Longitudinal measurement invariance of the Satisfaction With Life Scale in adolescence

Igor Esnaola[1] · Manuel Benito[2] · Iratxe Antonio-Agirre[3] · Inge Axpe[2] · Margarita Lorenzo[4]

Abstract

Purpose The main purpose of this research was to examine the longitudinal measurement invariance of the Satisfaction With Life Scale (SWLS) in adolescence.

Method The sample was composed by 484 adolescents from Spain, 46.7% were males. All participants belonged to six academic levels from Grade 7 to Grade 12, and answered the questionnaires at two different times: at the beginning ($M_{age1} = 14.95$, $SD_1 = 1.81$) and at the end of the school year ($M_{age2} = 15.61$, $SD_2 = 1.81$). The reliability of the scale was obtained through Cronbach's alpha, Guttman lambda, and MacDonald's Omega total. The multiple group confirmatory factor analysis (MGCFA) was used to examine the fit of the unifactorial model to data and to test the measurement of longitudinal invariance of the scale across two time points (at the beginning T1, and the end T2, of the academic year), and the time points and groups (gender and age).

Results The values of the single-factor SWLS structure were T1 ($CFI_1 = 1.000$, $TLI_1 = .997$, $RMSEA_1 = .080$, and $SRMR_1 = .028$), and T2 ($CFI_2 = .997$, $TLI_2 = .995$, $RMSEA_2 = .032$, and $SRMR_2 = .034$). On the other hand, values of the reliability and composite reliability when analyzing both time points together as well as separately were as follows: Cronbach's alpha = .86, Guttman's lambda = .84, McDonald's Omega total = .89. Results confirmed the longitudinal invariance of SWLS. The differences in gender and age were not significant and the small differences across time points showed that the means of the latent factor remained the same over time in both variables.

Conclusion The present study confirmed the single-factor structure of the SWLS in Spanish adolescents, as well as a good reliability and composite reliability. The full longitudinal measurement invariance was also found and there were negligible differences across time points considering gender and age. If these findings are further replicated, the scale could be used to compare the life satisfaction across two time points considering different age and gender groups.

Keywords SWLS · Longitudinal measurement invariance · Adolescence · Life satisfaction

✉ Igor Esnaola
igor.esnaola@ehu.eus

Manuel Benito
manubenito@gmail.com

Iratxe Antonio-Agirre
iratxe.antonio@ehu.eus

Inge Axpe
inge.axpe@ehu.eus

Margarita Lorenzo
lorenzo.reizabal@gmail.com

[1] UPV/EHU, Donostia-San Sebastián, Spain

[2] UPV/EHU, Bilbao, Spain

[3] UPV/EHU, Vitoria-Gasteiz, Spain

[4] Centro Superior de Música del País Vasco Musikene, Donostia-San Sebastián, Spain

Introduction

Life satisfaction can be defined as an individual's overall appraisal of the quality of its own life [1], including the perception that one is progressing towards important life goals [2]. It is an important construct for people of all ages, and correlates with many different indicators of mental health. Although different instruments have been developed, the Satisfaction With Life Scale (SWLS) is one of the most frequently used to measure the global cognitive judgment of satisfaction with one's life and it has been translated to over twenty languages worldwide. Gender, age, and culture have been the most common types of measurement invariance assessed [3]; however, the longitudinal invariance has not been sufficiently studied.

The reliability coefficient (Cronbach's alpha) of the scale in different studies ranged from .80 to .92 [4–11] and test–retest reliability coefficients were found to be of .83 after 2 weeks and .84 after 1 month [12]; in another study, it was found that the SWLS also had moderate stability (correlations between three wave factors) in a Taiwanese adolescent athlete sample, with coefficients ranged from .41 (between Time 1 and Time 3, 6-month interval) to .69 (between Time 1 and Time 2, 3-month interval) [13].

Likewise, adequate construct, convergent, and discriminant validity have been demonstrated [5, 9, 10, 14]. Finally, although not all items show the same level of fit in all studies, generally speaking confirmatory factor analysis have revealed a consistent single-factor structure [4–6, 8–10, 13–16]. In addition to these basic psychometric properties, the *measurement invariance* (MI) of the SWLS across groups has been examined in recent years.

Gender, age, and culture have been the most common types of MI assessed [3] and there are four levels [17]: configural, metric, scalar, and strict invariance. There is a hierarchy in the levels of invariance [18], which means that the existence of one higher-level invariance implies the existence of invariance at lower levels. The four levels of MI are tested as a nested hierarchy and the process is carried out by increasing levels of constraints. *Configural* invariance assumes that the factor structure is the same in every group; this is to say that there are the same loads on the same factors. Additionally, configural invariance tries to determine if the subjects in the different groups employ the same conceptual framework to answer the test items [19]. Therefore, configural invariance is not enough to estimate latent mean differences. *Metric* invariance requires factor loadings to be equivalent across groups (constrain factor loadings to be equal across groups). In other words, for all items, one unit change in the item score is scaled to an equal unit change in the factor score across groups [20], as the items are measuring the same underlying construct across them; so, the construct has the same meaning across groups and in this case no latent mean differences are estimated. *Scalar* invariance requires to constrain the factor loadings and items' intercepts to be equal across groups. This type of invariance is particularly important for assessing if the mean of the latent variable is different across groups; if scalar invariance holds, then the latent mean differences can be estimated [20]. Finally, *strict* invariance requires constraining the item factor loadings, intercepts, and residual variances to be equal across groups. Strict invariance is important for group comparisons based on the sum of observed item scores because the observed variance is a combination of true score variance and residual variance; in case of strict invariance, the latent mean differences can be estimated. These restrictions mean that, under the strict invariance, the group differences

in variances of observed variables are attributable only to group differences in variances of latent variables [17].

In adolescence, measurement invariance of the SWLS supported to the configural [21] and strict [8] levels in Spain; on the other hand, measurement invariance confirmed to scalar level in Serbia [14] and rural Norwegian adolescents [10].

Measurement invariance considering *age* has been studied basically in young and adult population with inconsistent results. In the meta-analysis carried out by [3], it is said that from the seven studies that specifically tested measurement invariance using multi-group confirmatory factor analysis across age groups, one of them [16] supported strict invariance; two showed metric invariance [9, 22]; and three configural invariance [6, 7, 15]. On the other hand, in another study no invariance was found across three age groups of Swedish Teaching students [23]. Partial invariance was observed in three age measurement invariance analysis, as non-invariance was observed for loadings and/or intercepts for item one [7], two [15], four [6], and five [6]. Specifically in adolescence, measurement invariance was reported to the partial scalar level in teenage Taiwanese athletes [13] and another study supported strong invariance in Spanish adolescents [8]. So, the SWLS seems to show a certain sensitivity to age, meaning that age impacts the interpretation of SWLS items.

However, very little attention has been paid to the longitudinal invariance of the SWLS. Measurement invariance to the scalar level was observed across two SWLS administrations over a 2-month period for undergraduates in Taiwan [13], and across three administrations over an 18-month period in Serbian undergraduate students [24], although in the latter, a modified single-factor model allowing correlated residuals between items 4 and 5 was supported. Among older adults in France [25], configural invariance was rejected (they observed partial scalar invariance when omitting item 5) across five SWLS administrations over 2-year intervals. To the best of the authors' knowledge, in the unique study carried out in adolescence [13], measurement invariance showed that the SWLS was partial strong longitudinal (6 month) invariant (equality of factor patterns, loadings across time for all items and equality of intercepts for items 2, 3, and 4 across time).

Although the SWLS has been used in longitudinal studies to assess individuals' life satisfaction across time [26–28], these studies do not examine the longitudinal invariance of the SWLS. Similar to measurement invariance across different groups, longitudinal measurement invariance analysis examines the equality of factor structure for a measurement, but its focus is on equality across time [13]; that is, cross-sectional invariance occurs across subgroups in a single time

point while longitudinal invariance occurs across multiple time points.

When mean differences are compared (in this paper especially across time), a key question is whether or not an instrument measures the same construct with equal structural validity across time points. Unless there is reasonable support for the temporal stability of the SWLS, it is not justified to compare life satisfaction scores. Hence, longitudinal measurement invariance of the scale is an important psychometric property and a prerequisite, especially when it is administered in a longitudinal study to the assessment of change over time [29]. As [30] claims, without prior testing longitudinal measurement invariance, it cannot be determined whether temporal change observed in a construct is due to true change or to changes in the structure or measurement of the construct over time. The research findings and conclusions regarding relationships between variables and mean differences in latent traits could be compromised and invalid [24]. To sum up, longitudinal measurement invariance entails that (1) individual SWLS scores at each time point are representing a similar underlying construct [31]; (2) the temporal changes in those scores are due to developmental and experiential changes in life satisfaction assessed by the SWLS, rather than due to temporal inconsistencies among the items comprising this scale [32]; and (3) since the metrics are the same at each time point and between groups, latent means can be compared [33].

Therefore, the main purpose of this research is to examine longitudinal measurement invariance of the SWLS across two time points (the beginning and the end of the academic year), and time points and groups (gender and age) in adolescence to determine whether the SWLS has satisfactory properties for longitudinal comparison and can be effectively used to examine adolescents' satisfaction with life change or development.

Method

Participants

The sample analyzed was composed of 484 Spanish adolescents, 46.7% male and 53.3% female, who completed the SWLS at the beginning (T1) and the end of the academic year (T2); age range from 11.85 to 18.83 in Time 1 ($M_{age1} = 14.95$, $SD_1 = 1.81$) and from 12.52 to 19.50 in Time 2 ($M_{age2} = 15.61$, $SD_2 = 1.81$). All participants belonged to six academic levels from Grade 7 to Grade 12 and the sample was split into three groups: those under 15 years old (40.5%) during the academic year (early adolescence); those under 15 years old at the beginning of the year that become older than 15 (11.8%) at the end of the academic

year (early-middle adolescence); and those over 15 years old (47.7%) throughout the academic year (middle adolescence).

Instruments

Life satisfaction was measured using the Spanish version of the SWLS [5], comprised of 5 items rated on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), with higher values representing greater satisfaction.

Procedure

Ethical permission to conduct the study was obtained from the Committee on Ethics of Research and Teaching (CEID) from the University of the Basque Country (Spain). Schools' participation was requested verbally and by writing. After obtaining schools' authorization, a letter of consent to participate was sent to all families. Subsequently, the battery of questionnaires was administered during school hours only to participants who brought the informed consent from their parents. Pupils answered the questionnaires individually, and during the process, anonymity of the responses and voluntary participation were ensured.

Data analysis

Statistical analysis were carried out with SPSS v22 for preliminary results and some packages of R (Lavaan and SEM Tools) for Multiple Group Confirmatory Factor Analysis (MGCFA), the standard method to analyze measurement invariance across groups and time points [34]. Previous descriptive analysis included means, variances, skewness, and kurtosis of the items. The reliability of the scale was obtained through Cronbach's alpha, Guttman lambda, and MacDonald's Omega total. The MGCFA analysis included confirmatory factor analysis to examine the fit of the unifactorial model to data and to test the measurement of longitudinal invariance (MLI) of the scale across time points and time points and groups (gender and age). The variables analyzed are considered to be continuous by nature given the number of response categories greater than five [35].

MLI requires testing successive models that are compared among them with increasingly stringent constraints. First, a baseline model is defined with a pattern of loadings similar for each time, although loadings, intercepts, and other parameters may be different across time. Next, the fit of each model is compared to the fit of the previous model. If model fit does not worsen, the subsequent model is selected (e.g., scalar invariance exists if the fit to the data is not worse than the fit of the metric invariance model). Although there are many different statistical methods to decide when model fit worsens, the usual procedure used in structural equation modeling (SEM) is to employ root mean square error of

approximation (RMSEA) and Comparative Fit Index (CFI); frequently these indexes are combined along with variations among the models through ΔRMSEA and ΔCFI. If the Δ in these indexes is low (usually accepted $\Delta < .01$), it is assumed that invariance exists [19]. According to [36] recommendations for non-invariance ($\Delta CFI \geq .01$, supplemented by $\Delta RMSEA \geq .015$), and [37, 38] suggestions, ΔCFI is the best indicator (when $\Delta CFI < .01$). These indexes have been used to analyze the longitudinal invariance of the scale at the two time points and also with regard to gender and age. The likelihood ratio Chi-square test has not been used due to the controversy of its usage as a measurement of the fit of a model [19].

If strict invariance stays, then it is possible to compare latent means through different time points. In this case, two models are fit: the first model constrains loadings, intercepts, factors means, and error variances to be equal between times; and the second model holds the same restrictions, but the means are allowed to vary. The size effect of the difference between latent means was calculated according to [39] proposal.

Results

Means, standard deviations, skewness, and kurtosis for SWLS items and the correlation matrix are presented in Table 1. All values of skewness and kurtosis are within an acceptable range ($-2, 2$) according to [40] recommendations. Then, a well-fitting baseline model was established. The single-factor SWLS structure with five items underpinned by the previous research showed a reasonable fit at both times T1 ($CFI_1 = 1.000$, $TLI_1 = .997$, $RMSEA_1 = .080$, and $SRMR_1 = .028$), and T2 ($CFI_2 = .997$, $TLI_2 = .995$, $RMSEA_2 = .032$, and $SRMR_2 = .034$). Likewise, the reliability and composite reliability showed good results when analyzing both time points together, as well as separately (Cronbach's alpha = .86, Guttman's lambda = .84, McDonald's Omega total = .89).

The goodness-of-fit statistics for tests of measurement longitudinal invariance of the single-factor model across time points, and gender and age are presented in Table 2, showing the fit index and the variations in CFI and RMSEA indexes between the models. As can be seen, the model comparisons indicate that the factor loadings can be assumed to be equal, since ΔCFI and ΔRMSEA are below the proposed cut-point of .01. These results allow analysis of the differences between time points and groups (gender and age), according to differences in scale values.

The latent factor means comparisons across time points have been made through the method described above. The differences in gender (group 2, estimate = .025, $z = .277$, $p = .78$, SE = .03) and age (group 2, estimate = .180,

Table 1 Descriptive analysis and correlation matrix

	M	SD	Sk	Ku	Swls1T1	Swls2T1	Swls3T1	Swls4T1	Swls5T1	Swls1T2	Swls2T2	Swls3T2	Swls4T2	Swls5T2
Swls1T1	4.87	1.48	−.781	−.055	1									
Swls2T1	5.34	1.35	−.948	.572	.616**	1								
Swls3T1	5.57	1.38	−1.279	1.475	.696**	.587**	1							
Swls4T1	5.16	1.44	−.907	.233	.533**	.375**	.561**	1						
Swls5T1	4.85	1.89	−.649	−.765	.584**	.456**	.571**	.547**	1					
Swls1T2	4.88	1.48	−.811	.117	.571**	.464**	.577**	.434**	.451**	1				
Swls2T2	5.46	1.31	−1.060	1.092	.419**	.580**	.475**	.315**	.365**	.601**	1			
Swls3T2	5.48	1.38	−1.213	1.312	.535**	.452**	.617**	.409**	.497**	.697**	.570**	1		
Swls4T2	5.16	1.44	−.784	.099	.423**	.274**	.434**	.472**	.403**	.557**	.348**	.634**	1	
Swls5T2	4.71	1.95	−.469	−1.034	.425**	.393**	.472**	.368**	.566**	.512**	.424**	.588**	.536**	1

$**p < 0.01$

Table 2 Goodness-of-fit statistics for tests of longitudinal invariance of a single-factor model

Model	CFI	RMSEA	ΔCFI	ΔRMSEA
Invariance across T1&T2				
Configural	.972	.073		
Weak	.973	.068	.001	.005
Strong	.969	.068	.004	.000
Strict	.970	.067	.001	.001
Invariance across T1&T2 and gender				
Configural	.963	.075		
Weak	.963	.073	.000	.002
Strong	.959	.075	.004	.002
Strict	.956	.076	.003	.001
Invariance across T1&T2 and age				
Configural	.958	.076		
Weak	.959	.074	.001	.002
Strong	.956	.076	.003	.002
Strict	.946	.083	.010	.008

This table shows the MLI across time points, and time points and groups

$z = 1.869$, $p = .062$, SE = .23; group 3, estimate = .121, $z = .782$, $p = .434$, SE = .15) have not been significant and the small differences across time points show that the latent factor means hold across time and in both groups (gender and age).

Discussion

Our results showed that all values of skewness and kurtosis were within an acceptable range. The single-factor SWLS structure with five items supported by the previous research in adolescents [4, 5, 8, 10, 11, 13–16] showed a reasonable fit at both times and the results of reliability and composite reliability of the scale were also good at both times corroborating previous findings [4, 5, 8, 10, 11]; so, the results of the present study suggest that the SWLS is a valid and reliable scale for studying life satisfaction in adolescents from Spain.

On the other hand, it is relevant to point out that although the SWLS have been used in longitudinal studies [26–28], these studies have not examined whether the SWLS has longitudinal invariance. In this regard, measurement invariance showed that SWLS was partial strong longitudinal (6 month) invariant (equality of factor patterns, loadings across time for all items, and equality of intercepts for items 2, 3, and 4 across time) in an adolescent athletes sample from Taiwan [13], being the unique study that has been carried out in this age previously. In the present study, the goodness-of-fit statistics for test of longitudinal measurement invariance of the single-factor model across time and groups (gender and age)

shows strict invariance, which means that time differences in variances of gender and age are attributable only to time differences in variances of their respective latent variables; this is to say that the verification of longitudinal invariance on the SWLS scale means that their items measure the same attribute across time. This is the first study that supports this level of invariance in adolescence population. The differences in gender and age have not been significant and the small differences across time points show that the latent factor means hold across time in both groups.

Our results could be surprising since adolescence is an important developmental phase, influenced by a multitude of significant physical, psychological, and social changes. Thus, given the changes that occur in the academic context and other areas (family, social, personal, biological, etc.), it would be expected to find some variation in the life satisfaction of adolescents. Some authors claim that life satisfaction may show fluctuation in relatively short periods of time, varying even from one semester to the next within the same school year [41]. Generally speaking, there is a consensus in relation to the significant downward trend during adolescence [28, 42, 43].

However, it should be noted the existence of discordant results, in which there are no significant fluctuations (15–18 years) [44], or even it has been observed among Finnish adolescents (15–17 years) an increase of satisfaction with life in the school transition from compulsory education or primary to secondary or vocational education [26]. The latter study reveals that the majority of adolescents (66%) show a high life satisfaction that remains stable over 2 years. Our results have confirmed that life satisfaction is relatively stable during one academic year in secondary school, although this period perhaps is too short to appreciate the possible fluctuations that may take place in the perception of life satisfaction during adolescence.

Limitations and future studies

This study has certain limitations. Although longitudinal measurement invariance has been analyzed through two repeated measures, more time point measures would be interesting to get more information about the scale, in order to have the possibility to use more complex statistical options like latent growth curve models analysis and to analyze the development of life satisfaction in adolescence. On one hand, due to the fact that longitudinal studies are conducted across different time intervals (for example, six months [13], one academic year in our current study, or 6 and 18 months [24]), broader time spans are desirable for testing the stability of invariance properties. In the present study, invariance was evaluated over relatively a short period of time (one academic year). Therefore, conclusions regarding long-term stability and structural invariance of the

SWLS cannot be assured. More research is clearly needed to examine longitudinal invariance of the SWLS through longer time periods.

On the other hand, the sample was incidental, not too large and focused in adolescence, consequently future studies should considerate including larger probabilistic sample, as well as wider range of ages.

Taking into account that there is a lack of studies that have analyzed the longitudinal measurement invariance of the SWLS, future studies should pay more attention to this property. Due to the fact that considerable changes in life satisfaction may occur, future investigations should assess the temporal stability of SWLS before analyzing the development of the global cognitive judgment of satisfaction with one's life.

Conclusion

In summary, the present study confirmed the following: the single-factor structure of the SWLS in Spanish adolescents; good reliability and composite reliability; the full longitudinal measurement invariance; and that there are negligible differences across time points considering gender and age. If these findings are further replicated, the scale can be used to compare life satisfaction across time points considering different age and gender groups.

Compliance with ethical standards

Conflict of interest The authors declare that they have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Ethical permission to conduct the study was obtained from the Committee on Ethics of Research and Teaching (CEID) from the University of University of the Basque Country.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49,* 71–75. https://doi.org/10.1207/s15327752jpa4901_13.
2. Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. E. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125,* 276–302. https://doi.org/10.1037/0033-2909.125.2.276.
3. Emerson, S. D., Guhn, M., & Gadermann, A. M. (2017). Measurement invariance of the Satisfaction With Life Scale: Reviewing three decades of research. *Quality of Life Research, 26,* 2251–2264. https://doi.org/10.1007/s11136-017-1552-2.
4. Atienza, F. L., Balaguer, I., Corte-Real, N., & Fonseca, A. M. (2016). Factorial invariance of the Satisfaction With Life Scale in adolescents from Spain and Portugal. *Psicothema, 28,* 353–356. https://doi.org/10.7334/psicothema2016.1.
5. Atienza, F. L., Pons, D., Balaguer, I., & García-Merita, M. L. (2000). Propiedades psicométricas de la escala de satisfacción con la vida en adolescentes. *Psicothema, 12,* 331–336.
6. Bai, X., Wu, C., Zheng, R., & Ren, X. (2011). The psychometric evaluation of the Satisfaction With Life Scale using a nationally representative sample of China. *Journal of Happiness Studies, 12,* 183–197. https://doi.org/10.1007/s10902-010-9186-x.
7. Clench-Aas, J., Nes, R. B., Dalgard, O. D., & Aarø, L. E. (2011). Dimensionality and measurement invariance in the Satisfaction With Life Scale in Norway. *Quality of Life Research, 20,* 1307–1317. https://doi.org/10.1007/s11136-011-9859-x.
8. Esnaola, I., Benito, M., Antonio-Agirre, I., Freeman, J., & Sarasa, M. (2017). Measurement invariance of the Satisfaction With Life Scale (SWLS) by country, gender and age. *Psicothema, 29,* 596–601. https://doi.org/10.7334/psicothema2016.394.
9. Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction With Life Scale (SWLS). *European Journal of Psychological Assessment, 27,* 127–132. https://doi.org/10.1027/1015-5759/a000058.
10. Moksnes, U. K., Løhre, A., Byrn, D. G., & Haugan, G. (2014). Satisfaction With Life Scale in adolescents: Evaluation of factor structure and gender invariance in a Norwegian sample. *Social Indicators Research, 118,* 657–671. https://doi.org/10.1007/s11205-013-0451-3.
11. Silva, A. D., Taveira, M. C., Marques, C., & Gouveia, V. V. (2015). Satisfaction With Life Scale among adolescents and young adults in Portugal: Extending evidence of construct validity. *Social Indicators Research, 92,* 489–496. https://doi.org/10.1007/s11205-014-0587-9.
12. Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment, 5,* 164–172. https://doi.org/10.1037/1040-3590.5.2.164.
13. Wu, C. H., Chen, L. H., & Tsai, Y. M. (2009). Longitudinal invariance analysis of the Satisfaction With Life Scale. *Personality and Individual Differences, 46,* 396–401. https://doi.org/10.1016/j.paid.2008.11.002.
14. Jovanović, V. (2016). The validity of the Satisfaction With Life Scale in adolescents and a comparison with single-item life satisfaction measures: A preliminary study. *Quality of Life Research, 25,* 3173–3180. https://doi.org/10.1007/s11136-016-1331-5.
15. Pons, D., Atienza, F. L., Balaguer, I., & García-Merita, M. L. (2000). Satisfaction With Life Scale: Analysis of factorial invariance for adolescents and elderly persons. *Perceptual and Motor Skills, 91,* 62–68. https://doi.org/10.1080/17405629.2012.707778.
16. Tomás, J. M., Gutiérrez, M., Sancho, P., & Romero, I. (2015). Measurement invariance of the Satisfaction With Life Scale (SWLS) by gender and age in Angola. *Personality and*

Individual Differences, 85, 182–186. https://doi.org/10.1016/j.paid.2015.05.008.

17. Timmons, A. C. (2010). Establishing factorial invariance for multiple-group confirmatory factor analysis. *KUant Guide*, *22.1.* Retrieved from http://crmda.dept.ku.edu/resources/kuantguides/22.Factorial_Invariance_Guide.pdf.

18. Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44,* S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89.

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

20. Wu, A. D., Li, Z., & Zumbo, B. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment Research & Evaluation, 12,* 1–26.

21. Atienza, F. L., Balaguer, I., & García-Merita, M. L. (2003). Satisfaction With Life Scale: Analysis of factorial invariance across sexes. *Personality and Individual Differences, 35,* 1255–1260. https://doi.org/10.1016/S0191-8869(02)00332-X.

22. Siedlecki, K. L., Tucker-Drob, E. M., Oishi, S., & Salthouse, T. A. (2008). Life satisfaction across adulthood: Different determinants at different ages? *The Journal of Positive Psychology, 3,* 153–164. https://doi.org/10.1080/17439760701834602.

23. Hultell, D., & Gustavsson, J. P. (2008). A psychometric evaluation of the Satisfaction With Life Scale in a Swedish nationwide sample of university students. *Personality and Individual Differences, 44,* 1070–1079. https://doi.org/10.1016/j.paid.2007.10.030.

24. Jovanović, V. (2017). Measurement invariance of the Serbian version of the Satisfaction With Life Scale across age, gender, and time. *European Journal of Psychological Assessment.* https://doi.org/10.1027/1015-5759/a000410.

25. Gana, K., Bailly, N., Saada, Y., Joulain, M., & Alaphilippe, D. (2013). Does life satisfaction change in old age: Results from an 8-year longitudinal study. *The Journals of Gerontology B, 68,* 540–552. https://doi.org/10.1093/geronb/gbs093.

26. Salmela-Aro, K., & Tuominen-Soini, H. (2010). Adolescents' life satisfaction during the transition to post-comprehensive education: Antecedents and consequences. *Journal of Happiness Studies, 11,* 683–701. https://doi.org/10.1007/s10902-009-9156-3.

27. Salmela-Aro, K., & Tynkkynen, L. (2010). Trajectories of life satisfaction across the transition to post-compulsory education: Do adolescents follow different pathways? *Journal of Youth and Adolescence, 39,* 870–881. https://doi.org/10.1007/s10964-009-9464-2.

28. Shek, D. T. L., & Liu, T. T. (2014). Life satisfaction in junior secondary school students in Hong Kong: A 3-year longitudinal study. *Social Indicator Research, 117,* 777–794. https://doi.org/10.1007/s11205-013-0398-4.

29. Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices and recommendations for organizational research. *Organizational Research Methods, 3,* 4–70. https://doi.org/10.1177/109442810031002.

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

31. Moreira, P. S., Santos, N., Castanho, T., Amorim, L., Portugal-Nunes, C., Sousa, N., et al. (2018). Longitudinal measurement invariance of memory performance and executive functioning in healthy aging. *PLoS ONE.* https://doi.org/10.1371/journal.pone.0204012.

32. Hawes, S. W., Byrd, A. L., Kelley, S. E., Gonzalez, R., Edens, J. F., & Pardini, D. A. (2018). Psychopathic features across development: Assessing longitudinal invariance among Caucasian and African American youths. *Journal of Research in Personality, 73,* 180–188. https://doi.org/10.1016/j.jrp.2018.02.003.

33. Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), 10–18. https://doi.org/10.1111/j.1750-8606.2009.00110.x.

34. Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.

35. Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods, 22*(3), 486–506.

36. Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14,* 464–504. https://doi.org/10.1080/10705510701301834.

37. Hirschfeld, G., & von Brachel, R. (2014). Multiple-group confirmatory factor analysis in R: A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation, 19,* 1–12.

38. Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1–55. https://doi.org/10.1080/10705519909540118.

39. Hancock, G. R. (2001). Effect size, power, and sample size determination for structured means modeling and mimic approaches to between-groups hypothesis testing of means on a single latent construct. *Psychometrika, 66,* 373–388. https://doi.org/10.1007/BF02294440.

40. Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38,* 171–189. https://doi.org/10.1111/j.2044-8317.1985.tb00832.x.

41. Siddall, J., Huebner, E. S., & Jiang, X. (2013). A prospective study of differential sources of school-related social support and adolescent global life satisfaction. *American Journal of Orthopsychiatry, 83,* 107–114. https://doi.org/10.1111/ajop.12006.

42. Suldo, S. M., Bateman, L. P., & Gelley, C. D. (2014). Understanding and promoting school satisfaction in children and adolescents. In M. J. Furlong, R. Gilman, & E. S. Huebner (Eds.), *Handbook of positive psychology in schools* (2nd ed., pp. 365–380). New York: Routledge.

43. Weber, M., & Huebner, E. S. (2015). Early adolescents' personality and life satisfaction: A closer look at global vs. domain-specific satisfaction. *Personality and Individual Differences, 83,* 31–36. https://doi.org/10.1016/j.paid.2015.03.042.

44. Porubanova-Norquist, M. (2012). Character as a predictor of life satisfaction in Czech adolescent sample: 3-Year follow-up study. *Personality and Individual Differences, 53,* 231–235. https://doi.org/10.1016/j.paid.2012.03.022.