

RESEARCH ARTICLE

## Can life satisfaction be measured fairly for different groups of South African first-year students? Testing the Satisfaction with Life Scale

Clarisse van Rensburg<sup>1</sup> & Karina Mostert<sup>2</sup>

**Article history:** Received 1 November 2022 | Accepted 31 May 2023 | Published 14 August 2023

### ABSTRACT

Student well-being has gradually become a topic of interest in higher education, and the accurate, valid, and reliable measure of well-being constructs is crucial in the South African context. This study examined item bias and configural, metric and scalar invariance of the Satisfaction with Life Scale (SWLS) for South African first-year university students. A cross-sectional design was used. A sample of 780 first-year South African university students was included. Confirmatory factor analysis, differential item functioning measurement invariance, and internal consistency were tested. A one-factor structure was confirmed. Item 1 of the SWLS was particularly problematic concerning bias (uniform and non-uniform bias). Measurement invariance was established; however, Item 1 was again problematic, resulting in partial metric and scalar invariance. The scale was reliable (Cronbach's  $\alpha$  was 0.83; McDonald's omega ( $\omega$ ) was 0.83). This study contributes to the limited research on the specific psychometric properties of the SWLS in a diverse higher education setting. The results could assist with valid and reliable measurements when developing interventions to enhance student well-being.

### KEYWORDS

*Satisfaction with Life Scale, item bias, differential item functioning, measurement invariance, first-year university students, student affairs*

### RÉSUMÉ

Le bien-être des étudiants est progressivement devenu un sujet d'intérêt dans l'enseignement supérieur, et la mesure précise, valide et fiable des constructions liées au bien-être est cruciale dans le contexte sud-africain. Cette étude a examiné le biais des items et l'invariance de la configuration, de la métrique et de l'échelle de satisfaction de vie (SWLS) pour les étudiants de première année d'université en Afrique du Sud. Une conception transversale a été utilisée. Un échantillon de 780 étudiants sud-africains de première année a été choisi. Une analyse factorielle confirmatoire, le fonctionnement différentiel des items, l'invariance de mesure et la cohérence interne ont été testées. Une structure à un facteur a été confirmée. L'item 1 de la SWLS était particulièrement problématique en termes de biais (biais uniforme et non uniforme). L'invariance de mesure a été établie ; cependant, l'item 1 était à nouveau problématique, entraînant une invariance métrique et scalaire partielle. L'échelle était fiable

1 Clarisse van Rensburg, JOBJACK, Cape Town, South Africa. Email: [clarisse0206@gmail.com](mailto:clarisse0206@gmail.com). ORCID: 0000-0003-4573-6575

2 Prof. Karina Mostert, North-West University, Potchefstroom Campus, South Africa. Email: [Karina.Mostert@nwu.ac.za](mailto:Karina.Mostert@nwu.ac.za). ORCID: 0000-0001-5673-5784

(l' $\alpha$  de Cronbach était de 0,83 ; l' $\omega$  de McDonald était de 0,83). Cette étude contribue à la recherche limitée sur les propriétés psychométriques spécifiques de la SWLS dans un contexte diversifié d'enseignement supérieur. Les résultats pourraient aider à obtenir des mesures valides et fiables lors de l'élaboration d'interventions visant à améliorer le bien-être des étudiants.

## MOTS-CLÉS

*Biais des items, échelle de satisfaction de vie, étudiants de première année d'université, fonctionnement différentiel des items, invariance de mesure, œuvres estudiantines, services étudiants*

## Introduction

Life satisfaction is an essential indicator of individual and social well-being and includes the perception that one is moving towards accomplishing significant life goals (Esnaola et al., 2019; Jovanović, 2019). Life satisfaction is also crucial to first-year university students, as they face a period of uncertainty in which they idealise the values of their lives, prepare for the world of work, and actively explore their adult roles (Gökalp & Topal, 2019). Studies show significant relationships between high levels of life satisfaction, taking on more responsibility, experiencing less stress and emotional loneliness and more resilience in overcoming academic challenges (Gökalp & Topal, 2019; Rode et al., 2005). There is also a relationship between life satisfaction and satisfaction with educational experiences, healthy relationships, self-esteem (Chow, 2005), engagement, motivation and study satisfaction (Lewis et al., 2011; Wach et al., 2016). Conversely, there are also associations between low levels of life satisfaction, perceived stress, anxiety, and burnout (Alleyne et al., 2010; Serin et al., 2010), higher levels of impaired concentration and deteriorated academic performance (Rode et al., 2005).

One of the most widely used scales in assessing life satisfaction is the Satisfaction with Life Scale (SWLS) (Diener et al., 1985). Periodic assessments are needed to accurately establish and measure well-being in the higher education sector, including measures of life satisfaction such as the SWLS (Băcilă et al., 2014). However, various factors challenge fair psychological testing in South Africa, such as the distribution of socio-economic resources, diversity in culture and language, and education and employment statuses (Foxcroft & Roodt, 2009). Psychological testing and other similar assessments are governed in South Africa by the Employment Equity Act No. 55 of 1998, Section 8 (President of the Republic of South Africa, 1998), which states that assessments are prohibited unless they can scientifically be shown to be reliable and valid, can be applied fairly to all ethnic groups and cultures, and are not biased against any person or group.

In addition, psychological assessments in South Africa are controversial due to past unfair, undiscerning, and biased use (Laher & Cockcroft, 2014). Historically, assessment practices in South Africa have been known to use measurement instruments from Western countries, often without any adaptation to South Africa's multi-cultural and diverse context (Blokland, 2016). Consequently, the majority of South Africa's population was excluded from these assessment practices, as they tend to cater mainly for the

Western, educated, industrialised, rich, and developed population sectors (Laher & Cockcroft, 2013; Laher & Cockcroft, 2014). Therefore, questions related to test bias and equivalence are raised when applying adapted and imported measurement instruments in South Africa (Teresi & Fleishman, 2007; Van De Vijver & Rothmann, 2004).

It is essential to distinguish between the concepts of item bias and equivalence. Bias refers to the presence of annoyance factors (items invoking added abilities or traits) when making cross-group comparisons (Schaap, 2011; Van De Vijver & Rothmann, 2004). In addition, another source of bias can be test items themselves – also referred to as **item bias** or **differential item functioning (DIF)** (Van De Vijver & Rothmann, 2004). Item bias signifies that the meaning of one or more scale items is not being understood identically across groups, is not applicable to a specific group or that semantic differences are present in how items are conceptualised (Cleary & Hilton, 1968). When respondents have the same standing on the underlying construct and are from different cultures but have different mean scores on the item, this could reflect actual differences in the construct or that item bias is present (Van De Vijver & Rothmann, 2004).

A distinction should be made between uniform and non-uniform bias. **Uniform bias** refers to the likelihood of similar responses for one group being systematically higher or lower at specific endorsement levels (the underlying construct) compared to other groups (Swaminathan & Rogers, 1990; Teresi & Fleishman, 2007). **Non-uniform bias** refers to the difference in the likelihood of similar answers across groups varying across all levels of endorsement (Swaminathan & Rogers, 1990; Teresi & Fleishman, 2007). The most common sources of item bias include ambiguities in the original item, poor item translation, the influence of cultural specifics (connotations or nuisance factors) associated with the wording of the item, and low appropriateness and familiarity of the item content in some cultures (Van De Vijver & Rothmann, 2004).

**Invariance (or equivalence)** indicates whether a construct is interpreted and understood similarly across different groups, which is essential for cross-group comparisons (Mellenbergh, 1989; Milfont & Fischer, 2010; Putnick & Bornstein, 2016; Van De Schoot et al., 2012). **Configural invariance** indicates the extent to which the factor structure of a measure can be replicated across different groups, which is crucial for meaningful comparisons (He & Van De Vijver, 2012; Schaap, 2011; Van De Vijver & Rothmann, 2004). **Metric invariance** involves equal factor loadings, across groups, for similar items (i.e. when individuals from different cultures who speak different languages complete the same questionnaire and conceptualise the construct the same way) (Milfont & Fisher, 2010; Morton et al., 2019). **Scalar invariance** establishes whether a test score has a similar meaning in its interpretation regardless of cultural background (He & Van De Vijver, 2012; Laher, 2008).

In essence, concerning the psychometric properties of assessment instruments, item bias and invariance testing will aid in establishing whether measures are fair to use for different sub-groups in the specific South African context (Schaap, 2011). Therefore, this study emphasizes the concepts of bias and invariance testing to ensure the validation of existing instruments in cross-cultural groups to ensure meaningful comparisons across sub-groups (Van De Vijver & Rothmann, 2004).

While scholars recently investigated the psychometric properties of the SWLS among South African samples, including an adult population (Schutte et al., 2021) and primary and secondary school teachers (Pretorius & Padmanabhanunni, 2022), studies testing item bias and equivalence are limited. The present study explores the psychometric properties, specifically item bias and invariance (configural, metric and scalar), of the SWLS in a sample of first-year South African students.

## Literature

### Measurement and psychometric properties of the Satisfaction with Life Scale

The SWLS measures a single life satisfaction construct that could indicate levels of satisfaction with life throughout one's life span (Tomás et al., 2015). The scale displays favourable psychometric properties, has been validated in various countries and is translated into numerous languages, including Spanish, Portuguese, Dutch, and German (Diener et al., 1985; Gouveia et al., 2009).

Studies testing bias and invariance for the SWLS are scarce and report mixed results. Regarding the item bias of the SWLS, a study conducted on a Turkish university student sample concluded that the items of the SWLS are unbiased across gender groups (Avcu, 2021). However, Hultell and Gustavsson (2008) found that Item 4 and Item 5 are sensitive to age.

With regards to **configural invariance**, most researchers have found the SWLS to be invariant across gender, age (Glaesmer et al., 2011; Hinz et al., 2018; Lorenzo-Seva et al., 2019; Wu & Yao, 2006), and countries such as Spain and Portugal (Atienza González et al., 2016), Germany (Glaesmer et al., 2011), Columbia (Ruiz et al., 2019), the United States, and Brazil (Zanon et al., 2014). However, other studies report configural invariance for gender groups, albeit not for age groups (Shevlin et al., 1998; Wu et al., 2009).

In addition, studies report **metric invariance** for different age groups (Pons et al., 2000; Glaesmer et al., 2011), gender groups (Emerson et al., 2017; Hinz et al., 2018; Jovanović, 2019; Moksnes et al., 2014; Ruiz et al., 2019), and cultures (Atienza González et al., 2016; Emerson et al., 2017; Jovanović & Brdar, 2018). However, Zanon et al. (2014) presented evidence against metric invariance between undergraduates from the United States and Brazil, specifically for Items 4 and 5.

Studies on the SWLS support the notion that **scalar invariance** is supported across gender groups (Clench-Aas et al., 2011; Hultell & Gustavsson, 2008; Shevlin et al., 1998; Zanon et al., 2014), age groups (Durak et al., 2010; Gouveia et al., 2009; Tomás et al., 2015; Wu et al., 2009), and several European countries (e.g., Austria, Bosnia and Herzegovina, Croatia, Montenegro, and Serbia; Jovanović & Brdar, 2018). However, some studies reported insufficient evidence for scalar invariance across age groups (Clench-Aas et al., 2011; Hultell & Gustavsson, 2008) and countries (Atienza González et al., 2016; Whisman & Judd, 2016).

## Method

### Research procedure and participants

Before data collection commenced, permission was obtained from the relevant university to conduct research. An ethics application was submitted and approved, focusing specifically on anonymity, confidentiality, and voluntary participation (ethics number: NWU-HS-2014-0165). A web-based survey link was sent via email and posted on the university's online platform for first-year modules. The study's goal, purpose, and value to the university were explained. The sample consisted of 780 first-year students aged 18 to 20. Of the 780 participants, 38.8% indicated that they spoke Afrikaans, 33.1% indicated that they spoke Setswana, and 6.2% indicated that they spoke Sesotho (three of the eleven official language groups in South Africa). The sample consisted of three campuses: Campus 1 (38.3%), Campus 2 (50.5%), and Campus 3 (9.7%). Concerning gender, 61.8% of the participants identified as women, and 38.2% identified as men.

### Measuring instrument

The SWLS was developed by Diener and colleagues (1985) and aimed to measure a single life satisfaction construct that could indicate levels of satisfaction with life throughout one's life span (Tomás et al., 2015). Participants are asked five questions (e.g. "The conditions of my life are excellent"). A seven-point Likert-type scale is used, ranging from 1 (strongly disagree) to 7 (strongly agree). Pavot and Diener (1993) confirmed the scale's reliability, reporting Cronbach's coefficient alphas ranging from 0.79 to 0.89.

### Statistical analysis

The statistical analyses were conducted using Mplus 8.6 (Muthén & Muthén, 2021). Before bias and invariance were tested, confirmatory factor analysis (CFA) was used to confirm the one-factor structure of the SWLS. Maximum likelihood with robust standard errors (MLR) was used in the CFA due to the small samples in some groups to supplement the item bias analyses. The following fit indices and cut-off scores were used to estimate the measurement model's goodness-of-fit: the traditional chi-square ( $\chi^2$ ) statistic, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the root mean square error of approximation (RMSEA) and the standardised root mean square residual (SRMR). Values of 0.90 and above indicate an acceptable fit for CFI and TLI (Byrne, 2001). Regarding the RMSEA scores, various researchers suggest using a cut-off score below 0.05 as the 'golden rule of thumb' to indicate model fit; however, values between 0.05 and 0.08 are considered to be an acceptable fit (Browne & Cudeck, 1993; Chen et al., 2008; Hu & Bentler, 1999; Steiger, 1989). Concerning SRMR, a cut-off value of 0.05 was used (Browne & Cudeck, 1993; Hu & Bentler, 1999).

Differential item functioning (DIF) was used to test for the presence of **item bias** by using the *lordif* package (Choi et al., 2011) in *RStudio* (<https://www.rstudio.com/>). The following formulas were used and compared to test for uniform and non-uniform bias, using ordinal logistic regression to generate three likelihood-ratio  $\chi^2$  statistics (Choi et al., 2011):

Model 0:  $\text{logit } P(u_i \geq k) = \alpha_k$

Model 1:  $\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 * \text{ability}$

Model 2:  $\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 * \text{ability} + \beta_2 * \text{group}$

Model 3:  $\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 * \text{ability} + \beta_2 * \text{group} + \beta_3 * \text{ability} * \text{group}$

Based on the formulas mentioned above, uniform bias can be indicated with a significant difference at  $p < 0.01$  when comparing logistic models 1 and 2 ( $\chi^2_{12}$ ;  $df = 1$ ), and non-uniform bias, when comparing models 2 and 3 ( $\chi^2_{23}$ ;  $df = 1$ ) (Choi et al., 2011). Total DIF is indicated with a significant difference at  $p < 0.01$  when comparing models 1 and 3 ( $\chi^2_{13}$ ;  $df = 2$ ) (Choi et al., 2011). The magnitude of DIF can be quantified using the pseudo-McFadden  $R^2$  statistic, which can be classified as either negligible ( $< 0.13$ ), moderate (between 0.13 and 0.26), or large ( $> 0.26$ ) (Zumbo, 1999). However, DIF can be under-identified by only using the pseudo-McFadden  $R^2$  statistic (Jodoin & Gierl, 2001; Kim et al., 2007). Therefore, to identify uniform DIF, the difference in the coefficient from Models 1 and 2 was used – with differences of 10% indicating a practically meaningful effect (Crane et al., 2004; Maldonado & Greenland, 1993). Lower than 5% and even 1% thresholds are also considered (Crane et al., 2007). In this study, a threshold of 5% was considered.

**Measurement invariance** was investigated for three different groupings: (1) language (Afrikaans, Sesotho, and Setswana, three indigenous South African language groups), (2) campus (three campuses were included), and (3) gender (men and women). A multi-group analysis framework was used to test for the configural invariance model (analogous factor structure), metric invariance model (similar factor loadings), and scalar invariance model (similar intercepts). CFI and RMSEA values were used to indicate measurement invariance. CFI is considered a good fit with values  $> 0.90$  and better if the values are  $> 0.95$ . Regarding RMSEA, cut-off values of 0.05 and 0.08 are considered acceptable (Van De Schoot, 2012). Changes in CFI ( $\Delta\text{CFI}$ ) were used as they are less susceptible to the effects of changes in  $df$  (Shi et al., 2019). A  $\Delta\text{CFI}$  value ( $> -0.01$ ; that is a worsening of the model fit according to CFI) between two nested models indicates that the added group constraints have led to a poorer fit; in other words, invariance has not been achieved, and the more constrained model is rejected. Additionally, it is essential to note that there are many small variances among groups regarding factor loadings or intercepts; therefore, partial measurement invariance (whether metric or scalar) can be achieved by freeing the loading or intercepts of items (Cheung & Rensvold, 2002; Preti et al., 2013; Van De Schoot et al., 2015).

Finally, the **internal consistency** of the SWLS was determined using Cronbach's alpha coefficients (McCrae et al., 2011; Revicki, 2014). Values where  $\alpha \geq 0.70$  were considered acceptable (Nunnally, 1978). In addition, McDonald's omega was calculated and reported for a more accurate estimation of internal consistency (Cortina et al., 2020). Reliability coefficients  $\geq 0.80$  indicate good internal consistency (Kline, 2015).

## Results

### Factorial validity

Before DIF and invariance testing, the factorial validity of the SWLS was tested with confirmatory factor analysis (CFA) to determine the model's goodness-of-fit. Even though it is best practice to test alternative measurement models (Marsh et al., 1998), given the limited items of the SWLS and based on the various studies supporting the five-item SWLS measuring one underlying factor, a one-factor measuring model was tested in this study. Also, with only five items, two factors would have one just-identified factor (3 items) and another under-identified factor (2 items).

The results indicate that the unidimensional structure of the SWLS is deemed to be a good fit for the data ( $\chi^2 = 806.844$ ;  $df = 10$ ; CFI = 0.964; TLI = 0.928; RMSEA = 0.086; SRMR = 0.035). The CFI value was the preferred index of choice to determine goodness-of-fit (Shi et al., 2019). Table 1 indicates the results for the standardised loadings of the items for the latent variables of the SWLS.

**Table 1: Standardised factor loadings for the latent variables of the SWLS**

Item	Item text	Loading	S.E.	<i>p</i>
Item 1	In most ways my life is close to my ideal	0.543	0.037	0.001
Item 2	The conditions of my life are excellent	0.831	0.024	0.001
Item 3	I am satisfied with my life	0.879	0.016	0.001
Item 4	So far I have gotten the important things I want in life	0.684	0.031	0.001
Item 5	If I could live my life over, I would change almost nothing	0.533	0.035	0.001

*Notes: S.E. = standard error; all p-values < 0.001*

Factor loadings ( $\lambda$ ) can either be classified as high (0.70), medium (0.50), or small (0.30) (Shevlin et al., 1998). The results show that the factor loadings for the SWLS ranged from medium to high.

### Item bias

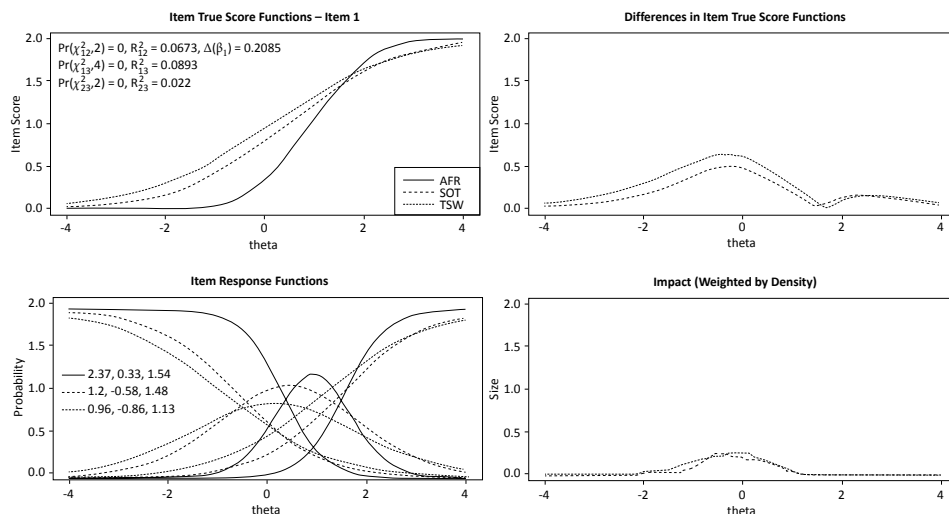
DIF was used to determine the item bias of the SWLS. Uniform and non-uniform bias were tested across different language, campus, and gender groups. Table 2 indicates the DIF of the SWLS.

Table 2 indicates that Item 1 was problematic for language and campus groups, and Item 3 for language groups. No DIF was detected across gender groups. Table 2 shows that Item 1 has uniform, non-uniform and total bias, based on the likelihood-ratio  $\chi^2$  difference testing across models 1, 2 and 3 ( $p < 0.01$ ). Figure 1 illustrates the bias present in Item 1 across the different language groups.

**Table 2: Summary of DIF for the SWLS**

Group	Item	$x_{12}^2$	$x_{13}^2$	$x_{23}^2$	$\beta_1$	$R_{12}^2$	$R_{13}^2$	$R_{23}^2$
Language	Item 1	0.0000	0.0000	0.0000	0.2085	0.0673	0.0893	0.0220
	Item 2	0.7623	0.7867	0.5545	0.0030	0.0003	0.0011	0.0007
	Item 3	0.0000	0.0000	0.1010	0.0483	0.0135	0.0160	0.0025
	Item 4	0.6712	0.1448	0.0488	0.0020	0.0004	0.0036	0.0032
	Item 5	0.5904	0.7034	0.5706	0.0063	0.0005	0.0010	0.0005
Campus	Item 1	0.0000	0.0000	0.0000	0.1201	0.0459	0.0627	0.0167
	Item 2	0.6982	0.8404	0.7035	0.0001	0.0004	0.0008	0.0004
	Item 3	0.1220	0.2405	0.5263	0.0102	0.0031	0.0041	0.0010
	Item 4	0.4245	0.7227	0.8363	0.0013	0.0013	0.0016	0.0003
	Item 5	0.0446	0.1706	0.9105	0.0051	0.0037	0.0039	0.0001
Gender	Item 1	0.0823	0.0202	0.0287	0.0052	0.0015	0.0038	0.0023
	Item 2	0.6606	0.8625	0.7482	0.0000	0.0001	0.0001	0.0000
	Item 3	0.9325	0.4037	0.1789	0.0000	0.0000	0.0008	0.0008
	Item 4	0.9949	0.9358	0.7158	0.0000	0.0000	0.0001	0.0001
	Item 5	0.0239	0.0772	0.8894	0.0004	0.0021	0.0021	0.0000

**Notes:** = chi-square of model 1 compared to model 2;  $x_{13}^2$  = chi-square of model 1 compared to model 3;  $x_{23}^2$  = chi-square of model 2 compared to model 3;  $\beta_1$  = change in beta coefficient;  $R_{12}^2$  = pseudo-Mcfadden  $R^2$  of model 1 compared to model 2;  $R_{13}^2$  = pseudo-Mcfadden  $R^2$  of model 1 compared to model 3;  $R_{23}^2$  = pseudo-Mcfadden  $R^2$  of model 2 compared to model 3

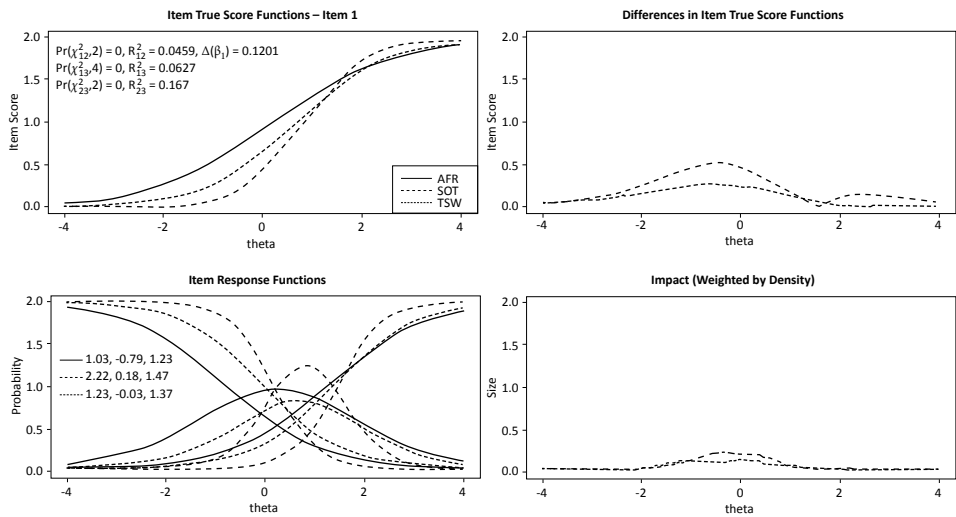


**Figure 1: Graphical display of Item 1, which shows uniform and non-uniform DIF for language groups**



The top left plot in Figure 1 indicates the item-true score functions based on group-specific item parameter estimates. The slope of the function for the Afrikaans group was significantly higher than that of the Sesotho and Setswana groups, indicating non-uniform DIF. The bottom left plot in Figure 1 compared the item response functions across the three language groups, which are noticeably different between the three language groups. The expected impact of DIF on scores is indicated by the top left plot in Figure 1 as the absolute difference between the item true-score functions (Kim et al., 2007). A difference can be seen at approximately  $\theta = 0.50$ ; however, the density-weighted impact (as indicated by the bottom right plot) can be interpreted as small. Even though the effect of bias can be regarded as small when taking the pseudo-McFadden  $R^2$  statistics ( $R^2 < 0.13$ ) into account, the change in beta coefficient is larger than 5% ( $\Delta\beta_1 = 20.85\%$ ), suggesting the effect is practically meaningful.

Item 1 was also problematic for campus groups. Based on the results in Table 2 and the plots in Figure 2, all likelihood-ratio  $\chi^2$  tests were statistically significant ( $< 0.01$ ;  $< 0.01$ ;  $< 0.01$ ), which indicates the presence of both uniform and non-uniform DIF. Similar to the results for DIF between language groups, the results in Table 2 and the graphs in Figure 2 demonstrate that even though the impact of bias can be regarded as small (pseudo-McFadden  $R^2$  statistics  $< 0.13$ ), the change in the beta coefficient ( $\Delta\beta_1 = 12.01\%$ ) indicates that the impact of the bias has a practically meaningful effect.



**Figure 2: Graphical display of Item 1, which shows uniform and non-uniform DIF for campus groups**

In addition, statistically significant bias was detected in Item 3 (Figure 3) across the different language groups, with significant likelihood ratio  $\chi^2$  tests when comparing Models 1 and 2 ( $< 0.01$ ) and Models 1 and 3 ( $< 0.01$ ), indicating mainly uniform bias. Noticeable differences between the language groups can be seen in the plots. However,

regarding the magnitude of these items, the density-weighted impact seen in the bottom right plot and pseudo-McFadden  $R^2$  statistic values  $< 0.13$  and  $\Delta\beta_1$  coefficient smaller than 5% indicate that the significant practical effect is negligible.

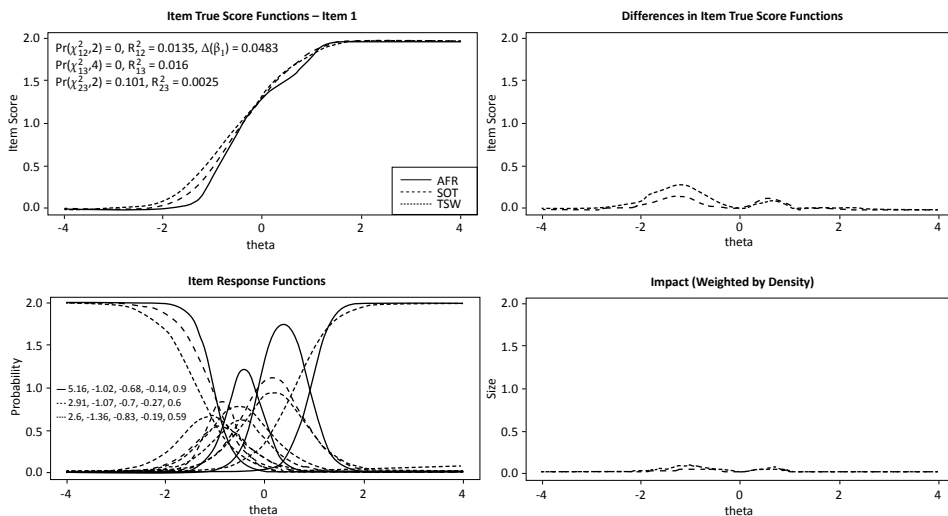


Figure 3: Graphical display of Item 3, which shows uniform DIF for language groups

### Measurement invariance

Table 3 shows the measurement invariance across the language, campus, and gender groups included in this study.

Table 3: Summary of measurement invariance analysis for the SWLS

Group		$\chi^2$	df	CFI	$\Delta$ CFI	RMSEA	$\Delta$ RMSEA
Language	Configural	52.85	20	0.962	—	0.102	—
	Metric	89.58	32	0.933	-0.029	0.106	0.004
	Scalar	178.36	44	0.843	-0.090	0.138	0.032
	Partial metric	60.67	31	0.965	0.003	0.078	-0.024
	Partial scalar	67.39	39	0.967	0.002	0.068	-0.010
Campus	Configural	37.29	15	0.975	—	0.076	—
	Metric	70.70	23	0.946	-0.029	0.090	0.014
	Scalar	154.54	31	0.861	-0.085	0.125	0.035
	Partial metric	47.84	22	0.971	-0.004	0.068	-0.008
	Partial scalar	50.28	25	0.972	0.001	0.063	-0.005
Gender	Configural	41.36	10	0.963	—	0.090	—

Group		$\chi^2$	<i>df</i>	CFI	$\Delta$ CFI	RMSEA	$\Delta$ RMSEA
	Metric	47.89	14	0.960	-0.003	0.079	-0.011
	Scalar	59.98	18	0.950	0.010	0.078	-0.001

**Notes:**  $\chi^2$  = chi-square; *df* = degrees of freedom; CFI = Comparative fit index;  $\Delta$ CFI = delta (change in) CFI; RMSEA = Root mean square error of approximation;  $\Delta$ RMSEA = delta (change in) RMSEA

First, configural invariance was tested. Table 3 indicates that the SWLS has configural invariance, with CFI scores ranging from 0.962 to 0.975, meaning that the scale consists of the same factor structure across all language, campus, and gender groups.

Concerning metric invariance, the results in Table 3 show that the SWLS has metric invariance across the gender groups ( $\Delta$ CFI = -0.003) but not across the language ( $\Delta$ CFI = -0.029) or campus groups ( $\Delta$ CFI = -0.029). By releasing the intercept of Item 1 in the Afrikaans groups and Campus 2, partial metric invariance was achieved for the SWLS across the different language and campus groups.

Concerning scalar invariance, Table 3 indicates that the SWLS achieved scalar invariance across gender groups ( $\Delta$ CFI = 0.010) but not across language ( $\Delta$ CFI = -0.090) and campus groups ( $\Delta$ CFI = -0.085). Therefore, to improve model fit, the intercepts of Item 1 and Item 3 were freed for the Afrikaans group to achieve partial scalar invariance. Similarly, partial scalar invariance was reached across different campus groups by releasing Items 1 and 5 intercepts for Campus 1 and Items 1 and 4 for Campuses 2 and 3.

### Internal consistency

Cronbach's alpha coefficients were calculated as a measure of internal consistency. A Cronbach's alpha coefficient of 0.83 was found for the SWLS, indicating an acceptable internal consistency ( $\alpha \geq 0.70$ ) (Nunnally, 1978). In addition, McDonald's omega ( $\omega$ ) was 0.83, showing good internal consistency (Kline, 2015).

### Discussion

Essentially, any assessment used in a diverse and cross-cultural higher education institution must be tested and analysed to ensure the scale measures the same constructs across diverse groups to be considered fair and unbiased (Hill et al., 2013). Therefore, this study presented preliminary evidence on the psychometric properties of the Satisfaction with Life Scale for first-year university students at a specific South African university, with a particular focus on item bias and invariance (including configural invariance, metric invariance and scalar invariance) between specific language, campus and gender groups.

Before differential item functioning and invariance testing, confirmatory factor analysis was used to provide evidence for a one-factor structure. This indicates that satisfaction with life can be measured as one general factor (Diener et al., 1985).

Differential item functioning was used to test for uniform and non-uniform bias. Item bias was detected in both language and campus groups but not across gender groups. Item 1 ("In most ways my life is close to my ideal") was problematic for language

and campus groups. Uniform bias was present between the Sesotho and Setswana groups, which indicates that the probability of a specific response from these two language groups was systematically higher or lower across all levels of endorsement (Swaminathan & Rogers, 1990; Teresi & Fleishman, 2007). However, non-uniform bias was present in the Afrikaans group, which indicates that at certain levels of endorsement, the relation between the Afrikaans group and the response to the item was dissimilar compared to the Sesotho and Setswana groups (Mellenbergh, 1989; Sireci & Rios, 2013). With regards to campus, the findings indicate that uniform bias is present between Campus 1 and 3, with non-uniform bias being present for Campus 2 – indicating that at specific levels of endorsement, the relation between Campus 2 and the response to the item was dissimilar compared to those for Campus 1 and 3 (Mellenbergh, 1989; Sireci & Rios, 2013).

In addition, uniform bias was observed in Item 3 (“I am satisfied with my life”). Even though this finding implies that the probability of a specific response to this item is found at all trait levels across the three language groups on a statistically significant level, based on the pseudo-McFadden  $R^2$  statistic as well as the changes in beta coefficient, this impact was negligible and therefore not of practical significance (Crane et al., 2007; Teresi & Fleishman, 2007). No bias was detected for any of the SWLS items for gender groups.

**Measurement invariance** included testing for configural, metric, and scalar invariance across the different language, campus and gender groups included in this study (Preti et al., 2013). The results showed that the SWLS has *configural invariance* across the different language, campus, and gender groups, indicating that the same factor structure is present across the groups included in this study.

Regarding **metric invariance**, the SWLS has metric invariance across gender groups but not across language and campus groups. When full invariance was not achieved, partial metric invariance was tested by assessing the factor structure of the SWLS based on changes in CFI ( $\Delta$ CFI) (Clench-Aas et al., 2011). Partial metric invariance was achieved by freeing the loading of Item 1 in both the Afrikaans group and Campus 2. Although some parameters can vary across groups (rejected constraints), at least two intercepts and factor loadings should be equally constrained across groups to make valid inferences (Byrne et al., 1989; Laguna et al., 2017). Therefore, factor loadings can still be fairly compared across language and campus groups with partial metric invariance (Van De Schoot et al., 2012). Full metric invariance was achieved across the gender groups included in this study, which indicates that each item similarly contributes to the latent construct of the SWLS across gender groups (Putnick & Bornstein, 2016).

Evidence was found for **scalar invariance** across gender groups but not across language and campus groups; hence, partial scalar invariance was tested. Constraints were rejected in both language and campus groups. Concerning the language groups, the intercepts of both Item 1 and Item 3 were freed for the Afrikaans group to improve model fit. Regarding the campus groups, partial scalar invariance was achieved by releasing the intercepts of Item 1 and Item 5 for the Campus 1 group as well as Item 1 and Item 4 for Campuses 2 and 3. This implies that fair comparisons across language

and campus groups can still be made (Van De Schoot et al., 2012). Full scalar invariance was achieved across the gender groups, which indicates that the factor loadings and intercepts of the five items of the SWLS can be meaningfully compared across gender groups (Putnick & Bornstein, 2016). These results are in line with other studies, where some items did not appear to be equivalent across cultures. More specifically, with regards to cross-cultural studies, variance was reported for Item 2, Item 3, Item 4 and Item 5 (Atienza González et al., 2016; Dimitrova & Domínguez, 2015; Whisman & Judd, 2016; Zanon et al., 2014).

Cronbach's alpha coefficients and McDonald's omega ( $\omega$ ) were used to test for the **internal consistency** (a measure of reliability) of the Satisfaction with Life Scale. The findings indicate a Cronbach's alpha coefficient of 0.83 and McDonald's omega of 0.83, indicating that the Satisfaction with Life Scale is reliable.

### **Limitations and recommendations**

The findings indicate that across the different language, campus, and gender groups included in the study, bias was detected in both Item 1 (language and campus groups) and Item 3 (language groups). In addition, evidence was provided for configural invariance across all groups but not metric and scalar invariance. As a result, partial, metric and scalar invariance was detected across language and campus groups (probably due to Item 1 being problematic). Although the SWLS has been validated across many countries, languages, and cultures, future researchers should validate the psychometric properties of the SWLS to ensure it is valid and reliable for use across different diverse groups and settings of university students. The present study serves as preliminary evidence of item bias and invariance of the SWLS. However, future research should focus on its differential prediction for different academic outcomes, such as test and academic performance, because the slope and intercept of relations still need to be determined (Berry 2015; Theron, 2006). Furthermore, future research could inform the nomological network of the SWLS by exploring its relationships to other variables related to student well-being, success, and goal commitment (Jonker et al., 2015; Van Lill et al., 2020).

Robust maximum likelihood (MLR) was used as the CFA measurement model's estimation technique and to test the SWLS's invariance based on the small sample sizes in this study. Even though MLR has been introduced into CFA models when the normality assumption of data distribution is moderately violated, as can be the case when using small sample sizes (Knief & Forstmeier, 2021), future studies with sufficient sample sizes could use a weighted least squares with mean- and variance adjusted (WLSMV) method of estimation when the data are of ordinal nature (see Li, 2016). Indeed, the results of the current partial measurement invariance might, with larger sample sizes in some groups, either reach the threshold of non-partial invariance or point out new nuanced differences in the interpretation of items between languages.

### **Practical implications**

Globally, the multicultural nature of populations has become more salient (Van De Vijver & Rothmann, 2004). For example, the increasing demographic diversity in the United

States has been well documented, specifically among the population that does not use English as their native or primary language (Nwosu et al., 2014; Pascarella, 2006). Liu et al. (2019) stipulate that literature regarding language diversity among college students in the United States is scarce and under-researched – the reason being that students with diverse language backgrounds are often combined in the discussion of low-income students, racial minorities, and other underrepresented student groups (Kanno & Cromley, 2013). In addition, non-native English speakers' cultural values and norms are not perfectly aligned with the English-only college environment (Liu et al., 2019).

Through the two mentioned examples, educators and researchers should be cautious when applying any instrument, in our case, the SWLS, to university settings without paying special attention to student diversity and testing for psychometric properties like item bias and invariance. Additionally, the current study only included Afrikaans, Sesotho, and Setswana language groups; therefore, future researchers should include English as a language group for cross-cultural comparisons.

On a practical level, psychologists and practitioners should take great care when applying concepts and instruments from Western countries without testing the applicability of those measures in a diverse setting. Without adequate testing, systematic measurement variability can cause several challenges, including flawed population forecasts, errors in hypothesis testing, planning and implementation of policies, and misguided research on discrepancies (Perkins et al., 2006). Instruments developed in other countries could either be culturally biased, produce inconsistent results when groups are compared, and might be unable to adequately measure a construct when the culture or language differs from the country of origin (Blokland, 2016; Moletsane, 2016; Van De Vijver & Rothmann, 2004). Therefore, it is essential to ensure equivalent measurement before comparing groups or individuals to avoid ambiguous comparisons (Gregorich, 2006; Teresi & Fleishman, 2007).

### **Acknowledgement**

The authors would like to thank Prof. L.T. de Beer for his assistance with the statistical analysis and interpretation of the results.

### **Ethics statement**

The study was approved by the Ethics Committee, Faculty of Economic and Management Sciences (EMS-REC) (Ethics no.: NWU-HS-2014-0165-A4).

### **Potential conflict of interest**

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

### **Funding acknowledgement**

The material described in this article is based on work supported by the office of the Deputy Vice-Chancellor: Teaching and Learning at the North-West University, South Africa; and the National Research Foundation, under reference number RA180103297058

(Grant No.: 118953). The views and opinions expressed in this research are those of the researchers and do not necessarily reflect the opinions or views of the funders.

## References

- Alleyne, M., Alleyne, P., & Greenidge, D. (2010). Life satisfaction and perceived stress among university students in Barbados. *Journal of Psychology in Africa, 20*(2), 291–297. <https://doi.org/10.1080/14330237.2010.10820378>
- Atienza González, F. L., Solá, I. B., Corte-Real, N., & Fonseca, A. M. (2016). Factorial invariance of the Satisfaction with Life Scale in adolescents from Spain and Portugal. *Psicothema, 28*(3), 353–358. <https://doi.org/10.7334/psicothema2016.1>
- Avcu, A. (2021). Item response theory-based psychometric investigation of SWLS for university students. *International Journal of Psychology and Educational Studies, 8*(2), 27–37. <https://dx.doi.org/10.52380/ijpes.2021.8.2.265>
- Băcilă, M., Pop, M. C., Scridon, M. A., & Ciornea, R. (2014). Development of an instrument for measuring student satisfaction in business educational institutions. *Contemporary Priorities in Business Education, 16*(37), 841–856.
- Berry, C. M. (2015). Differential validity and differential prediction of cognitive ability tests: Understanding test bias in the employment context. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 435–463. <https://doi.org/10.1146/annurev-orgpsych-032414-111256>
- Blokland, L. M. (2016). Non-Western (African) views of psychological constructs: Current context of psychological assessment in South Africa. In R. Ferreira (Ed.), *Psychological assessment: Thinking innovatively in context of diversity* (pp. 37–51). JUTA.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing, 1*(1), 55–86. [https://doi.org/10.1207/S15327574IJT0101\\_4](https://doi.org/10.1207/S15327574IJT0101_4)
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. DOI:10.1037/0033-2909.105.3.456
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cut-off points in RMSEA test statistic in structural equation models. *Sociology Methods & Research, 36*(4), 462–464. <https://doi.org/10.1177/0049124108314720>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indices for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). **lordif**: An R Package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software, 39*(8), 1–30.
- Chow, H. P. H. (2005). Life satisfaction among university students in a Canadian prairie city: A multivariate analysis. *Social Indicators Research, 70*(2), 139–150. <https://doi.org/10.1007/s11205-004-7526-0>
- Cleary, T. A., & Hilton, T. L. (1968). An investigation of item bias. *Educational and Psychological Measurement, 28*(1), 61–75. <https://doi.org/10.1177/001316446802800106>
- Clench-Aas, J., Nes, R. B., Dalgard, O. S., & Aarø, L. E. (2011). Dimensionality and measurement invariance in the Satisfaction with Life Scale in Norway. *Quality of Life Research, 20*(8), 1307–1317. <https://doi.org/10.1007/s11136-011-9859-x>

- Cortina, J. M., Sheng, Z., Keener, S. K., Keeler, K. R., Grubb, L. K., Schmitt, N., Tonidandel, S., Summerville, K. M., Heggstad, E. D., & Banks, G. C. (2020). From alpha to omega and beyond! A look at the past, present, and (possible) future of psychometric soundness in the Journal of Applied Psychology. *Journal of Applied Psychology, 105*(12), 1351–1381. <https://doi.org/10.1037/apl0000815>
- Crane, P. K., Gibbons, L. E., Ocepek-Welikson, K., Cook, K., Cella, D., Narasimhalu, K., Hays, R. D., & Teresi, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research, 16*(1), 69–84. <https://doi.org/10.1007/s11136-007-9185-5>
- Crane, P. K., Van Belle, G., & Larson, E. B. (2004). Test bias in a cognitive test: Differential item functioning in the CASI. *Statistics in Medicine, 23*(2), 241–256. <https://doi.org/10.1002/sim.1713>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49*(1), 71–75. [https://doi.org/10.1207/s15327752jpa4901\\_13](https://doi.org/10.1207/s15327752jpa4901_13)
- Dimitrova, R., & Domínguez, A. (2015). Measurement invariance of the Satisfaction with Life Scale in Argentina, Mexico and Nicaragua. *Social Inquiry into Well-being, 1*, 32–39.
- Durak, M., Senol-Durak, E., & Gencoz, T. (2010). Psychometric properties of the Satisfaction with Life Scale among Turkish university students, correctional officers, and elderly adults. *Social Indicators Research, 99*(3), 413–429. <https://doi.org/10.1007/s11205-010-9589-4>
- 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*(9), 2251–2264. <https://doi.org/10.1007/s11136-017-1552-2>
- Eснаоla, I., Benito, M., Antonio-Agirre, I., Xpe, I., & Lorenzo, M. (2019). Longitudinal measurement invariance of the Satisfaction with Life Scale in adolescence. *Quality of Life Research, 28*(10), 2831–2837. <https://doi.org/10.1007/s11136-019-02224-7>
- Foxcroft, C., & Roodt, G. (2009). *Introduction to psychological assessment in the South African context* (3rd ed.). Oxford University Press.
- Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction with Life Scale (SWLS): Psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment, 27*(2), 127–132. <https://doi.org/10.1027/1015-5759/a000058>
- Gökalp, M., & Topal, T. (2019). Investigation of life satisfaction of university students according to various variables. *The Turkish Online Journal of Educational Technology, 2*, 191–204.
- Gouveia, V. V., Milfont, T. L., da Fonseca, P. N., & de Miranda Coelho, J. A. P. (2009). Life satisfaction in Brazil: Testing the psychometric properties of the Satisfaction with Life Scale (SWLS) in five Brazilian samples. *Social Indicators Research, 90*(2), 267–277. <https://doi.org/10.1007/s11205-008-9257-0>
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(11 Suppl 3), S78–94. <https://doi.org/10.1097/01.mlr.0000245454.12228.8f>
- He, J., & Van De Vijver, F. (2012). Bias and equivalence in cross-cultural research. *Online Readings in Psychology and Culture, 2*(2). <https://doi.org/10.9707/2307-0919.1111>
- Hill, C., Nel, J. A., Van de Vijver, F. J. R., Meiring, D., Valchev, V. H., Adams, B. G., & De Bruin, G. P. (2013). Developing and testing items for the South African Personality Inventory (SAPI). *SA Journal of Industrial Psychology, 39*(1), 13 pages. <https://doi.org/10.4102/sajip.v39i1.1122>



- Hinz, A., Conrad, I., Schroeter, M. L., Glaesmer, H., Brähler, E., Zenger, M., Kocalevent, R.-D., & Herzberg, P. Y. (2018). Psychometric properties of the Satisfaction with Life Scale (SWLS), derived from a large German community sample. *Quality of Life Research, 27*(6), 1661–1670. <https://doi.org/10.1007/s11136-018-1844-1>
- Hu, L. T., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- 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*(5), 1070–1079. <https://doi.org/10.1016/j.paid.2007.10.030>
- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I Error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*(4), 329–349. [https://doi.org/10.1207/S15324818AME1404\\_2](https://doi.org/10.1207/S15324818AME1404_2)
- Jonker, C. S., Koekemoer, E., & Nel, J. A. (2015). Exploring a positive SWB model in a sample of university students in South Africa. *Social Indicators Research, 121*(3), 815–832. <https://doi.org/10.1007/s11205-014-0658-y>
- Jovanović, V. (2019). Measurement invariance of the Serbian version of the Satisfaction with Life Scale across age, gender, and time. *European Journal of Psychological Assessment, 35*(4), 555–563. <https://doi.org/10.1027/1015-5759/a000410>
- Jovanović, V., & Brdar, I. (2018). The cross-national measurement invariance of the Satisfaction with Life Scale in a sample of undergraduate students. *Personality and Individual Differences, 128*, 7–9. <https://doi.org/10.1016/j.paid.2018.02.010>
- Kanno, Y., & Cromley, J. (2013). English language learners' access to and attainment in postsecondary education. *TESOL Quarterly, 47*, 89–121.
- Kim, S. H., Cohen, A. S., Alagoz, C., & Kim, S. (2007). DIF detection effect size measures for polytomously scored items. *Journal of Educational Measurement, 44*(2), 93–116.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Knief, U., & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods, 53*(6), 2576–2590. <https://doi.org/10.3758/s13428-021-01587-5>
- Laguna, M., Mielniczuk, E., Razmus, W., Moriano, J. A., & Gorgievski, M. J. (2017). Cross-culture and gender invariance of the Warr (1990) job-related well-being measure. *Journal of Occupational and Organizational Psychology, 90*, 117–125. <https://doi.org/10.1111/joop.12166>
- Laher, S. (2008). Structural equivalence and the Neo-Pi-R: Implications for the applicability of the five-factor model of personality in an African context. *SA Journal of Industrial Psychology, 34*(1), 76–80. <https://doi.org/10.4102/sajip.v34i1.429>
- Laher, S., & Cockcroft, K. (2013). Current and future trends in psychological assessment in South Africa: Challenges and opportunities. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 535–552). Wits University Press.
- Laher, S., & Cockcroft, K. (2014). Psychological assessment in post-apartheid South Africa: The way forward. *South African Journal of Psychology, 44*(3), 303–314. <https://doi.org/10.1177/0081246314533634>
- Lewis, A. D., Huebner, E. S., Malone, P. S., & Valois, R. F. (2011). Life satisfaction and student engagement in adolescents. *Journal of Youth and Adolescence, 40*(3), 249–262. <https://doi.org/10.1007/s10964-010-9517-6>

- Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Liu, J., Hu, S., & Pascarella, E. T. (2019). Are non-native English speaking students disadvantaged in college experiences and cognitive outcomes? *Journal of Diversity in Higher Education*, 14(3), 398–407. <http://dx.doi.org/10.1037/dhe0000164>
- Lorenzo-Seva, U., Calderon, C., Ferrando, P. J., del Mar Muñoz, M., Beato, C., Ghanem, I., Castelo, B., Carmona-Bayonas, A., Hernández, R., & Jiménez-Fonseca, P. (2019). Psychometric properties and factorial analysis of invariance of the Satisfaction with Life Scale (SWLS) in cancer patients. *Quality of Life Research*, 28(5), 1255–1264. <https://doi.org/10.1007/s11136-019-02106-y>
- Maldonado, G., & Greenland, S. (1993). Simulation study of confounder-selection strategies. *American Journal of Epidemiology*, 138(11), 923–936. <https://doi.org/10.1093/oxfordjournals.aje.a116813>
- Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33(2), 181–220. [https://doi.org/10.1207/s15327906mbr3302\\_1](https://doi.org/10.1207/s15327906mbr3302_1)
- McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, 15(1), 28–50. <https://doi.org/10.1177/1088868310366253>
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, 13(2), 127–143. [https://doi.org/10.1016/0883-0355\(89\)90002-5](https://doi.org/10.1016/0883-0355(89)90002-5)
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111–130. <https://doi.org/10.21500/20112084.857>
- Moksnes, U. K., Løhre, A., Byrne, 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(2), 657–671. <https://doi.org/10.1007/s11205-013-0451-3>
- Moletsane, M. (2016). Understanding the role of indigenous knowledge in psychological assessment and intervention in a multicultural South African context. In R. Ferreira (Ed.), *Psychological assessment: Thinking innovatively in the context of diversity* (pp. 20–36). JUTA.
- Morton, N., Hill, C., Meiring, D., & Van De Vijver, F. J. (2019). Investigating measurement invariance in the South African Personality Inventory: English version. *South African Journal of Psychology*, 50(2), 274–289. <https://doi.org/10.1177/0081246319877537>
- Muthén, L. K., & Muthén, B. O. (2021). *Mplus user's guide* (8th ed.). [https://www.statmodel.com/download/usersguide/MplusUserGuideVer\\_8.pdf](https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf)
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.) McGraw-Hill.
- Nwosu, C., Batalova, J., & Auclair, G. (2014). *Frequently requested statistics on immigrants and immigration in the United States*. Migration Policy Institute. <http://www.migrationpolicy.org/article/frequently-requested-statistics-immigrants-and-immigration-unitedstates>
- Pascarella, E. T. (2006). How college affects students: Ten directions for future research. *Journal of College Student Development*, 47, 508–520.
- Pavot, W., & Diener, E. (1993). The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research*, 28, 1–20.
- Perkins, A. J., Stump, T. E., Monahan, P. O., & McHorney, C. A. (2006). Assessment of differential item functioning for demographic comparisons in the MOS SF-36 health survey. *Quality of Life Research*, 15(3), 331–348. <https://doi.org/10.1007/s11136-005-1551-6>

- Pons, D., Aienza, 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, 87*, 519–529. <https://doi.org/10.2466/pms.2000.91.1.62>
- President of the Republic of South Africa. (1998, October 19). Employment Equity Act (Act no. 55 of 1998). *Government Gazette*, 19370, p. 16. [https://www.gov.za/sites/default/files/gcis\\_document/2014/09/a55-980.pdf](https://www.gov.za/sites/default/files/gcis_document/2014/09/a55-980.pdf)
- Preti, A., Vellante, M., Gabbriellini, M., Lai, V., Muratore, T., Pintus, E., Pintus, M., Sanna, S., Scanu, R., Tronci, D., Corrias, I., Petretto, D. R., & Carta, M. G. (2013). Confirmatory factor analysis and measurement invariance by gender, age and levels of psychological distress of the short TEMPS-A. *Journal of Affective Disorders, 151*, 995–1002.
- Pretorius, T. B., & Padmanabhanunni, A. (2022). Assessing the cognitive component of subjective well-being: Revisiting the Satisfaction with Life Scale with classical test theory and item response theory. *African Journal of Psychological Assessment, 4*(0), a106. <https://doi.org/10.4102/ajopa.v4i0.106>
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71–90. <https://dx.doi.org/10.1016/j.dr.2016.06.004>
- Revicki, D. (2014). Internal consistency reliability. In A. C. Michalos (Ed.), *Encyclopedia of quality of life and well-being research*. (pp. 3305–3306). Springer. [https://doi.org/10.1007/978-94-007-0753-5\\_1494](https://doi.org/10.1007/978-94-007-0753-5_1494)
- Rode, J. C., Arthaud-Day, M. L., Mooney, C. H., Near, J. P., Baldwin, T. T., Bommer, W. H., & Rubin, R. S. (2005). Life satisfaction and student performance. *Academy of Management Learning & Education, 4*(4), 421–433. <https://doi.org/10.5465/amle.2005.19086784>
- Ruiz, F. J., Suárez-Falcón, J. C., Flórez, C. L., Odriozola-González, P., Tovar, D., López-González, S., & Baeza-Martin, R. (2019). Validity of the Satisfaction with Life Scale in Colombia and the factorial equivalence with Spanish data. *Revista Latinoamericana de Psicología, 51*(2), 58–65. <http://dx.doi.org/10.14349/rlp.2019.v51.n2.1>
- Safak-Ayvazoglu, A., & Kunuroglu, F. (2019). Acculturation experiences and psychological well-being of Syrian refugees attending university in Turkey: A qualitative study. *Journal of Diversity in Higher Education, 14*(1), 96–109. <http://dx.doi.org/10.1037/dhe0000148>
- Schaap, P. (2011). The differential item functioning and structural equivalence of a nonverbal cognitive ability test for five language groups. *SA Journal of Industrial Psychology, 37*(1), 1–16. <https://doi.org/10.4102/sajip.v37i1.881>
- Schutte, L., Negri, L., Delle Fave, A., & Wissing, M. P. (2021). Rasch analysis of the Satisfaction with Life Scale across countries: Findings from South Africa and Italy. *Current Psychology, 40*(10), 4908–4917. <https://doi.org/10.1007/s12144-019-00424-5>
- Serin, N. B., Serin, O., & Özbaş, L. F. (2010). Predicting university students' life satisfaction by their anxiety and depression level. *Procedia Social and Behavioural Sciences, 9*, 579–582. <https://doi.org/10.1016/j.sbspro.2010.12.200>
- Shevlin, M., Brunson, V., & Miles, J. N. V. (1998). Satisfaction with Life Scale: Analysis of factorial invariance, mean structures and reliability. *Personality and Individual Differences, 25*(5), 911–916. [https://doi.org/10.1016/S0191-8869\(98\)00088-9](https://doi.org/10.1016/S0191-8869(98)00088-9)
- Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement, 79*(2), 310–334. <https://doi.org/10.1177/0013164418783530>

- Sireci, S. G., & Rios, J. A. (2013). Decisions that make a difference in detecting differential item functioning. *Educational Research and Evaluation, 19*(2–3), 170–187. <https://doi.org/10.1080/13803611.2013.767621>
- Steiger, J. H. (1989). *EzPATH: A supplementary module for SYSTAT and SYGRAPH*. Systat, Inc.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*(4), 361–370. <https://www.jstor.org/stable/1434855>
- Teresi, J. A., & Fleishman, J. A. (2007). Differential item functioning and health assessment. *Quality of Life Research, 16*, 33–42. <https://doi.org/10.1007/s11136-007-9184-6>
- Theron, C. (2007). Confessions, scapegoats, and flying pigs: Psychometric testing and the law. *SA Journal of Industrial Psychology, 33*(1), 102–117. <https://doi.org/10.4102/sajip.v33i1.260>
- 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>
- Van De Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology, 9*(4), 1–7. <http://dx.doi.org/10.1080/17405629.2012.686740>
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology, 6*(1064), 1–4. <https://doi.org/10.3389/fpsyg.2015.01064>
- Van De Vijver, A. J. R., & Rothmann, S. (2004). Assessment in multicultural groups: The South African case. *SA Journal of Industrial Psychology, 30*(4), 1–7. <https://doi.org/10.4102/sajip.v30i4.169>
- Van Lill, X., Roodt, G., & de Bruin, G. P. (2020). Is there a general factor in goal commitment? *SA Journal of Industrial Psychology, 46*, a1765. <https://doi.org/10.4102/sajip.v46i0.1765>
- Wach, F.-S., Karbach, J., Ruffling, S., Brücken, R., & Spinath, F. M. (2016). University students' satisfaction with their academic studies: Personality and motivation matter. *Frontiers in Psychology, 7*(55), 1–12. <https://doi.org/10.3389/fpsyg.2016.00055>
- Whisman, M. A., & Judd, C. M. (2016). A cross-national analysis of measurement invariance of the Satisfaction with Life Scale. *Psychological Assessment, 28*(2), 239–244.
- Wu, C., & Yao, G. (2006). Analysis of factorial invariance across gender in the Taiwan version of the Satisfaction with Life Scale. *Personality and Individual Differences, 40*(6), 1259–1268. <https://doi.org/10.1016/j.paid.2005.11.012>
- Wu, C.-H., Chen, L. H., & Tsai, Y.-M. (2009). Longitudinal invariance analysis of the Satisfaction with Life Scale. *Personality and Individual Differences, 46*(4), 396–401. <https://doi.org/10.1016/j.paid.2008.11.002>
- Zanon, C., Bardagi, M. P., Layous, K., & Hutz, C. S. (2014). Validation of the Satisfaction with Life Scale to Brazilians: Evidences of measurement noninvariance across Brazil and US. *Social Indicators Research, 119*, 443–453. <https://doi.org/10.1007/s11205-013-0478-5>
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modelling as a unitary framework for binary and Likert-type (ordinal) item scores*. Directorate of human resources research and evaluation, Department of National Defence.

#### How to cite:

van Rensburg, C., & Mostert, K. (2023). Can life satisfaction be measured fairly for different groups of South African first-year students? Testing the Satisfaction with Life scale. *Journal of Student Affairs in Africa, 11*(1), 109–128. DOI: 10.24085/jsaa.v11i1.4327