Psychology Research and Behavior Management Dovepress

Open Access Full Text Article ORIGINAL RESEARCH

Factorial and Criterion Validities of the Chinese Version of Rosenberg Self-Esteem Scale Among Undergraduate Students

Wanrong Peng[1,*], You Zhou[1,2,*], Jun Chu[1], Zhaoxia Liu[1], Kaili Zheng[1], Shuqiao Yao[1,3,4,†], Jinyao Yi[1,3,4]

[1]Medical Psychological Center, the Second Xiangya Hospital, Central South University, Changsha, 410011, People's Republic of China; [2]Xiangya School of Medicine, Central South University, Changsha, 410013, People's Republic of China; [3]Medical Psychological Institute, Central South University, Changsha, 410011, People's Republic of China; [4]National Clinical Research Center for Mental Disorders, Changsha, 410011, People's Republic of China

*Wanrong Peng and You Zhou contributed equally to this work

†Prof. Dr. Shuqiao Yao† passed away on 20 January 2024

Correspondence: Jinyao Yi, Email jinyaoyi@csu.edu.cn

Psychology Research and Behavior Management downloaded from https://www.dovepress.com/
For personal use only.

Purpose: To systematically investigate the factorial and criterion validity of the Chinese version of the Rosenberg Self-Esteem Scale (RSES) among undergraduate students.

Methods: Altogether 1415 undergraduates completed the RSES, the Short Form of Mood and Anxiety Symptoms Questionnaire, the Perceived Stress Scale and the Quality of Life Enjoyment and Satisfaction Questionnaire. Confirmatory factor analysis was conducted to examine the factorial structure, gender invariance, and latent mean invariance of RSES. An independent *t*-test was employed to compare differences in MASQ-SF scores between the high- and low-level self-esteem groups, and a path analysis was performed to confirm the vulnerability model.

Results: Results of Confirmatory Factor Analysis showed that the correlated trait-correlated method model, which includes both global self-esteem and two positive and negative factors, fits the data best (optimum model). The results of the Multigroup Confirmatory Factor Analyses demonstrated that this factor structure achieved measurement invariance and latent mean equivalence across genders among university students. Independent sample *t*-tests revealed that undergraduates with higher self-esteem exhibited lower levels of anxiety and depression than those with lower self-esteem did. Path analysis showed that high self-esteem mitigated depressive symptoms induced by stress, thereby enhancing life satisfaction.

Conclusion: The Chinese version of the RSES exhibits robust factorial and excellent criterion validity, making it a suitable tool for assessing self-esteem among undergraduate students to promote mental and physical well-being.

Keywords: Rosenberg self-esteem scale, factorial structure, gender invariance, criterion validity

Introduction

Self-esteem, encompassing self-worth and self-confidence, refers to one's positive or negative attitude towards oneself and is one of the most critical traits in estimate self-concept.[1] It is a pivotal trait for assessing self-concept, and serves as a crucial indicator of mental health and subjective well-being.[2] From the perspective of positive psychology, which is considered a vital protective and promotive factor for mental health, individuals with higher self-esteem tend to report greater self-acceptance and satisfaction. Conversely, various theories have identified low self-esteem as a contributing factor to depressive disorders.[3]

Several self-report questionnaires and interviews have been developed to effectively measure self-esteem. Among the questionnaires, the Rosenberg Self-Esteem Scale (RSES), the Self-Perception Profile, and the Self-Esteem Inventory are notable examples. Interview-based measures included self-evaluation, social support, and personal interviews. In particular, RSES has gained widespread usage owing to its efficiency and ease of administration.[1,4] Initially designed

Received: 3 September 2024
Accepted: 21 November 2024
Published: 5 December 2024

© 2024 Peng et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

to gauge the global self-esteem of adolescents,[1] the RSES has demonstrated acceptable reliability and validity across diverse age groups and countries, suggesting its utility in assessing self-esteem across a wide spectrum of individuals.[5]

Early research by Rosenberg considered the RSES to be unidimensional[1] (see Figure 1, RSES-GSE Model). Subsequent studies, however, have revealed a bi-factorial structure that distinguishes between self-confidence and self-deprecation based on positively and negatively wording methods[6] (see Figure 1, RSES-PN Model). Meanwhile, the wording effects introduced by positively and negatively keyed items would further result in method effects, that is, systematic errors related to the measurement process rather than to the scale structure, which then yield undesirable systematic variation and weaken the measurement of the constructs. To address the problems caused by method effects, correlated trait-correlated method (CTCM) and correlated trait-correlated uniqueness (CTCU) models derived from the multitrait-multimethod (MTMM) framework have been proposed.[7,8] Specifically, CTCM models treat items with the same wording method as a latent variable, which works as an independent factor along with the content factor.[7] Therefore, CTCM models allow the estimation of the variance arising from method effects, yet may carry the risk of the model not being convergent. CTCU models, instead, correlate the residuals of items with the same wording method to each other, which helps CTCU models converge well,[9] but cannot estimate the variance resulting from wording effects.[7]

Using confirmatory factor analysis (CFA), researchers have tested the construct validity of the RSES, with various models receiving support. The CTCU model with negatively worded items (see Figure 1, CTCU-N Model) showed a better fit, whereas Mullen and his colleagues[10] supported the CTCU model with positive item wording (see Figure 1, CTCU-P Model). In contrast, Gomez-Lugo[11] and Vasconcelos-Raposo[12] preferred the CTCU model with positive and negative items (see Figure 1, CTCU-PN Model). Boduszek regarded the CTCM model with positively and negatively worded items (see Figure 1, CTCM-PN Model) as the best fitting one.[13] In contrast, other studies have approved the bi-factor model of the RSES with positive and negative factors[8,11] (see Figure 1, RSES-PN Model) and the bi-factor model with only positive or negative items.[6] Overall, these competitive models of RSES with both pros and cons were supported by corresponding studies, indicating that

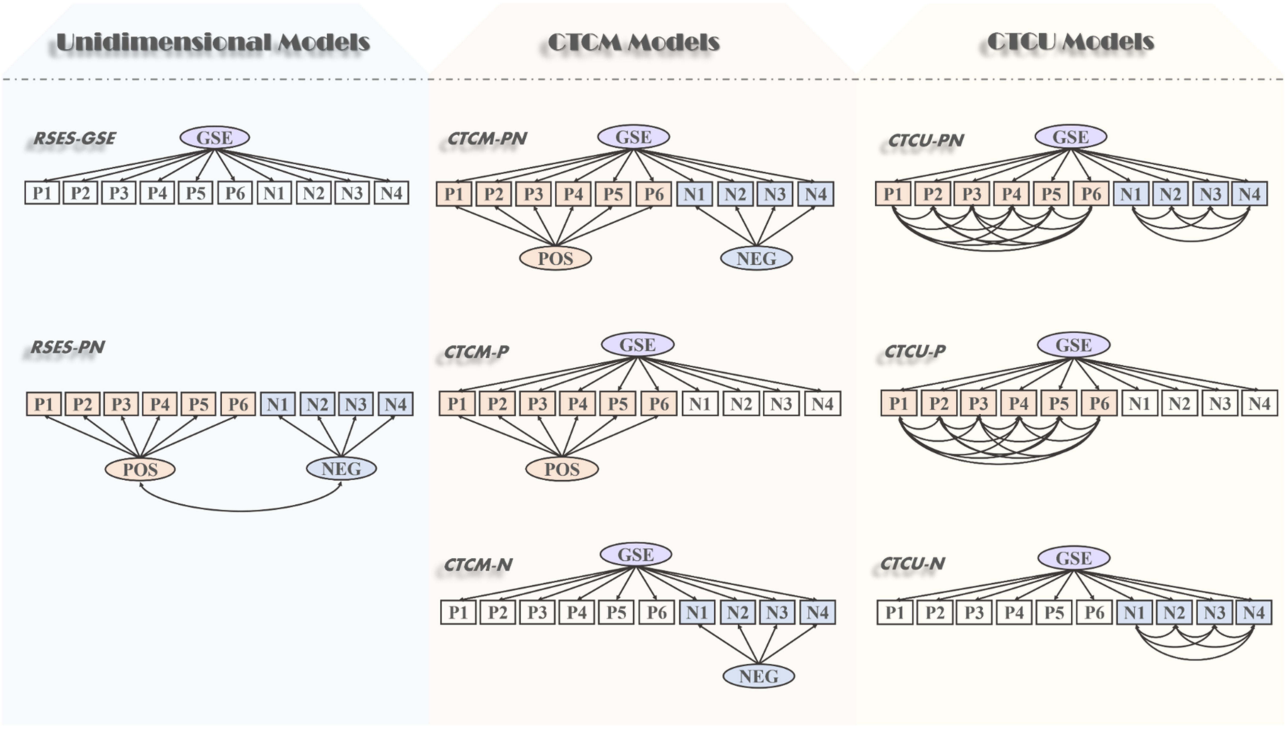

Figure 1 Eight Alternative Models of RSES. RSES-GSE Model = a unidimensional model with a global self-esteem (GSE) factor; RSES-PN Model = two separate factors representing positively and negatively worded items; CTCM-PN Model= correlated traits-correlated method (CTCM) bi-factor model containing a GSE factor and two respective models describing positive and negative method effects; CTCM-P Model = CTCM model with a GSE factor and only a positive factor of method effect. CTCM-N Model = CTCM model with a GSE factor and only a negative factor of method effect; CTCU-PN Model = correlated uniqueness (CTCU) model including one trait factor and correlated uniqueness among positive and negative subscales; CTCU-P Model = CTCU model including one trait factor and correlated uniqueness among positively worded items; CTCU-N Model = CTCU model including one trait factor and correlated uniqueness among negatively worded items.

the factorial structure of the RSES is contaminated to some extent by method effects.[7] Hence, it is necessary to examine all the above models by performing CFA to compare the fitting of each model and obtain the best model for future studies. In addition to the inconformity of RSES dimensionality, whether gender differences exist in self-esteem remains controversial. Support for their existence has been provided by several studies employing the unidimensional model,[14] bi-factor model,[8] and CTCU models with positive items. Other studies, however, yielded no significant differences in self-esteem across genders using unidimensional and CTCM models with negatively worded items.[15] As suggested by prior research, the method effects associated with the wording of items may vary in populations or groups, for example, between males and females. For example, DiStefano and Motl[15] found that gender differences in RSES total scores might be related to different responses to negatively worded items. Therefore, when comparing intergroup differences in the RSES, models controlling for method effects would be preferred first. Further comparisons of latent mean differences should be conducted after cross-group measurement invariance of the model has been established, while directly comparing the observed scores may still suffer from systematic errors.

The establishment of measurement invariance in the factorial structure is a prerequisite for the subsequent analysis of group differences in factor scores, as it confirms that the attributes measured by the scale are invariant across groups.[9] Therefore, the gender invariance of the factorial structure and model of the RSES should be confirmed before comparing gender differences, based on which the apparent difference between males and females in self-esteem can truly reflect the discrepancy in the evaluative component of self-concept.[16] Some previous studies have focused on gender invariance in the RSES; however, their results remain blended. Several studies found that the RSES was gender-invariant in undergraduates for both trait and method factors using the CTCM model with negatively worded items,[15,17] while some studies reported that the CTCM model of positive and negative factors showed a strong invariance across gender.[16] One study found that the CTCU model with positively and negatively worded items had partially weak invariance across gender.[12] Blended results may stem from differences in the model selection and sample size, warranting further investigation.

Since its translation into Chinese, the RSES has been widely applied in China to evaluate the self-esteem level of individuals and its relationship with other psychological variables.[18] Several studies have examined its factorial structure and compared alternative models, supporting the bi-factor model as the best fit in Chinese populations, such as adolescents and patients with acute coronary syndrome patients.[19] However, only one study has examined the measurement invariance among Chinese urban and rural adolescents.[19]

As stated above, low self-esteem is a significant clinical risk factor for depression. The vulnerability model of depression, a prominent framework for elucidating the link between self-esteem and depression, posits that low self-esteem is a stable personality trait that predisposes individuals to experience depressive symptoms. It shapes an enduring susceptibility to mental health challenges,[20,21] suggesting that individuals with low self-esteem are not only directly more likely to develop depression but also have an increased sensitivity to stress, thereby heightening the likelihood of depressive reactions under adverse conditions. Empirical evidence also supports the roles of self-esteem and depressive symptoms in bridging the relationship between stress and life satisfaction.[22] Additionally, the cross-lagged model has revealed that self-esteem not only mediates these effects but also predicts life satisfaction over time, highlighting its dynamic influence on well-being trajectories.[14] However, the specific mechanisms through which self-esteem interacts with stress, depression, and life satisfaction remain unclear. To address this gap, this study aimed to verify the criterion validity of the RSES by re-examining the complex interplay among self-esteem, stress, depression and life satisfaction.

In summary, although the RSES has been widely used to estimate individuals' self-esteem levels in China, no studies have systematically evaluated its factorial validity, let alone gender invariance. Hence, using a Chinese undergraduate sample, this study aimed to verify the factorial and criterion validities of the RSES by (1) comparing the factorial structures of eight alternative models of the RSES and choosing the optimum fitting model using CFA; (2) testing the gender invariance of the optimum model of the RSES using multigroup CFA; (3) verifying whether the RSES has acceptable internal consistency reliability; and (4) examining the criterion validity of the RSES by comparing the group differences in anxiety and depression (high self-esteem vs low self-esteem) and by verifying the vulnerable model of depression.

Materials and Methods

Procedure and Participants

A priori power analysis was conducted to determine the sufficient sample size for the intended statistical analyses with G*power software.[23] It was estimated that a total of 436 participants would be necessary to achieve a statistical power of 95% for the Goodness-of-Fit tests (39 degrees of freedom).

Participants were recruited from a university in Hunan Province, mainland China, using convenience sampling. To ensure a diverse and representative sample, this research included participants from all the four undergraduate year levels, ranging from first year to senior students. The administration of data collection was conducted under the supervision of two psychology graduate students per class. All participants acknowledged and signed the informed consent form at the beginning of the study.

A total of 1493 participants were enrolled in this research. After excluding those who failed to complete all items and those suspected of random responding, a total of 1415 participants (mean age = 20.52 ± 1.03), including 696 males (mean age = 20.71 ± 0.98) and 719 females (mean age = 20.34 ± 1.04), were included in the analysis. Their scores for subjective economic status (mean = 5.52) and subjective social status (mean = 6.15) both ranged from 1 to 10 and were normally distributed.

This research complies with the ethical standards of the Declaration of Helsinki and has been approved by the Ethics Committee of the Second Xiangya Hospital of Central South University.

Measure

The Rosenberg's Self-Esteem Scale (RSES)

The RSES is a 10-item self-report scale that measures self-esteem level.[1] Each item is scored on a 4-point Likert scale ranging from 1 (totally disagree) to 4 (totally agree). Notably, the original RSES contains five positive wording items and five negative wording items, of which Item 8 was scored in reverse. However, considering the cultural setting, item 8 of the Chinese version of the RSES was scored positively, since the Chinese tended to understand it in a positive way.[24] Therefore, in the Chinese version of the RSES, Items 3, 5, 9, and 10 are scored in reverse, so that higher scores on these four items indicate lower self-respect levels. The total RSES score ranges from 10 to 40, with higher scores indicating positive self-esteem or a general perception of self-worth.

The Short Form of Mood and Anxiety Symptoms Questionnaire (MASQ-SF)

MASQ-SF was used to evaluate a range of symptoms related to depression and anxiety.[25] The MASQ-SF includes 62 items, with the score for each item ranging from one (strong disagreement) to five (extreme agreement). The items were divided into four subscales: General Distress (GD), Anhedonic Depression (AD), General Anxiety (GA), and Anxious Arousal (AA). The Chinese version of the MASQ-SF has been shown to have adequate psychometric properties. In this study, the Cronbach's α α values of the MASQ-SF was 0.94 in the total sample, 0.94 in males and 0.93 in females, respectively.

Perceived Stress Scale (PSS)

PSS was employed to evaluate the level of perceived stress during the past month. It consists of 10 items rated on a Likert 5-point scale (0–4) with a score range of 0–40, with higher scores reflecting a higher level of perceived stress. Cronbach's α of the PSS was 0.73 in the whole sample, 0.74 in males and 0.72 in females.

Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form (Q-LES-Q-SF)

The Q-LES-Q-SF is a 16-item self-report questionnaire that evaluates overall enjoyment and satisfaction experienced in the past week.[26,27] The responses of the Q-LES-Q-SF were scored on a 5-point scale (1–5), where the first 14 items summed the total score, while the 2 last items involving medications and overall life satisfaction were considered independently. Higher total scores indicate better life satisfaction (ranging from to 14–70). The Cronbach's α α of the Q-LES-Q-SF was 0.88 in the whole sample, 0.88 in males and 0.87 in females.

Statistical Analysis

Statistical analyses were conducted using SPSS 26.0, and Mplus 7.4. The skewness and kurtosis values for each item of RSES, PSS, Q-LES-Q-SF and majority items of MASQ-SF were within acceptable limits, with absolute values less than 1, indicating that the data of these items were normally distributed. The tolerance values (0.73 ~ 0.86) and variance inflation factor (VIF) values (1.16 ~ 1.37), calculated based on total scores of each scale, were within acceptable ranges (tolerance > 0.1, VIF < 10), indicating no risk of multicollinearity. Additionally, Harman's single factor test, conducted using all items across four scales, revealed 18 factors with eigenvalues greater than 1 before rotation, and the first factor explained 18.20% of the variance, which was below the critical threshold of 40%, suggesting that common method bias was not a severe concern in the data for this study.

A series of Robust Maximum Likelihood CFAs were first conducted to examine the factorial structure of the eight competitive models of the RSES. As shown in Figure 1, the *RSES-GSE Model* is unidimensional with a global self-esteem (GSE) factor. The *RSES-PN Model* has two relevant factors that represent positively and negatively worded items, respectively. The *CTCM-PN Model* is a CTCM bi-factor model containing a GSE factor and two respective factors that describe the positive and negative method effects. The *CTCM-P Model* and *CTCM*-N *Model* are CTCM bi-factor models with a GSE factor and only a positive OR negative factor for the method effect. The *CTCU-PN Model* is a CTCU model included one trait factor and correlated uniqueness among positively and negatively worded items. The *CTCU-P Model* and *CTCU*-N *Model* are CTCU models with trait factors and correlated uniqueness among positively or negatively worded items, respectively. The variance of factors was restricted to 1 in all models, and factors were forbidden from being correlated in the CTCM-PN, CTCM-P, and CTCM-N models. All factor loadings were freely estimated. The following indices were selected to evaluate the model fit: ratio of χ^2 to df (χ^2/df), Comparative Fit Index (CFI), Tucker Lewis Index (TLI), root mean square error of approximation (RMSEA), Standardized Root Mean Squared Residuals (SRMR), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Generally, χ^2/df < 5, CFI > 0.90, TLI > 0.90, RMSEA < 0.08, SRMR < 0.08, and AIC and BIC values that were as small as possible were considered acceptable.[28] The best-fit model was selected by comparing the fitting indices.

Second, the gender invariance and latent means invariance of the factorial structure of the RSES were tested with the best fitting model using Robust Maximum Likelihood multigroup CFAs. The baseline model was set with no constraint for configural invariance, the weak invariance model was configured to have invariant factor loadings between groups to confirm that the RSES shared identical structure and meaning across gender; the strong invariance model added the intercept invariance of items while the weak invariance was satisfied, and the strict invariance model then imposed the residual invariance of items upon the assumption that the strong invariance was held. A successful demonstration of these models would yield measurement invariance across sexes. Models of variance-covariance invariance were then examined to test for factorial invariance. Subsequently, latent mean invariance was examined to determine whether the latent means were equal across the samples. If all the models were established, the optimum model of the RSES was confirmed to have factor invariance. Model fit was assessed using several indices: BIC, χ^2, χ^2/df, CFI, TLI, and RMSEA with significance χ^2, CFI > 0.90, TLI > 0.90, and RMSEA < 0.08 are regarded as satisfactory. For model estimation, changes in CFI (ΔCFI) ≤ 0.01, changes in TLI (ΔTLI) ≤ 0.01, and a smaller BIC are considered acceptable invariance.[29] Finally, manifest means that differences in RSES scores across the sexes were examined.

Third, to examine the internal reliability of the RSES, the Cronbach's α and the McDonald's ω coefficients were both computed.

Fourth, criterion validity was confirmed by comparing group differences in anxiety and depression (high vs low self-esteem) and validating the vulnerability model of depression. Participants with the highest 25% of the total RSES scores were defined as the high self-esteem group, whereas those with the lowest 25% of the scores were defined as the low self-esteem group. The independent samples *t* test was then used to compare group differences in MASQ-SF scores. Cohen's *d* value was used to estimate the effect size.[30] As a specific instance of structural equation modeling, path analysis was employed to examine the mediating roles of self-esteem and depressive symptoms in the relationship between perceived stress and life satisfaction. We initially tested several candidate models: *perceived stress → life satisfaction; perceived stress → self-esteem / depression → life satisfaction; perceived stress → self-esteem → depression → life satisfaction*. Based on the model fit indices and the significance of individual paths, we refined our models to arrive at the most

parsimonious and well-fitting model. The model-fitting criteria for path analysis were χ^2/df < 5, CFI > 0.90, TLI > 0.90, RMSEA < 0.08, and SRMR < 0.08.

Results

Model Fitting of RSES

As seen in Table 1, the RSES-GSE Model showed a poor fit, with both CFI and TLI lower than 0.90 and RMSEA (= 0.164) much above the listed criterion. The fit of the RSES-PN Model improved, but its CFI, TLI, and RMSEA did not meet these standards. The fit indices became much more satisfactory after the method effects were considered, except for the CTCU-PN Model, which could not be identified using Mplus. All remaining models met the criteria of CFI and TLI (all > 0.90), while only the CTCM-PN and CTCU-P models met RMSEA norms (0.061 for Model 3, 0.082 for Model 7, both < 0.100). Overall, the CTCM-PN Model, with a significantly smaller RMSEA and χ^2/df and a larger df, was the optimum fitting model (χ^2/df = 6.237, CFI = 0.986, TLI = 0.974, RMSEA = 0.061). As the CTCM-PN Model was selected, the fit indices for each gender indicated that the fit in males (CFI = 0.988, TLI = 0.979, χ^2 = 74.420) was better than that in females (CFI = 0.979, TLI = 0.962, χ^2 = 128.938), as shown in Table 2. Additionally, the loadings on the GSE factor range from 0.323 to 0.890.

Measurement Invariance Across Gender of the CTCM-PN Model

As shown in Table 3, the results of the measurement and structural invariance analyses of the CTCM-PN Model across genders all met the criteria (all χ^2/df < 5, CFI > 0.90, TLI > 0.90, RMSEA < 0.08, and SRMR < 0.08). Moreover, weak invariance, strong invariance, strict invariance, factor variance-covariance, and latent mean invariance of the CTCM-PN Model between males and females were established, with ΔCFI and ΔTLI < 0.01. The difference in manifest self-esteem levels between males and females was not statistically significant (P = 0.34, t = 0.95).

Table 1 Confirmatory Factors Analysis for Alternative Models

Models	AIC	BIC	χ^2 (*df*)	χ^2/*df*	CFI	TLI	SRMR	RMSEA (90% CI)
RSES-GSE	22784.491	22,942.138	1370.857*** (35)	39.167	0.852	0.810	0.068	0.164 (0.157 0.172)
RSES-PN	21927.121	22,090.023	511.487*** (34)	15.044	0.947	0.930	0.042	0.100 (0.092 0.107)
CTCM-PN	21589.230	21,799.426	155.926*** (25)	6.237	0.986	0.974	0.023	0.061 (0.052 0.070)
CTCM-P	21879.517	22,068.693	453.882 *** (29)	15.651	0.953	0.927	0.038	0.102 (0.094 0.110)
CTCM-N	21823.105	22,001.771	401.470*** (31)	12.951	0.959	0.941	0.031	0.092 (0.084 0.100)
CTCU-PN	CTCU-PN Model was not identified.							
CTCU-P	21655.218	21,891.687	211.583*** (20)	10.579	0.979	0.952	0.034	0.082 (0.072 0.092)
CTCU-N	21795.822	21,984.998	370.187 *** (29)	12.765	0.962	0.942	0.030	0.091 (0.083 0.100)

Note: ***P < 0.001.

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; *df*, degrees of freedom; CFI, Comparative Fit Index; TLI, Tucker–Lewis Index; SRMR, Standardized Root Mean Squared Residual; RMSEA, Root Mean Square Error of Approximation; RSES-GSE, unidimensional model with a global self-esteem (GSE) factor; RSES-PN, two separate factors representing positively and negatively worded items; CTCM-PN, CTCM bi-factor model containing a GSE factor and two factors describing positive and negative method effects; CTCM-P, CTCM bi-factor models with a GSE factor and only a positive factor of method effect; CTCM-N, CTCM bi-factor models with a GSE factor and only a negative factor of method effect; CTCU-PN, CTCU bi-factor model including one trait factor and correlated uniqueness among POS and NEG subscales; CTCU-P, CTCU model with one trait factor and correlated uniqueness among positively worded item; CTCU-N, CTCU model with one trait factor and correlated uniqueness among negatively worded item.

Table 2 Fit Indices of the CTCM-PN Model in Each Gender

	AIC	BIC	χ^2 (*df*)	χ^2/*df*	CFI	TLI	SRMR	RMSEA (90% CI)
Male	11104.940	11,286.754	74.420*** (25)	2.977	0.988	0.979	0.020	0.053 (0.040 0.067)
Female	10427.543	10,610.657	128.938*** (25)	5.158	0.979	0.962	0.031	0.076 (0.063 0.089)

Note: ***P < 0.001.

Abbreviations: AIC, Akaike Information Criterion; BIC, Bayesian Information Criterion; *df*, degrees of freedom; CFI, Comparative Fit Index; TLI, Tucker–Lewis Index; SRMR, Standardized Root Mean Squared Residual; RMSEA, Root Mean Square Error of Approximation; CI, Confidence Interval.

Table 3 Fit Indices for Measurement Invariance Tests of CTCM-PN Model of RSES Across Gender

Model	χ^2 (*df*)	χ^2/*df*	TLI	CFI	RMSEA (90% CI)		ΔCFI	ΔTLI	BIC
Model A	155.247 (51)***	3.044	0.973	0.985	0.054 (0.044 0.064)				21,946.795
Model B	172.006 (68)***	2.530	0.980	0.985	0.046 (0.038 0.055)	B vs A	0.000	0.007	21,844.171
Model C	188.608 (74)***	2.549	0.980	0.983	0.047 (0.039 0.055)	C vs B	0.002	0.000	21,817.557
Model D	236.385 (84)***	2.814	0.976	0.978	0.051 (0.043 0.058)	D vs C	0.005	0.004	21,822.368
Model E	245.802 (87)***	2.825	0.976	0.977	0.051 (0.043 0.058)	E vs C	0.006	0.004	21,818.415
Model F	244.024 (88)***	2.773	0.977	0.977	0.050 (0.043 0.058)	F vs E	0.000	0.001	21,804.214

Notes: Model A = Configural Invariance; Model B = Weak Invariance; Model C = Strong Invariance; Model D = Strict Invariance; Model E = Factor Variance-Covariance Invariance; Model F = Latent Means Invariance ***P < 0.001.
Abbreviations: *df*, degrees of freedom; TLI, Tucker–Lewis Index; CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation; CI, Confidence Interval; ΔCFI, Changes in CFI; ΔTLI, Changes in TLI; BIC, Bayesian Information Criterion.

Table 4 Differences in MASQ-SF's Scores Between the High-Level and the Low-Level Self-Esteem Groups

	High-Level Self-Esteem Group (n=333)	Low-Level Self-Esteem Group (n=338)	t	|Cohen's *d*|
MASQ-SF	107.61 ± 19.66	135.64 ± 29.75	14.41***	1.11
GD	17.31 ± 5.22	25.17 ± 9.40	13.42***	1.04
AD	51.80 ± 10.99	65.59 ± 12.23	15.36***	1.19
GA	16.34 ± 5.12	19.21 ± 6.51	6.36***	0.49
AA	22.16 ± 5.80	25.67 ± 9.41	5.81***	0.45

Note: ***P < 0.001.
Abbreviations: MASQ-SF, the total score of MASQ-SF; GD, General Distress; AD, Anhedonic Depression; GA, General Anxiety; AA, Anxious Arousal; |Cohen's *d*|, absolute values of Cohen's *d*.

Reliability of RSES

The Cronbach's α of the RSES was 0.92 for the whole sample, with 0.91 for males and 0.92 for females. The Cronbach's α of the positively worded items of the RSES was 0.91 for the whole sample, 0.91 for males, and 0.92 for females. Correspondingly, the Cronbach's α α of the negatively worded items of the RSES was 0.83 for the whole sample, 0.83 for males, and 0.83 for females. Meanwhile, the McDonald's Omega (*ω*) coefficients were from 0.84 to 0.92 for each subscale of the RSES.

Criterion Validity of RSES

According to the criteria mentioned above, 333 subjects were classified into the high-level self-esteem group (total scores of RSES in the top 25%), and 338 subjects were classified into the low-level self-esteem group (total scores of RSES in the last 25%). The results of the independent samples *t* test indicated significant group differences in MASQ-SF's scores, with the effect size ranging from 0.45 to 1.19 (see Table 4).

As illustrated in Figure 2(a) and Table 5, the model fit indicators of the original model all met the fit criteria, while the path from self-esteem to quality of life was not significant. Therefore, that path was removed, and the model fit was

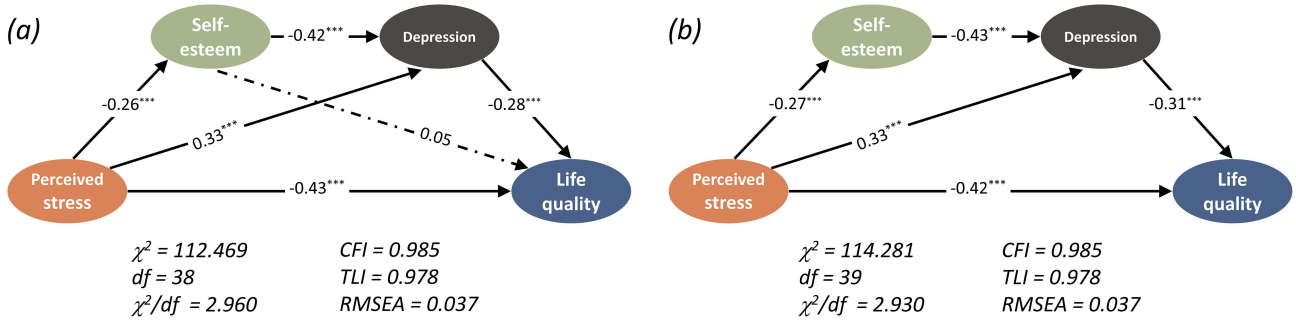

Figure 2 Results of Path Analyses. (**a**) Results of the original model; (**b**) Results of the modified model.
Notes: ***P < 0.001

Table 5 Standardized Indirect Effects and 95% Confidence Intervals of Structural Equation Model

Path	Estimated	95% CI	Effect Size
Perceived Stress→Depression→Life Quality	−0.093	−0.125, −0.062	74.40%
Perceived Stress→Self-esteem→Depression→Life Quality	−0.032	−0.048, −0.016	25.60%
Total indirect effects	−0.125	−0.161, −0.090	100.00%

Notes: Estimated was the standardized indirect effect of each mediate pathway; effect size was calculated by using the ratio of the total effect of the mediating effect. Empirical 95% CI did not overlap with zero; all pathways were significant.

reversed. The fit results of the modified model are presented in Figure 2(b) and Table 5, where all paths were significant and the model fit indicators fulfilled the fit criterion.

Discussion

In the current study, we comprehensively investigated the factorial and criterion validities of the Chinese version of the RSES in undergraduates using the following analyses. Initially, we compared the model fittings of eight competitive models using CFA with a sample of undergraduates, revealing that the CTCM-PN Model best fitted the data. We then confirmed that the gender invariance of the CTCM bi-factor model was established at configural, metric, scalar, and strict levels, and further confirmed latent means invariance. Next, we measured the Cronbach's alpha α of the RSES and confirmed that it had acceptable reliability among Chinese undergraduates. Subsequently, we verified the criterion validity of the RSES by comparing anxiety and depressive symptoms between individuals with high and low self-esteem. Finally, we provided further robustness to the criterion validity of the RSES by unraveling the essential contribution of self-esteem to alleviating depressive symptoms and ultimately enhancing life satisfaction from the perspective of vulnerable models.

Consistent with prior research, the current study supported the CTCM bi-factor model, including a GSE factor with two positive and negative factors (CTCM-PN) that fitted optimally.[16] In the CTCM-PN Model, all items had significant and considerable loadings on the GSE factor as they all reflected self-esteem levels. With this model, it is possible to distinguish the items affected by wording effects. In addition, it was observed that negatively worded items were dominant in CTCM models, whereas the opposite pattern existed in CTCU models in this study. However, in most previous studies, both CTCU and CTCM models showed a tendency for models with negatively worded items to have a better fitting.[5,10] The different dominance of items with different valences in this study could be attributed to the unsatisfactory fitting of CTCU models, which were less stable and unable to analyze method effects directly, whereas CTCM models were more suitable for reflecting convergent, discriminant, and criterion validity, where negatively worded items were more important.[9]

Because the CTCM-PN Model of RSES was the optimum model, gender invariance analyses were then performed using multigroup CFAs. Controlling for method effects introduced by negatively and positively wording, the CTCM-PN model of the RSES was configurationally, weakly, strongly, and strictly invariant across genders, implying that self-esteem measured by the Chinese version of the RSES shared identical potential theoretical constructs and psychological implications across genders. Measurement invariance of the RSES across genders has been confirmed in other countries,[15,16] and the present study further confirmed the factorial invariance by verifying the factor variance-covariance invariance, which not only confirms the robust measurement properties of the RSES in the Chinese population but also establishes the foundation for further applied research.

Differences in latent means and manifest scores were then examined upon the establishment of factorial invariance, which uniformly suggested that self-esteem levels of male and female undergraduates were comparable. This finding was consistent with the results of previous research showing latent mean invariance of RSES in undergraduates.[17] However, it contradicted with several studies comparing observed scores, which reported significantly higher self-esteem levels in females than in males.[14,31,32] This disparity may stem from the comparison of arms, which means that comparisons based on measurement invariance eliminate systematic errors caused by different units or reference points across groups, reflecting authentic group differences.[33] However, direct comparisons of manifest means used in prior research could not eliminate the systematic errors caused by method effects; therefore, a more careful interpretation of manifest mean differences is required.

The findings of the criterion validity analyses provided further evidence for the extended utility of the RSES within the realm of mental health assessment. Consistent with prior research,[16] the current study revealed that individuals with comparatively

lower levels of self-esteem would experience higher levels of anxiety and depression. Typically, individuals with lower self-esteem and negative self-perception are at higher risk of self-derogation, sociopathic behavior, anxiety, depression, and other mental problems.[16] To further investigate the implications of self-esteem on mental health and daily life of undergraduate students, a vulnerability model for depression was validated in the current research. Consistent with the results of previous studies,[20] our findings revealed a well-fitted vulnerability model that explained the relationship between self-esteem and depression. Undergraduates with high self-esteem are more likely to alleviate stress-induced depressive symptoms and ultimately enhance their life satisfaction. This finding provides ideas for subsequent efforts to improve the physical and mental health of undergraduates, especially in what is still the post-epidemic period with relatively high levels of mental stress.

Limitation

Despite the advantages of this study, it has several limitations must be mentioned. First, the participants included in this study were drawn from a demographically homogeneous sample of college students, suggesting that future research with more diverse and heterogeneous samples would be beneficial in ascertaining the broader applicability of the CTCM bi-factor model across various populations. Second, the equivalence analysis was limited to the gender dimension in this study, further research should extend this analysis to include other dimensions, such as temporal stability, age variations, and clinical populations. Third, this study did not specifically address gender differences in self-esteem, future research should endeavor to increase sample size to allow for a more detailed examination of gender disparities in self-esteem levels. This examination, grounded in positive psychology, could inform targeted interventions to enhance self-esteem and well-being across gender. Such inquiry is crucial not only for elucidating the complex interplay between gender and self-esteem but also for developing nuanced strategies that promote psychological resilience among diverse populations.

Conclusion

This study systematically confirmed the factorial and criterion validities of the RSES within a Chinese undergraduate sample by comparing the model fitting of eight models of the RSES, establishing the gender invariance and latent mean invariance of the CTCM bi-factor of the RSES, and verifying the criterion validity of the RSES. These findings pave the way for future research on self-esteem measurement within the Chinese cultural context, contributing to a deeper understanding of the nuances of self-esteem in diverse cultural settings.

Acknowledgments

We sincerely extend our heartfelt gratitude to all the participants involved in this research.

Funding

This study was supported by grants from the National Natural Science Foundation of China (Grant No. 82271573) and the Fundamental Research Funds for the Central Universities of Central South University (Grant No. 2021zzts0375).

Disclosure

The authors have no conflict of interest to declare.

References

1. Rosenberg M. Society and the adolescent self-image. *Princet Univ*. 1965.
2. Mayordomo T, Gutierrez M, Sales A. Adapting and validating the Rosenberg self-esteem scale for elderly Spanish population. *Int Psychogeriatr*. 2020;32(2):183–190. doi:10.1017/S1041610219001170
3. Reid-Russell A, Miller AB, Cvencek D, Meltzoff AN, McLaughlin KA. Lower implicit self-esteem as a pathway linking childhood abuse to depression and suicidal ideation. *Dev Psychopathol*. 2022;34(4):1272–1286. doi:10.1017/S0954579420002217
4. Syropoulou A, Vernadakis N, Papastergiou M, Kourtessis T. Psychometric evaluation of the Rosenberg self-esteem scale in primary school students with mild intellectual disability: first evidence. *Res Dev Disabil*. 2021;114:103964. doi:10.1016/j.ridd.2021.103964
5. Alessandri G, Vecchione M, Eisenberg N, Łaguna M. On the factor structure of the Rosenberg (1965) general self-esteem scale. *Psychol Assess*. 2015;27(2):621. doi:10.1037/pas0000073
6. Xu ML, Leung SO. Effects of varying numbers of Likert scale points on factor structure of the Rosenberg self-esteem scale. *Asian J Soc Psychol*. 2018;21(3):119–128. doi:10.1111/ajsp.12214

7. Marsh HW, Scalas LF, Nagengast B. Longitudinal tests of competing factor structures for the Rosenberg self-esteem scale: traits, ephemeral artifacts, and stable response styles. *Psychol Assess*. 2010;22(2):366. doi:10.1037/a0019225
8. Urbán R, Szigeti R, Kökönyei G, Demetrovics Z. Global self-esteem and method effects: competing factor structures, longitudinal invariance, and response styles in adolescents. *Behav Res Methods*. 2014;46(2):488–498. doi:10.3758/s13428-013-0391-5
9. Wang M-C, Deng Q, Bi X. Latent variable modeling using Bayesian methods. *Adv Psychol Sci*. 2017;25(10):1682. doi:10.3724/SP.J.1042.2017.01682
10. Mullen SP, Gothe NP, McAuley E. Evaluation of the factor structure of the Rosenberg self-esteem scale in older adults. *Personal Individ Differ*. 2013;54(2):153–157. doi:10.1016/j.paid.2012.08.009
11. Gómez-Lugo M, Espada JP, Morales A, et al. Adaptation, validation, reliability and factorial equivalence of the Rosenberg self-esteem scale in Colombian and Spanish population. *Span J Psychol*. 2016;19:E66. doi:10.1017/sjp.2016.67
12. Vasconcelos-Raposo J, Fernandes HM, Teixeira CM, Bertelli R. Factorial validity and invariance of the Rosenberg self-esteem scale among Portuguese youngsters. *Soc Indic Res*. 2012;105(3):483–498. doi:10.1007/s11205-011-9782-0
13. Boduszek D, Hyland P, Dhingra K, Mallett J. The factor structure and composite reliability of the Rosenberg self-esteem scale among ex-prisoners. *Personal Individ Differ*. 2013;55(8):877–881. doi:10.1016/j.paid.2013.07.014
14. Ye S, Yu L, Li -K-K. A cross-lagged model of self-esteem and life satisfaction: gender differences among Chinese university students. *Personal Individ Differ*. 2012;52(4):546–551. doi:10.1016/j.paid.2011.11.018
15. DiStefano C, Motl RW. Self-esteem and method effects associated with negatively worded items: investigating factorial invariance by sex. *Struct Equ Model Multidiscip J*. 2009;16(1):134–146. doi:10.1080/10705510802565403
16. Michaelides MP, Koutsogiorgi C, Panayiotou G. Method effects on an adaptation of the Rosenberg self-esteem scale in Greek and the role of personality traits. *J Pers Assess*. 2016;98(2):178–188. doi:10.1080/00223891.2015.1089248
17. Lima TJSD, Souza LECD. Rosenberg self-esteem scale: method effect and gender invariance. *Psico-USF*. 2019;24(3):517–528. doi:10.1590/1413-82712019240309
18. Ji Y, Rana C, Shi C, Zhong Y. Self-esteem mediates the relationships between social support, subjective well-being, and perceived discrimination in Chinese people with physical disability. *Front Psychol*. 2019;10:2230. doi:10.3389/fpsyg.2019.02230
19. Wu Y, Zuo B, Wen F, Yan L. Rosenberg self-esteem scale: method effects, factorial structure and scale invariance across migrant child and urban child populations in China. *J Pers Assess*. 2017;99(1):83–93. doi:10.1080/00223891.2016.1217420
20. Orth U, Robins RW, Meier LL. Disentangling the effects of low self-esteem and stressful events on depression: findings from three longitudinal studies. *J Pers Soc Psychol*. 2009;97(2):307. doi:10.1037/a0015645
21. Orth U, Robins RW. Understanding the link between low self-esteem and depression. *Curr Dir Psychol Sci*. 2013;22(6):455–460. doi:10.1177/0963721413492763
22. Arslan G. Mediating role of the self–esteem and resilience in the association between social exclusion and life satisfaction among adolescents. *Personal Individ Differ*. 2019;151:109514. doi:10.1016/j.paid.2019.109514
23. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41(4):1149–1160. doi:10.3758/BRM.41.4.1149
24. Shen Z-L, Cai T-S. Disposal to the 8th item of Rosenberg self-esteem scale (Chinese version). *Chin Ment Health J*. 2008.
25. Zhu X, Auerbach RP, Yao S, et al. Psychometric properties of the cognitive emotion regulation questionnaire: Chinese version. *Cogn Emot*. 2008;22(2):288–307. doi:10.1080/02699930701369035
26. Stevanovic D. Quality of life enjoyment and satisfaction questionnaire–short form for quality of life assessments in clinical practice: a psychometric study. *J Psychiatr Ment Health Nurs*. 2011;18(8):744–750. doi:10.1111/j.1365-2850.2011.01735.x
27. Lee Y-T, Liu S-I, Huang H-C, et al. Validity and reliability of the Chinese version of the short form of quality of life enjoyment and satisfaction questionnaire (q-LES-q-SF). *Qual Life Res*. 2014;23(3):907–916. doi:10.1007/s11136-013-0528-0
28. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1–55. doi:10.1080/10705519909540118
29. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Model*. 2002;9(2):233–255. doi:10.1207/S15328007SEM0902_5
30. Cohen J. A power primer. 2016.
31. Gao W, Luo Y, Cao X, Liu X. Gender differences in the relationship between self-esteem and depression among college students: a cross-lagged study from China. *J Res Personal*. 2022;97:104202. doi:10.1016/j.jrp.2022.104202
32. Juszkiewicz A, Lachowicz-Tabaczek K, Wróbel M. Self-esteem, gender, and emotional contagion: what predicts people's proneness to "catch" the feelings of others? *Personal Individ Differ*. 2020;157:109803. doi:10.1016/j.paid.2019.109803
33. Brown TA. *Confirmatory Factor Analysis for Applied Research*. Guilford publications; 2015.

Psychology Research and Behavior Management Dovepress

Publish your work in this journal

Psychology Research and Behavior Management is an international, peer-reviewed, open access journal focusing on the science of psychology and its application in behavior management to develop improved outcomes in the clinical, educational, sports and business arenas. Specific topics covered in the journal include: Neuroscience, memory and decision making; Behavior modification and management; Clinical applications; Business and sports performance management; Social and developmental studies; Animal studies. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/psychology-research-and-behavior-management-journal