve children's well-being by influencing opinion leaders, decision makers, professionals and the general public, both in the project countries and internationally.

2.1 Data Sets

Following initial phases of small scale qualitative testing, the pilot wave of the survey was conducted in 14 countries between the winter of 2011 and the winter of 2012 with children in three age groups (8, 10 and 12) with three different versions of the questionnaire. Not all countries were able to cover all three age groups. For simplicity the analysis in this article uses only the data gathered from children aged around 12 years old.

A total of 11 countries undertook a pilot of the 12-year-old version of the questionnaire –Algeria (Oran County), Brazil (Rio Grande do Sul State), Chile, England, Israel, Romania (Cluj County), South Africa (Western Cape), South Korea, Spain, Uganda (Eastern Uganda) and USA (South Dakota). In total,

almost 17,000 children participated in these 11 countries. The target was to include a sample of at least 1000 children in the age group. However in practice, three countries did not achieve this target – Algeria (428 cases), USA (784) and Chile (827). Several countries achieved much larger samples – Spain (5727), South Korea (2602). In all countries, the survey was undertaken in schools.

In each country, approval was obtained from an appropriate ethics committee prior to the survey being conducted.

In ten countries paper questionnaires were used while in England the survey was administered via computer.

The original questionnaire was written in English. In countries other than England and the USA, the questionnaires were translated into other languages as required. The translation process involved an initial translation from English, a translation of the resulting questionnaire back into English and then a resolution of any resulting anomalies in wording between the original version and the back-translated version. This resolution included taking into account discussions with children where possible and the knowledge of the local research teams. As well as these measures regarding translation issues, an effort was made to ensure that the visual presentation of the questionnaire and labelling of responses (e.g. anchoring points for scales) was consistent in all countries as variations in these aspects may result in variations in response patterns.

As this was a pilot survey conducted in most cases without substantial funding and with the main purpose of assessing the viability of the survey questionnaires and survey method, sampling strategies were not specified in advance and varied from one country to another. In most countries, some form of random sampling (usually stratified cluster sampling) aimed at achieving a representative sample was utilised. However in some countries, a more pragmatic approach was taken to achieving the target sample. This factor does place some limitations on the generalizability of the analysis. However, given the scarcity of international data sets on children's subjective well-being, the ISCWeB pilot offers an opportunity to undertake preliminary analysis which can then be pursued in the future with more representative samples.

2.2 Measures

2.2.1 Huebner's SLSS

The Student Life Satisfaction Scale (SLSS) was developed by Scott Huebner in the US (Huebner 1991). The scale consists of seven items designed to refer to overall (not domain-specific) life satisfaction (example item: *'My life is going well'*) to which respondents are asked to agree or disagree. The initial version used a four-point frequency response scale, but more recently a six-point agree-disagree scale has been recommended by Huebner. The scale has been shown to have good reliability and validity in general samples of youth in the US (see Huebner and Hills 2013 for a recent summary) and has also been used in a number of other countries including Portugal (Marques et al. 2011) and Hong Kong (Park and Huebner 2005).

In the ISCWeB survey, to preserve consistency with other statement-based items in the questionnaire, a five-point response scale from ‘Strongly disagree’ to ‘Very much agree’ was used. The seven items were included in questionnaire in all countries except Spain where only five items were used.

2.2.2 Huebner’s BMSLSS (Modified Version)

Huebner and colleagues have also developed the Brief Multidimensional Student Life Satisfaction Scale (BMSLSS) which contains single-item measures of five key domains in children’s lives – family, friends, school, self and living environment (Seligson et al. 2003). The scale has subsequently been tested and validated in the US with children and young people aged 8 to 18 (Huebner et al. 2006, 2011). Response options for the original version are on a seven-point scale from ‘Terrible’ to ‘Delighted’ originally suggested by Andrews & Withey. Outside the USA, the BMSLSS has been utilised in studies in Turkey (Siyez and Kaya 2008) and China (Kwan and Ip 2009).

In the ISCWeB survey, all domain satisfaction items were measured on a scale from zero to ten where zero equals ‘Completely dissatisfied’ and ten equals ‘Completely satisfied’. Five of the domain items approximate to the domains proposed in the BMSLSS. These were ‘People I live with’, ‘Friends’, ‘The school I go to’, ‘Myself’ and ‘The area I live in’. We have used these items with the 11-point response scale as a modified version of the BMSLSS for testing purposes. We will refer to this as the Modified BMSLSS.

2.2.3 Cummins and Lau’s PWI-SC

Cummins and colleagues developed the Personal Well-Being Index (PWI) to measure the subjective well-being of adults in Australia (Cummins et al. 2003). Subsequently, Cummins also developed alternative versions for different populations – including the Personal Well-Being Index – School Children (PWI-SC) for school-aged children and adolescents (Cummins and Lau 2005). Each version of the PWI consists of seven items, with a zero to ten response scale. The PWI-SC has been validated in Australia (Tomy and Cummins 2011).

The seven items forming the PWI-SC were included in the questionnaire in ten countries (excluding Spain). The response scale was as described above. The items ask about satisfaction with material possessions, health, achievements, relationships, safety, community connectedness and future security.

2.2.4 PWI-SC5 (Alternative Version)

The ISCWeB questionnaire also included questions about satisfaction with other aspects of life. It was therefore possible also to test an alternative abbreviated version of the PWI-SC with items available for all 11 countries. The domains covered were material possessions, health, relationships, school and time use. Some of the wordings of the domains belong to the PWI (adults’ version) that has been validated with adolescents in studies in Spain and Romania (Casas et al. 2009). The last item has been suggested as important in research on children’s subjective well-being (Casas et al. 2009). We will refer to this as the Alternative PWI-SC5.

2.2.5 Overall Life Satisfaction

Also included in the survey questionnaire was a single-item measure of overall life satisfaction (on a scale from zero to ten as described above). We have used this variable as a dependent variable in the analysis. We will refer to this as OLS.

3 Analysis

The main statistical technique we have chosen to use to address the research aims earlier is Structural Equation Modelling (SEM), and in order to test the appropriate fit of the multi-item psychometric scales we will use SEM for Confirmatory Factor Analysis (CFA). In contrast to exploratory factor analysis, which is commonly used to attempt to identify separate factors of a particular construct, CFA can be used to assess the extent to which an existing hypothesised model fits a particular data set. Here we use this technique to check for the fit of the four proposed multi-item measures of children's well-being discussed above to the data from the ISCWeB pilot survey. In addition, using CFA, it is possible to examine the extent to which the hypothesised model fits sub-groups of the data set in a similar way. In this case we have used multi-group CFA to examine simultaneously the relative fit of the measures separately in each of the countries represented in the data set with the aim of establishing whether it is possible to use these measures for comparative cross-national analysis. In order to assess a particular CFA model, and the incremental improvements between variations of a model, various statistical indexes can be generated to inform a decision about whether the model adequately fits the data. We provide further details on the measures of model fit used in this analysis below.

3.1 Data Preparation

The data set used for this analysis had been cleaned and prepared as part of the international project. This process included identifying and excluding cases with high proportions of missing data, and identifying and excluding cases with systematic response patterns. The latter point is particularly relevant to the topic of this article because, for example, Cummins and Lau (2005) recommend that respondents who score at the top or bottom of the scale for all items in the PWI-SC should be excluded from analysis. There are problems with this approach, however, in that it automatically excludes anyone who expresses complete satisfaction with all aspects of their life covered by the instrument. Within the current data set it was possible to take a broader approach because the questionnaire included a number of different sets of items. So the approach taken was that uniform response patterns were identified for five different sets of items in the data set and cases were excluded if they exhibited such patterns for more than one of the five item sets. Thus cases excluded on this basis are not included in the analysis presented in this paper.

In preparing for the confirmatory factor analysis, a further assessment was made of instances of missing values within the data set. Two issues were identified. First, 28 cases had a missing value for gender. Since, there are often significant variations in subjective well-being by gender, it was decided to exclude these cases in order to

facilitate inclusion of gender in some of the models tested. Second, 171 cases had more than three missing values on the satisfaction items being tested and these were also excluded from the analysis. This left a data set with a total of 16,704 cases. A breakdown of cases by country and gender is shown in Table 1.

In addition, basic checks were undertaken on the level of missing data for each variable included in the analysis. It was decided to exclude from the analysis any variable with more than 5 % missing values. On this basis, one variable from Huebner's SLSS was excluded from the analysis (see further details below).

Missing values for satisfaction items were then substituted by regression as implemented in SPSS version 19.

Skewness of the variables ranged from -2.242 to -0.301 and kurtosis from -1.316 to 5.147 . These departures from normality were handled in structural equation models by means of bootstrap, as implemented in the AMOS 19 software.

Descriptive statistics for the measures used in the analysis are shown in Table 2

3.2 Analytical Methods

In order to first assess the validity of the factorial structure of multi-item scales, we tested different confirmatory factor analysis (CFA) models for each of the measures. We used the AMOS 19 software with maximum likelihood estimation and the bootstrap method to compute standard errors. The fit indices considered were TLI (Tucker & Lewis Index), CFI (Comparative Fix Index), RMSEA (Root Mean Square Error of Approximation) and SRMR (standardised Root Mean Square Residual). We assumed that results higher than 0.950 for CFI and results below 0.05 for RMSEA and SRMR are excellent, in accordance with Batista-Foguet and Coenders (2000), Arbuckle (2010) and Byrne (2010). However, it is worth bearing in mind that for some other authors RMSEA values up to 0.08 represent acceptable errors of approximation in larger samples (Browne and Cudeck 1993; Byrne 2010; Marsh et al. 2010), while CFI greater than 0.90 reflects acceptable fit to the data (Marsh et al. 2010).

Table 1 Final sample by country and gender

| Country | Females | Males | Total |
|--------------|---------|-------|--------|
| Algeria | 211 | 215 | 426 |
| Brazil | 467 | 535 | 1002 |
| Chile | 477 | 349 | 826 |
| England | 456 | 646 | 1102 |
| Israel | 495 | 488 | 983 |
| Romania | 669 | 682 | 1351 |
| South Africa | 457 | 536 | 993 |
| South Korea | 1074 | 1505 | 2579 |
| Spain | 2798 | 2841 | 5639 |
| Uganda | 523 | 510 | 1033 |
| USA | 376 | 394 | 770 |
| Total | 8003 | 8701 | 16,704 |

Table 2 Descriptive statistics for measures used in the analysis

| | Mean | Standard deviation |
|--|------|--------------------|
| My life is going well | 3.25 | 0.908 |
| My life is just right | 2.98 | 1.038 |
| I wish I had a different kind of life (Reversed) | 2.35 | 1.469 |
| I have a good life | 3.27 | 0.959 |
| I have what I want in life | 2.85 | 1.142 |
| Satisfaction with: The people you live with | 8.97 | 1.860 |
| Satisfaction with: Your friends | 8.68 | 1.943 |
| Satisfaction with: The school you go to | 8.2 | 2.363 |
| Satisfaction with: Yourself | 8.47 | 2.205 |
| Satisfaction with: The area you live in general | 8.03 | 2.574 |
| Satisfaction with: All the things you have | 8.64 | 2.110 |
| Satisfaction with: Your health | 8.77 | 1.985 |
| Satisfaction with: The things you want to be good at (*) | 8.16 | 2.227 |
| Satisfaction with: Your relationships with people in general | 8.61 | 1.870 |
| Satisfaction with: How safe you feel (*) | 8.09 | 2.306 |
| Satisfaction with: Doing things away from your home (*) | 7.77 | 2.534 |
| Satisfaction with: What may happen to you later in life (*) | 8.04 | 2.325 |
| Satisfaction with: How you use your time | 8.15 | 2.137 |
| Overall life satisfaction (OLS) | 8.5 | 2.221 |

$N=16,704$, excepting for (*) which have a $N=11,065$, because these questions were not asked in Spain

In order to compare CFA coefficients between groups (in our case, between countries), we first checked whether factor invariance is observed. Factor invariance refers to the degree that items in the questionnaire have the same meaning for members of different groups and is a prerequisite for factor comparison to make sense. If it is not observed, means or correlation coefficients may be attributed to real distribution differences or to different meanings of the variables (Meredith 1993). We will consider two kinds of factor invariance: metric factor invariance (a requisite for comparing variances, covariance or regression coefficients) and strong factor invariance (a requisite for comparing means between groups) (Coenders et al. 2005). Therefore, we will test each multigroup model in three steps.

The analytical procedure for each of the four measures being tested was as follows. First, CFA of each measure was conducted using the pooled sample to test model fit and identify any problematic items. Then, once any modifications had been made to the initial model (exclusion of items and/or inclusion of error covariance constraints), multi-group CFA was conducted to test measurement invariance across countries. This involved three steps (see Byrne 2010) testing sequentially (a) an unconstrained multi-group model (configural invariance), (b) a model with constrained unstandardised factor loadings (metric invariance), and (c) a model with constrained factor loadings and intercepts (scalar invariance). These three steps involve applying increasingly restrictive constraints to provide information about the extent to which the measures are suitable for comparative analysis. Configural invariance provides a basic test that

the model generated by the pooled sample is valid separately for each group (country). If this is so, then testing for metric invariance addresses the question as to whether the correlations and regressions associated with the measure are comparable between countries. Finally, testing for scalar invariance addresses the question as to whether the mean scores on the measure are comparable between countries.

We then proceeded to test each measure in a structural equation model including OLS and gender.

For information purposes, we also report reliability coefficients (Cronbach's alpha) for each of the measures for the pooled sample.

4 Results

4.1 Huebner's SLSS

One of the seven items in the SLSS - '*My life is better than most kids*' - had more than 5 % missing values in the pooled sample and therefore, based on the threshold agreed for the analysis, this item was excluded from the analysis process. It is interesting to note that some of our qualitative pilot testing suggested that this particular item may be problematic. Children asked for clarification regarding who they should be comparing themselves with - for example, other children they knew, children in their country in general, or children globally. It may be useful to consider more carefully the use of comparative statements such as this in research on subjective well-being.

4.1.1 Pooled Sample Analysis

Because one of the remaining six SLSS items had not been asked in Spain, the initial analysis of the pooled sample included ten countries (excluding Spain). A CFA of the initial six-item model using the pooled sample of ten countries did not fit well statistically ($\chi^2=2330.47$, $df=9$, $p<0.001$, $CFI=0.895$, $RMSEA=0.152$, $SRMR=0.071$). Adding several error covariances generated a moderate model fit, but more detailed analysis indicated that all of the error covariances were too high and suggested that better models might be obtained by deleting items. The most problematic two items were those with reverse coding - '*I would like to change things in my life*' and '*I wish I had a different kind of life*' - where agreement implies negative rather than positive subjective well-being.

A further CFA model was tested with the pooled sample for all 11 countries excluding these two items and therefore consisting of four positively worded items. This model, with the inclusion of one error covariance, shows excellent fit statistics ($\chi^2=31.32$, $df=1$, $p<0.001$, $CFI=0.999$, $RMSEA=0.042$, $SRMR=0.006$) (see Fig. 1).

Cronbach's alpha for this four item scale for the pooled sample was 0.823 indicating good statistical reliability.

4.1.2 Multi-Group Analysis

We proceeded to test a multi-group model of the reduced four-item version of the SLSS described above for the 11 countries. The multi-group model (Table 3, Model 1) also

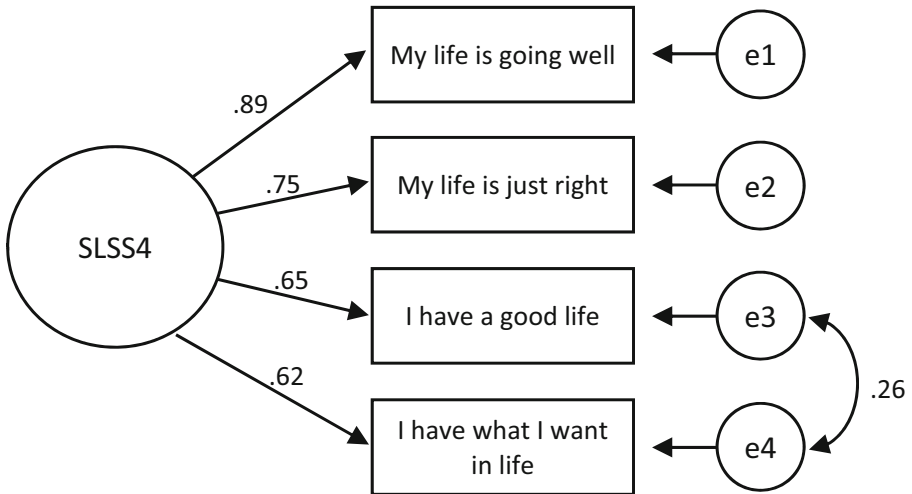


Fig. 1 CFA of the SLSS4 with the pooled sample of 11 countries. Unconstrained

showed good fit statistics without constraints. However, with constrained loadings (Table 3, Model 2) the model showed a change in the CFI substantially higher than 0.01 and was therefore not acceptable (Chen 2007; Cheung and Rensvold 2002). These results suggest that, although the SLSS4 is functioning well enough in all studied countries, we cannot compare results among the 11 countries. This may be due to a different response style among children in different countries and it could be useful to undertake further exploration of these differences using differential item functioning analysis.

Further analysis indicated that, when England and Uganda were excluded, the model (Table 3, Model 4) fitted well with constrained loadings (CFI=0.987 compared to 0.994 for unconstrained model, RMSEA=0.026). However, a model with constrained loadings and intercepts (Table 3, Model 5) did not fit adequately (CFI=0.940, RMSEA=0.043).

We can conclude from this analysis that mean values of the SSL4 are not comparable among the studied countries but correlations and regressions are comparable among nine of the 11 countries (i.e. excluding England and Uganda).

Finally, we have tested a multi-group structural equation model (SEM) including SLSS4, OLS and gender, with 9 countries. Results show that the standardised weights of the latent variable SLSS4 on OLS vary among countries from 0.389 in South Africa

Table 3 Multi-group CFA models: SLSS4

| Model | χ^2 | df | p-value | CFI | RMSEA (confidence interval) | SRMR |
|---|----------|----|---------|-------|-----------------------------|-------|
| 1 11 countries, unconstrained | 166.99 | 11 | 0.000 | 0.994 | 0.029 (0.025–0.033) | 0.003 |
| 2 11 countries, constrained factor loadings | 1221.14 | 41 | 0.000 | 0.958 | 0.042 (0.040–0.044) | 0.020 |
| 3 9 countries, unconstrained | 158.71 | 9 | 0.000 | 0.994 | 0.034 (0.029–0.039) | 0.025 |
| 4 9 countries, constrained factor loadings | 359.74 | 33 | 0.000 | 0.987 | 0.026 (0.024–0.029) | 0.026 |
| 5 9 countries, constrained factor loadings and intercepts | 1575.34 | 57 | 0.000 | 0.940 | 0.043 (0.041–0.045) | 0.030 |

to 0.756 in South Korea. Additionally, the relative contributions of different items to SLSS4 vary between countries (see Table A1 in Appendix). The highest contribution is from the item ‘My life is going well’ in all countries excepting in South Africa and South Korea, where it is ‘My life is just right’, although generally the contribution of these two items are fairly similar in most countries. The lowest contribution in all countries is from the item ‘I have what I want in life’.

Gender does not show a significant standardised loading on SLSS4 in Algeria, Chile, South Africa, South Korea, Spain or USA. However, it does in Brazil, Israel and Romania – Israel being the only country where the loading for girls was higher than for boys. Correlations between OLS and gender are significant in Brazil, Chile, South Korea and USA, and non-significant in Algeria, Israel, Romania, South Africa and Spain.

4.2 Huebner’s BMSLSS (Modified Version)

4.2.1 Pooled Sample Analysis

A CFA of the model using the five items described earlier for the pooled sample of 11 countries showed very good fit statistics ($\chi^2=49.95$, $df=5$, $p<0.001$, $CFI=0.996$, $RMSEA=0.023$, $SRMR=0.009$) (see Fig. 2), with no error covariances required.

Cronbach’s alpha for the BMSLSS items for the pooled sample was 0.706 indicating acceptable reliability.

4.2.2 Multi-Group Analysis

The unconstrained multi-group CFA for the 11 countries using the same items also showed good fit (Table 4, Model 1). However, the multi-group model with constrained

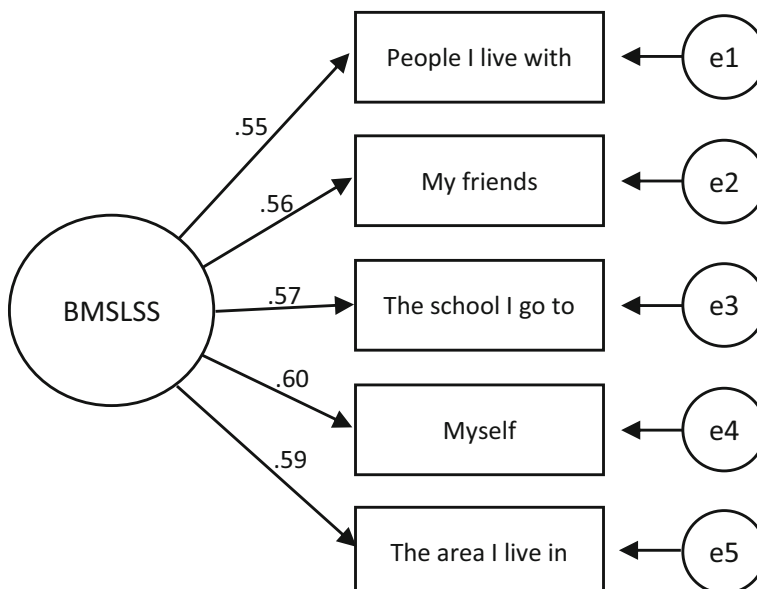


Fig. 2 CFA of the BMSLSS with the pooled sample of 11 countries. Unconstrained

Table 4 Multi-group CFA models: BMSLSS (modified version)

| Model | χ^2 | df | p-value | CFI | RMSEA (confidence interval) | SRMR |
|--|----------|----|---------|-------|--------------------------------|-------|
| 1 11 countries, unconstrained | 277.99 | 55 | 0.000 | 0.979 | 0.016 (0.014–0.017) | 0.016 |
| 2 11 countries, constrained factor loadings | 519.54 | 95 | 0.000 | 0.961 | 0.016 (0.015–0.018) | 0.058 |
| 3 7 countries, unconstrained | 136.88 | 35 | 0.000 | 0.988 | 0.015 (0.013–0.018) | 0.015 |
| 4 7 countries, constrained factor loadings | 236.96 | 59 | 0.000 | 0.979 | 0.015 (0.013–0.018) | 0.035 |
| 5 7 countries, constrained factor loadings and intercepts | 909.98 | 83 | 0.000 | 0.905 | 0.028 (0.027–0.030) | 0.043 |

loadings (Table 4, Model 2) did not fit adequately as the reduction in the CFI statistic in comparison with the unconstrained model was greater than 0.01.

The next step was to check which country results were comparable. The results of this detailed analysis suggest that correlations and regressions are comparable among seven of the countries - Algeria, Brazil, Chile, Romania, South Korea, Spain, and USA (Table 4, Model 4). However, a further multi-group model for these seven countries with constrained loadings and intercepts did not fit adequately (Table 4, Model 5) indicating that it is not appropriate to compare mean BMSLSS scores between countries.

We went on to test a multi-group SEM including BMSLSS5, OLS and gender, with the above seven countries. Results show that the standardised weights of the latent variable BMSLSS5 on OLS vary among countries from 0.690 in Algeria and Brazil to 0.871 in the USA (Table A2 in the Appendix). The highest contribution to BMSLSS5 is from the item ‘Satisfaction with myself’ in all 7 countries, with much higher standardised loadings than any other item, suggesting the items composing this scale are dissimilar in their importance in explaining the latent variable. In this context it is worth noting that, in the four countries excluded from this analysis - England, Israel, South Africa and Uganda - the item with the highest standardised loading on the BMSLSS5 is not ‘satisfaction with myself’ but ‘satisfaction with the area I live in’. Therefore there appear to be two separate clusters of countries in which the BMSLSS5 items function differently.

Amongst the group of seven countries included in Models 3 to 5 above, gender does not show a significant standardised loading on BMSLSS5 in Algeria, Brazil, Chile, Romania or South Korea. However, it does in Spain and USA. Correlations between OLS and gender are significant in all countries excepting Algeria and Romania.

4.3 Cummins' PWI-SC

Our analysis here covers ten countries - excluding Spain where not all questions in the PWI-SC were asked.

4.3.1 Pooled Sample Analysis

An initial CFA model using the referred 7 items with the pooled sample of 10 countries shows only moderate fit statistics ($\chi^2=473.34$, $df=14$, $p<0.001$, $CFI=0.980$, $RMSEA=0.054$, $SRMR=0.024$). However, including 3 error covariances in the

model, the fit statistics are improved ($\chi^2=296.95$, $df=11$, $p<0.001$, $CFI=0.991$, $RMSEA=0.040$, $SRMR=0.015$) (see Fig. 3).

Cronbach's alpha for the PWI-SC with the pooled sample of 10 countries was 0.826 indicating good reliability.

4.3.2 Multi-Group Analysis

An unconstrained multi-group version of the same model (including error covariances) also fits very well (Table 5, Model 1). However, the model shows a change in the CFI substantially higher than 0.01 with constrained loadings and is therefore not acceptable (Chen 2007; Cheung and Rensvold 2002). (Model 4 in Table 5). These results suggest that, although the PWI-SC7 is functioning well enough in all studied countries, we cannot compare results among the ten countries.

We proceeded again to check which country results were comparable. When three countries were excluded - Chile, England and Uganda - model fit with constrained factor loadings was adequate (Table 5, Models 3 to 5). However, even excluding these countries, a model with constrained factor loadings and intercepts still showed a change in the CFI substantially higher than 0.01 (Table 5, Models 6 to 8). Thus, we conclude that the mean values of the PWI-SC are not comparable among the surveyed countries but correlations and regressions are comparable among seven of the countries.

A multi-group SEM including PWI-SC, OLS and gender for seven countries (excluding Chile, England and Uganda) produced standardised weights of the latent

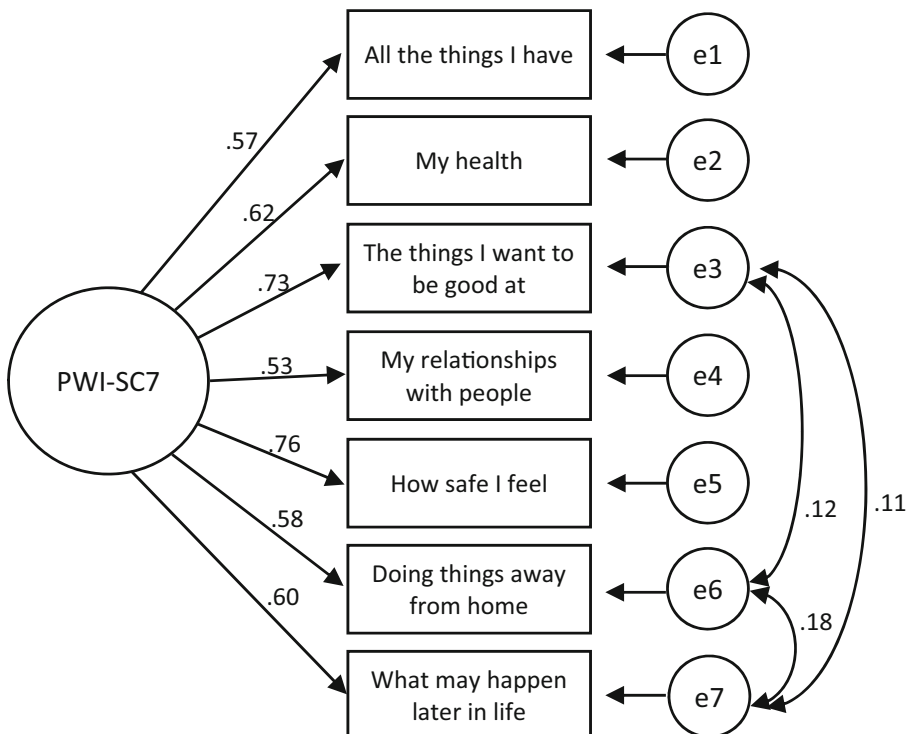


Fig. 3 CFA of the PWI-SC with the pooled sample of 10 countries. Unconstrained

Table 5 Multi-group CFA models: PWI-SC

| Model | χ^2 | df | <i>p</i> -value | CFI | RMSEA (confidence interval) | SRMR |
|--|----------|-----|-----------------|-------|--------------------------------|-------|
| 1 10 countries, unconstrained | 445.35 | 110 | 0.000 | 0.985 | 0.017 (0.015–0.018) | 0.019 |
| 2 10 countries, constrained factor loadings | 859.48 | 164 | 0.000 | 0.970 | 0.020 (0.018–0.021) | 0.056 |
| 3 7 countries, unconstrained | 331.04 | 77 | 0.000 | 0.986 | 0.020 (0.018–0.022) | 0.020 |
| 4 7 countries, constrained factor loadings | 529.16 | 113 | 0.000 | 0.976 | 0.021 (0.020–0.023) | 0.053 |
| 5 7 countries, constrained factor loadings and intercepts | 1409.66 | 149 | 0.000 | 0.928 | 0.032 (0.031–0.034) | 0.052 |

variable PWI-SC on OLS varying from 0.499 for South Africa to 0.786 for USA (Table A3 in the Appendix). Again, in addition, the different items making up the PWI-SC7 make different contributions to the latent variable in different countries. In Algeria, Brazil, South Korea and USA the highest contribution to PWI-SC7 is from the item “Satisfaction with my safety”; while it is “Satisfaction with things I am good at” in Israel, Romania and South Africa. The lowest contribution to the latent variables is “Satisfaction with my relationships in general” in Brazil, Israel, Romania and USA, while it is “Satisfaction with things I do away from home” in Algeria and South Africa and with “Satisfaction with things I have” in South Korea.

Gender showed a significant standardised loading on PWI-SC in three countries - Israel, South Korea and USA; and the correlation between gender and OLS was significant in four countries - Algeria, Israel, Romania and South Africa.

Our provisional conclusion is that PWI-SC7 is functioning well within all ten countries, but only in seven countries is it suitable for cross-national comparative analysis on the basis of correlations and regressions.

4.4 PWI-SC5 (Alternative Version)

Finally, as explained earlier, we wished to test an alternative to the PWI-SC using five items that have been found to be important components of children's subjective well-being in previous studies and which were available for all 11 countries.

4.4.1 Pooled Sample Analysis

A CFA of the five item model with the pooled sample from 11 countries showed excellent fit statistics with no error covariances ($\chi^2=49.95$, $df=5$, $p<0.001$, $CFI=0.996$, $RMSEA=0.026$, $SRMR=0.009$) (see Fig. 4).

Cronbach's alpha for this five item scale for the pooled sample was 0.714 indicating acceptable reliability.

4.4.2 Multi-Group Analysis

The same model also fit very well as an unconstrained multi-group model (Table 6, Model 1). However as with the other scales tested, the model showed a change in the

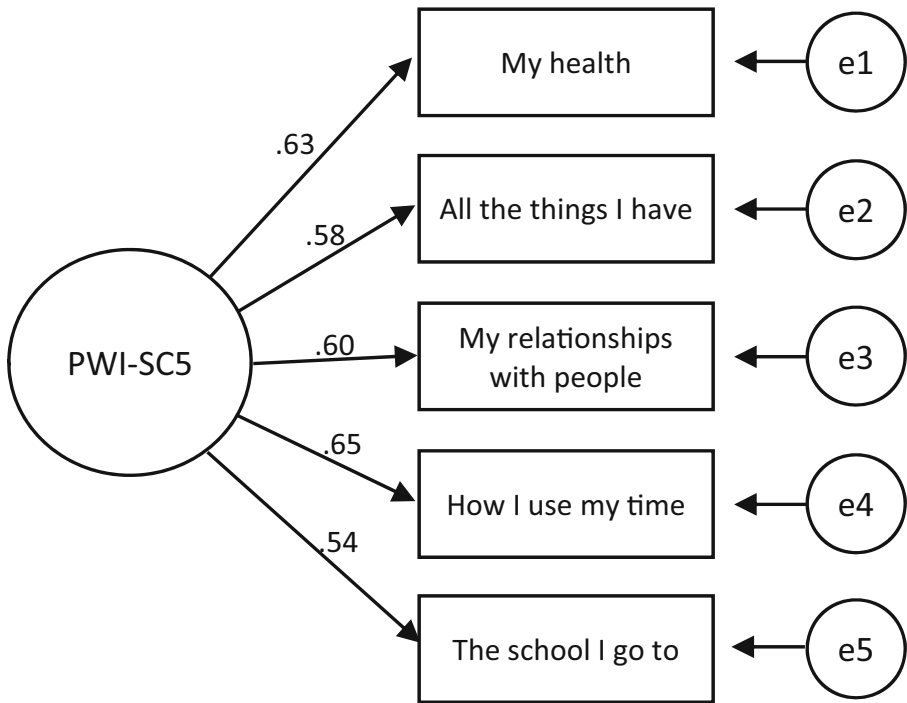


Fig. 4 CFA of the PWI-SC5 (alternative version) with the pooled sample of 11 countries. Unconstrained

CFI substantially higher than 0.01 with constrained factor loadings (Table 6, Model 2). These results suggest that, although the PWI-SC5 is functioning well enough in all studied countries, we cannot compare results among the 11 countries.

Again we checked which country results were comparable. We found that when England, South Korea and USA were excluded, the fit of the model with constrained factor loadings was acceptable (Table 6, Models 3 and 4). However, once again, a

Table 6 Multi-group CFA models: Alternative PWI-SC5

| Model | χ^2 | df | <i>p</i> -value | CFI | RMSEA (confidence interval) | SRMR |
|---|----------|----|-----------------|-------|--------------------------------|-------|
| 1 11 countries, unconstrained | 184.14 | 55 | 0.000 | 0.989 | 0.012 (0.010–0.014) | 0.023 |
| 2 11 countries, constrained factor loadings | 494.67 | 95 | 0.000 | 0.967 | 0.016 (0.015–0.017) | 0.045 |
| 3 8 countries, unconstrained | 84.76 | 40 | 0.000 | 0.994 | 0.010 (0.007–0.012) | 0.023 |
| 4 8 countries, constrained factor loadings | 183.49 | 68 | 0.000 | 0.984 | 0.012 (0.10–0.014) | 0.039 |
| 5 8 countries, constrained factor loadings and intercepts | 1848.86 | 96 | 0.000 | 0.750 | 0.039 (0.037–0.040) | 0.038 |
| 6 England and USA, unconstrained | 30.52 | 10 | 0.000 | 0.989 | 0.033 (0.020–0.047) | 0.015 |
| 7 England and USA, constrained factor loadings | 42.61 | 14 | 0.000 | 0.985 | 0.033 (0.022–0.045) | 0.020 |
| 8 England and USA, constrained factor loadings and intercepts | 57.94 | 18 | 0.000 | 0.979 | 0.034 (0.025–0.045) | 0.019 |

model with constrained factor loadings and intercepts showed again a change in the CFI and in the RMSEA substantially higher than 0.01.

We also tested whether England, South Korea and USA data are comparable. A multi-group model with the three countries does not fit. However a model using the data only from England and USA offers very good fit with constrained loadings, and even with constrained loadings and intercepts suggesting that means, correlations and regressions are comparable between these two countries (Table 6, Models 6 to 8).

Finally, we have tested a multi-group SEM including the Alternative PWI-SC5, OLS and gender, with 8 countries (Table A4 in the Appendix). The standardised weights of the latent variable Alternative PWI-SC5 on OLS range from 0.475 in Uganda to 0.725 in Chile. The highest contribution to PWI-SC5 is from the item 'Satisfaction with time use' in all 8 countries. This suggests that time use may be an important issue for children and that it could potentially make a useful addition to the PWI-SC as suggested in Casas et al. (2011). The lowest contribution to the latent variable varies depending on the country – it is 'Satisfaction with school' in Israel, Romania, and Spain, 'Satisfaction with my health' in South Africa and Uganda, 'Satisfaction with the things I have' in Algeria and Chile, and 'Satisfaction with my relationships in general' in Brazil.

Gender does not show a significant standardised loading on PWI-SC5 in Algeria, South Africa or Uganda. However, it does in the other countries. Correlations between OLS and gender are only significant in two countries - Brazil and Chile.

As our results indicate that it is possible to compare the mean scores on the PWI-SC5 for England and the USA, we analysed the difference in the mean scores for these two countries. We found that the mean value for the latent variable was 0.274 lower in England than in the USA and that this difference was statistically significant ($p < 0.01$).

5 Discussion

5.1 Conclusions

The findings we have presented on four different sets of items designed to measure children's subjective well-being are relatively consistent and several key points have emerged from the analysis.

An important initial point is that the confirmatory factor analysis suggests that, with some modifications, all measures examined here appear to offer acceptable fit for within-country analysis in all countries included in the survey.

For Huebner's Student Life Satisfaction Scale we did not find a good fit for the full seven-item measure using CFA on the pooled sample. This seemed to relate primarily to two issues – the two items that were negatively framed and therefore required reverse scoring, and the one item that asked children to compare themselves with other children. Excluding these three items, we found that a CFA with the remaining four positively-framed non-comparative items had good fit with one error covariance. These results raise questions about the suitability of the SLSS in a cross-cultural context and further research on this issue would be helpful.

The other three measures tested focused on domain satisfactions and the CFA revealed good fit and few problems for these measures. The analysis did identify some

issues, however, with one item of Huebner's BMSLSS relating to satisfaction with oneself. This item tended to have much higher factor loadings than the other items in the scale in most countries and there is a question regarding whether the concept of satisfaction with self is sufficiently distinct from satisfaction with life as a whole from children's perspective. This type of finding can contribute to the development of the Personal Well-Being Index. A key concept of the International Well-Being Group testing the index in different countries and contexts (see <http://www.deakin.edu.au/research/acqol/iwbg/>) is that the PWI will evolve as result of new theoretical developments and new evidence.

In general, the CFA models also highlight cross-national variations in the contribution of different items to the respective latent variable. This issue requires further exploration and it may be that further improvements could be made to item wordings to enable the measures to be more accurately comparable between countries. Our analysis suggests that each of the measures is also suitable for cross-national comparative analysis of correlations and regressions for many, but not all, of the countries included in the sample. There were some exceptions where the measures did not appear to be comparable and this may be due to different answering styles and/or different understandings of the wordings of the items by children in different cultures or countries. For example, the previous research cited in the introduction suggests that respondents in different cultures may have different response tendencies in terms of likelihood of selecting extreme response options. It is possible that, for children, variations in prior experience of self-completion surveys and of particular response scales may also contribute to different response patterns. Difficulties in translating particular wordings may be a factor which leads to different understandings of questions across nations and cultures. Further research might helpfully explore these issues and the extent to which they contribute to the observed differences in the functioning of the measures.

With only one exception, our results suggest that it is not appropriate to compare mean scores on the four measures between countries. The one exception is the PWI-SC5 scale for which it may be acceptable to compare means between England and USA only. Overall, this is an important finding and highlights the importance of exercising caution in making comparisons of mean levels of subjective well-being among children in different countries.

In the analysis of each measure we also present structural equation models exploring the associations between the measure, overall life satisfaction and gender. These models show a range of strengths of association between the latent variables and life satisfaction both between measures and between countries. This suggests that the items contained within the measures capture different degrees of the overall concept of life satisfaction in different cultural contexts. Gender has shown significant standardised loadings on the latent variables only in some countries and only with some of the scales here used. Something similar is observed with the correlation of gender with OLS which is only significant in some countries, but not in the same way when the psychometric scale included in the model changes. These results suggest that different scales are more or less sensitive to gender differences in subjective well-being depending on the cultural context in which data are collected and that conclusions on gender relationships with subjective well-being variables may differ when using different scales.

Additionally, the analysis indicates that, for each of the four measures tested, there are differences in the extent to which individual items make a contribution to the latent

variable in different countries. Particularly in relation to the domain-based measures, this suggests that different aspects of life may have a differential importance for children's subjective well-being in different national and cultural contexts.

The above two points highlight the need for further cross-cultural research to identify the most appropriate sets of domains for inclusion in measures for comparative research. The models also reveal differences in terms of the extent of gender variations in patterns of well-being in different countries which also warrants further exploration.

5.2 Limitations

In terms of the data set several limitations should be noted. First, the data has been taken from a pilot study which, although large-scale, did not have a uniform sampling strategy in all countries. Some countries aimed for representative samples of children through stratified random sampling, whereas other countries adopted purposive sampling techniques. Second, in several countries, the survey was only undertaken in a specific state or region. Third, because the survey was school-based in all countries it naturally excludes those children not attending a mainstream school. Given these considerations, while the analysis can highlight cross-national similarities or differences, the findings for specific countries should not be taken to be strictly representative of the child population in those countries.

Additionally, the sample sizes achieved were quite variable and in several countries did not reach the minimum target of 1000 children. This might suggest some caution in interpreting findings for these specific countries.

Finally, the data set analysed here only relates to children around 12 years of age and similar further research is needed with a wider age range.

Despite these limitations the findings in this paper extend the limited evidence base on cross-national, cross-cultural and cross-language comparisons of children's subjective well-being and highlight important areas for future research and analysis. The most important conclusions to be drawn from our analysis are that simple comparisons of mean scores for children's subjective well-being between countries may not be reliable, but that it should generally be acceptable to use subjective well-being measures to compare correlations and regressions between countries. This second point offers considerable scope for international comparative work on children's subjective well-being.

References

- Adamson, P. (Ed.) (2007). *Child poverty in perspective: an overview of child well-being in rich countries*. Report Card 7. Innocenti Research Centre. UNICEF.
- Angelini, V., Cavapozzi, D., Corazzini, L., & Paccagnella, O. (2013). Do Danes and Italians rate life satisfaction in the same way? Using vignettes to correct for individual-specific scale biases. *Oxford Bulletin of Economics and Statistics*. doi:10.1111/obes.12039.
- Arbuckle, J. L. (2010). *IBM SPSS Amos 19 User's Guide*. Chicago: Amos Development Corporation.
- Batista-Foguet, J. M., & Coenders, G. (2000). *Modelos de Ecuaciones Estructurales*. Madrid: La Muralla; Hespérides [ISBN: 84-7133-694-4].
- Boehm, J. K., Lyubomirsky, S., & Sheldon, K. M. (2011). A longitudinal experimental study comparing the effectiveness of happiness-enhancing strategies in Anglo Americans and Asian Americans. *Cognition and Emotion*, 25(7), 1263–1272. doi:10.1080/02699931.2010.541227.

- Bolle, F., & Kemp, S. (2009). Can we compare life satisfaction between nationalities? Evaluating actual and imagined situations. *Social Indicators Research*, *90*(3), 397–408. doi:10.1007/s11205-008-9265-0.
- Bradshaw, J., Martorano, L., Natali, L., & de Neubourg, C. (2013). *Children's Subjective Well-being in Rich Countries, Working Paper 2013–03*. UNICEF Office of Research, Florence. Retrieved from http://www.unicef-irc.org/publications/pdf/iwp_2013_3.pdf.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of estimating model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models*. Newbury Park: Sage Publications.
- Byrne, B. M. (2010). *Structural equation modelling with AMOS: Basic concepts, applications and programming* (2nd ed.). New York: Routledge.
- Casas, F., Sarriera, J. C., Abs, D., Coenders, G., Alfaro, J., Saforcada, E., & Tonon, G. (2011). Subjective indicators of personal well-being among adolescents. Performance and results for different scales in Latin-language speaking countries: a contribution to the international debate. *Child Indicators Research*, *5*(1), 1–28. doi:10.1007/s12187-011-9119-1.
- Casas, F., Tiliouine, H., & Figuer, C. (2013). The subjective well-being of adolescents from two different cultures: applying three versions of the PWI in Algeria and Spain. *Social Indicators Research*. doi:10.1007/s11205-012-0229-z.
- Casas, F., Baltatescu, S., González, M., & Hatos, A. (2009). *Similarities and differences in the PWI of Romanian and Spanish adolescents aged 13–16 years-old*. Conference presentation. International Society for Quality of Life Studies. Florence, July 19–23.
- Chen, F. F. (2007). Sensitivity of goodness of Fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. doi:10.1080/10705510701301834.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-Fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. doi:10.1207/S15328007SEM0902_5.
- Chong, W. C., Osman, M. M. B., Tong, E. M. W., & Tan, D. (2011). Self-construal and subjective well-being in two ethnic communities in Singapore. *Psychology*, *2*(2), 67–70.
- Coenders, G., Batista-Foguet, J. M., & Saris, W. (2005). *Temas avanzados en modelos de ecuaciones estructurales*. Madrid: La Muralla.
- Cummins, R. A., & Lau, A. L. (2005). *Personal wellbeing index—School children*. Manual (3rd ed.). Melbourne: School of Psychology, Deakin University.
- 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*(2), 159–190. doi:10.1023/A:1024704320683.
- Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: a critical examination. *Psychological Bulletin*, *130*(4), 574–600. doi:10.1037/0033-2909.130.4.574.
- Helliwell J, Layard R. & Sachs J. (2012). *World Happiness Report*. The Earth Institute / CIFAR / CEP.
- Hommerich, C., & Klient, S. (2012). Happiness: Does culture matter? *International Journal of Well-Being*, *2*(4), 292–298.
- Huebner, E. S. (1991). Initial development of the Student's life satisfaction scale. *School Psychology International*, *12*(3), 231–240. doi:10.1177/0143034391123010.
- 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., Antaramian, S. P., Hills, K. J., Lewis, A. D., & Saha, R. (2011). Stability and predictive validity of the brief multidimensional Students' life satisfaction scale. *Child Indicators Research*, *4*(1), 161–168. doi:10.1007/s12187-010-9082-2.
- Huebner, E.S. & Hills, K.J. (2013) 'Assessment of subjective well-being in adolescents'. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwab (eds.). *The Oxford Handbook of Child Psychological Assessment*. Oxford University Press.
- Kapteyn, A., Smith, J.P., & van Soest, A. (2009) *Comparing Life Satisfaction (RAND Working Paper No. WR-623-1)*. Available at SSRN: <http://ssrn.com/abstract=1289423> or doi:10.2139/ssrn.1289423.
- Kim, H., Schimmack, U., & Oishi, S. (2012). Cultural differences in self- and other-evaluations and well-being: a study of European and Asian Canadians. *Journal of Personality and Social Psychology*, *102*(4), 856–873. doi:10.1037/a0026803.
- Kwan, Y., & Ip, W. (2009). Life satisfaction, perceived health, violent and altruistic behaviour of Hong Kong Chinese adolescents: only children versus children with siblings. *Child Indicators Research*, *2*(4), 375–389. doi:10.1007/s12187-009-9041-y.
- Marques, S. C., Lopez, S. J., & Pais-Ribeiro, J. L. (2011). "Building hope for the future": a program to foster strengths in middle-school students. *Journal of Happiness Studies*, *12*(1), 139–152. doi:10.1007/s10902-009-9180-3.

- Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*(3), 471–491.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.
- Park, N., & Huebner, E. S. (2005). A cross-cultural study of the levels and correlates of life satisfaction among adolescents. *Journal of Cross-Cultural Psychology, 36*(4), 444–456. doi:10.1177/0022022105275961.
- 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*(2), 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*(2), 139–147. doi:10.1177/0734282907307802.
- Tomyn, A. J., & Cummins, R. A. (2011). The subjective wellbeing of high-school students: validating the personal wellbeing index—school children. *Social Indicators Research, 101*(3), 405–418. doi:10.1007/s11205-010-9668-6.
- Tov, W., & Diener, E. (2007). Culture and subjective well-being. In S. Kitayama & E. Cohen (Eds.), *Handbook of Cultural Psychology* (pp. 691–713). New York: Guilford.
- Uchida, Y., Norasakkunkit, V., & Kitayama, S. (2004). Cultural constructions of happiness: theory and empirical evidence. *Journal of Happiness Studies, 5*(3), 223–239. doi:10.1007/s10902-004-8785-9.
- Veenhoven, R. (2009). Comparability of happiness across nations. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-being, 211–234*.
- Vittersø, J., Biswas-Diener, R., & Diener, E. (2005). The divergent meanings of life satisfaction: item response modeling of the satisfaction with life scale in Greenland and Norway. *Social Indicators Research, 74*(2), 327–348. doi:10.1007/s11205-004-4644-7.