



# Psychometric Properties of the Chinese Version of the Neuroticism Subscale of the NEO-PI

Chang Xi<sup>1</sup>, Mingtian Zhong<sup>2</sup>, Xiaoxia Lei<sup>1</sup>, Ying Liu<sup>1</sup>, Yu Ling<sup>3</sup>, Xiongzhuo Zhu<sup>1,4</sup>, Shuqiao Yao<sup>1,4</sup> and Jinyao Yi<sup>1,4\*</sup>

<sup>1</sup> Medical Psychological Center, The Second Xiangya Hospital, Central South University, Changsha, China, <sup>2</sup> Center for Studies of Psychological Application, School of Psychology, South China Normal University, Guangzhou, China, <sup>3</sup> Education Institute, Hunan Agricultural University, Changsha, China, <sup>4</sup> Medical Psychological Institute of Central South University, Changsha, China

## OPEN ACCESS

### Edited by:

Laura Badenes-Ribera,  
Universitat de València, Spain

### Reviewed by:

Agata Debowska,  
University of Sheffield,  
United Kingdom  
Jocelyn Holden Bolin,  
Ball State University, United States

### \*Correspondence:

Jinyao Yi  
jinyaoyi@csu.edu.cn;  
jinyaoyi2001@163.com

### Specialty section:

This article was submitted to  
Quantitative Psychology and  
Measurement,  
a section of the journal  
Frontiers in Psychology

**Received:** 18 September 2017

**Accepted:** 24 July 2018

**Published:** 17 August 2018

### Citation:

Xi C, Zhong M, Lei X, Liu Y, Ling Y,  
Zhu X, Yao S and Yi J (2018)  
Psychometric Properties of the  
Chinese Version of the Neuroticism  
Subscale of the NEO-PI.  
Front. Psychol. 9:1454.  
doi: 10.3389/fpsyg.2018.01454

Neuroticism is an important concept in psychology, self-report measures of neuroticism are important for both research and clinical practice. The neuroticism subscale of the Neuroticism-Extraversion-Openness Personality Inventory (NEO-PI) is a brief measure of neuroticism, and it was widely used in the world. This study was aimed to examine the psychometric properties of the Chinese version of the neuroticism subscale of the NEO-PI. A total of 5,494 undergraduates from three universities and 551 clinical patients with mental disorders from a psychological clinic had completed the Chinese version of the neuroticism subscale of the NEO-PI. Confirmatory factor analysis was performed to examine how well the three hypothetical models fit the data and the measurement equivalence of neuroticism subscale across gender. The internal consistency and test-retest reliability were also evaluated. Both the six-facet model and the bi-factor model (six-facet model with one general factor) achieved satisfactory fit, while the six-facet model had best fit (Undergraduate sample: TLI = 0.919, CFI = 0.933, RMSEA = 0.044, SRMR = 0.033; Clinical sample: TLI = 0.921, CFI = 0.935, RMSEA = 0.047, SRMR = 0.041), and it had measurement equivalence across gender. The neuroticism subscale also showed acceptable internal consistency and good stability. Within the undergraduate sample, there were statistically significant gender differences in neuroticism total scores and scores of six facets, while there were no significant gender differences in the neuroticism scores in the clinical sample. Both in the undergraduate sample and the clinical sample, anxiety facet, depression facet and vulnerability facet of the neuroticism subscale significantly predicted the depression level, while anxiety facet, angry-hostility facet and vulnerability facet significantly predicted the anxiety level. In conclusion, the Chinese version of the neuroticism subscale is a reliable and valid measurement of neuroticism in both undergraduate and clinical population.

**Keywords:** neuroticism, big five, measurement equivalence, factor structure, reliability, validity

## INTRODUCTION

While the term neuroticism dates back to Freudian theory, the modern concept of neuroticism was introduced by Eysenck. Eysenck recognized neuroticism as a trait of emotionality, specifically the tendency to arouse quickly when stimulated and to inhibit emotions slowly (Eysenck and Michael, 1985). Neuroticism is operationally defined by the factors of items referring to irritability, anger, sadness, anxiety, worry, hostility, and self-consciousness (Costa and McCrae, 1992b).

Researches on neuroticism are important to public health due to the robust correlation between neuroticism and a wide variety of both mental and physical health problems (Malouff et al., 2005, 2006). Neurotic individuals have a limited tolerance for aversive stimuli and tend to experience negative emotion, such as anger, anxiety and depression (McCrae and Costa, 1997, 2008). Individuals with higher level of neuroticism not only tend to develop mood disorders, but also exhibit other disorders, such as substance use disorder and eating disorders (Widiger, 2001). Meanwhile, neurotic individuals are at higher risk of engaging in potentially destructive behaviors that may have a negative impact on health and lead to a decrease in life expectancy, such as smoking and excessive drinking (Daniel et al., 2009).

Self-reported questionnaires are the most commonly used tools for neuroticism assessment. The Eysenck Personality Questionnaire (Eysenck et al., 1985), the Neuroticism-Extraversion-Openness Personality Inventory (NEO-PI; John and Kentle, 1991), and the Dutch Personality Questionnaire (Kerkhof, 2003) have proven adequate psychometric properties in assessing neuroticism. The most widely used instrument to date is the 48-item neuroticism subscale of the NEO-PI, which assesses individual differences in a predisposition to experience negative emotional states associated with symptoms of depression, anxiety and high arousal (Pervin, 1999; Matthews, 2000). The NEO-PI is a 240-item self-reported questionnaire based on the five-factor model (the Big Five) of personality: Extraversion, Neuroticism, Conscientiousness, Agreeableness, and Openness/ Intellect. Each factor of NEO-PI includes six facets. Over the past two decades, the Big Five factors have become the most prominent model for describing the structure of personality traits (Egger et al., 2003; Malouff et al., 2006; Ortet et al., 2012). Among the Big Five factors, neuroticism has been the most studied factor by researchers around the world and the neuroticism subscale of the NEO-PI has been the most widely used instrument for the measurement of neuroticism. The neuroticism subscale of the NEO-PI measures six facets of neuroticism: anxiety, angry-hostility, depression, self-consciousness, impulsiveness and vulnerability (Costa and McCrae, 1992b). The psychometric properties of the NEO-PI scale have been proven validity across cultures and ages, each facet scale includes eight items to ensure that the scales are comparable in many respects (Terracciano, 2011). Considering the wide use of the neuroticism subscale of the NEO-PI, studies focusing on the psychometric properties of the neuroticism subscale are warranted.

Prior researches have found that women scored higher than men on almost all facets of neuroticism and on total level of

neuroticism (Grossman and Wood, 1993; Lynn and Martin, 1997; Costa et al., 2001; Terracciano, 2003). One study including data from 26 countries concluded that women scored higher than men on neuroticism as measured by the neuroticism subscale of the NEO-PI, as well as on most facets of neuroticism (Costa et al., 2001). A study in Italy also found that women scored significantly higher than men on total neuroticism and on each of its facets (Terracciano, 2003). Since neuroticism increases the risk of a wide range of mental disorders, gender difference in neuroticism may lead to the gender difference seen in major psychopathologies. For example, women diagnosed with many mental disorders (e.g., phobias, major depression, dysthymic disorder, generalized anxiety disorder, borderline personality disorder) were associated with neuroticism at a higher rate than men (Grossman and Wood, 1993). Although there were many previous researches supporting the difference in neuroticism between men and women (Foot and Koszycki, 2010; Banzhaf et al., 2012; Konishi et al., 2014), only when the latent structure are equivalent across gender, the observed gender difference in neuroticism can truly predict the true difference between males and females.

Measurement invariance is “the mathematical equality of corresponding measurement parameters for a given factorially defined construct across groups” (Little, 1997, p. 55). Generally, measurement invariance is used to estimate parameters that reflect difference in the latent construct, and examine whether an instrument’s items operate in the same way in different groups (Meredith, 1993). Simply put, it is used to examine whether an instrument’s items measure the same parameters in different groups. If there is poor measurement invariance in the neuroticism scale across gender, the observed difference between males and females may be confounded by the fact that different constructs are being measured.

Therefore, the purpose of this study was to examine the psychometric properties of the 48-item neuroticism subscale of the NEO-PI in Chinese population. One undergraduate sample and one clinical sample were used in this study. Specifically, the current study would examine the measurement invariance across gender in the undergraduate sample.

## METHODS

### Participants

From March 2015 to October 2015, 5,612 undergraduates from three universities in Changsha were invited to participate in the study. A total of 5,494 (99.7%) students [2,878 (52.4%) men, 2,616 (47.6%) women; aged 19–30 (Mean = 25.0; standard deviation (SD) = 1.02)] provided complete data anonymously. To estimate test–retest reliability, a subgroup of 865 students who had been participating in a psychology course completed the neuroticism subscale of the NEO-PI twice with a 2-month interval.

In addition to the subjects described above, from March 2015 to March 2017, 568 outpatients, who had been referred for assessment and treatment in the psychological clinic of Second Xiangya Hospital, were also asked to participate in the study. Those patients who could not understand the question well were excluded, such as patients with intellectual disability. A total

of 551 (97%) outpatients provided complete data, consisting of 281 men (51%) and 270 women (49%), ranging in age from 16 to 77 (Mean = 31.41;  $SD = 18.52$ ). The diagnoses of the clinical sample were schizophrenia, depressive disorder, anxiety disorder, obsessive-compulsive disorder, trauma and stress-related disorder, somatic symptom and related disorder, personality disorder and other mental disorders. There was a significant age difference between the undergraduate sample and the clinical sample, the clinical sample was significantly older than the undergraduate sample ( $t = 5.772, p < 0.001$ ), but no significant gender difference was found between the two samples ( $\chi^2 = 0.386, df = 1, p = 0.535$ ).

The study was approved by the Ethics committee of Second Xiangya Hospital, Central South University. All participants provided written informed consent at the time of enrollment.

## Instruments

### The Neuroticism Subscale of the NEO-PI

The NEO-PI is the most comprehensive self-report questionnaire that measures the five-factor model of personality, in which neuroticism is included (Costa and McCrae, 1997; Xu and Potenza, 2012). The NEO-PI consists of 240 items and has been extensively validated (John and Kentle, 1991; Luchetti et al., 2018). It measures five major factors and 30 facets. The internal consistency of the NEO-PI was high (0.86 to 0.92), and the internal consistency of its facets ranged from 0.58 to 0.81 (Costa and McCrae, 1992a). The neuroticism subscale of the NEO-PI includes 48 items, each is rated on a 5-point Likert scale (0–4), with total score ranging from 0 to 192. Higher scores are indicative of higher level of neuroticism. The neuroticism subscale includes six facets: anxiety, angry-hostility, depression, self-consciousness, impulsiveness, and vulnerability. Each facet is measured by eight items, and the score of each facet ranges from 0 to 32 (Dai and Yao, 2004). The neuroticism subscale of the NEO-PI has previously been proven to be valid in the assessment of neuroticism among Chinese people (Dai et al., 1999).

### The Center for Epidemiological Studies Depression Scale

The 20-item Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) assesses various current depressive symptoms in general population. Each item is rated on a 4-point scale. Total score ranges from 0 to 80. Higher scores are indicative of a greater presence of depressive symptoms. The Chinese version of the CES-D has acceptable reliability (Cronbach's  $\alpha = 0.88$ ) and validity in Chinese population (Wang et al., 2013).

### The Self-Rating Anxiety Scale

The Zung Self-Rating Anxiety Scale (SAS) was designed to quantify an individual's level of anxiety (Zung, 1965). Each question on the SAS is scored on a Likert-type scale of 1–4 (1 = "rarely," 2 = "some of the time," 3 = "often," and 4 = "most of the time"). The total score ranges from 20 to 80. Higher scores are indicative of more symptoms of anxiety. The SAS has demonstrated acceptable reliability (Cronbach's  $\alpha = 0.78$ ) and

validity in the assessment of anxiety in Chinese-speaking samples (Luo et al., 2006; Wang and Tang, 2011).

## STATISTICAL ANALYSES

To evaluate the construct validity of the neuroticism subscale, we performed robust maximum-likelihood (MLR) confirmatory factor analysis (CFA) using M-plus 7.11 software. The CFA was used to determine how well the factor models fit the data, and to examine the measurement equivalence of neuroticism subscale across gender. We examined three hypothetical models: the original single-factor model (one general factor), the six-facet model (six facets: anxiety, angry-hostility, depression, self-consciousness, impulsiveness, and vulnerability), and the bi-factor model (six-facet model with one general factor, each item loading on both one facet and the general factor) (Reise et al., 2010). Each model included 24 parcels, which were created by randomly selecting two items within a facet for each parcel (Kishton and Widaman, 1994; Floyd and Widaman, 1995), so each facet included four parcels, see **Supplementary Material**. Several model fit indices were used to evaluate the model fit: Tucker-Lewis index (TLI), comparative fit index (CFI), and root-mean-square error of approximation (RMSEA) (Chou et al., 1991; Hu and Bentler, 1998). The criteria used to evaluate model fit were:  $TLI \geq 0.90$ ,  $CFI \geq 0.90$ , and  $RMSEA \leq 0.08$  (Browne and Cudeck, 1993; Hu and Bentler, 1999). To further compare the fit of the competing models, the  $\chi^2$  test is also reported.

To assess the measurement invariance of neuroticism subscale across gender in the undergraduate sample, the multi-group confirmatory factor analysis (MG-CFA) was used. The test of measurement invariance was divided into four levels from low to high: configural invariance, weak factorial invariance, strong factorial invariance, and strict invariance (Joreskog, 1971). The four levels have a hierarchical relationship so that data analyses were carried out step by step: (1) configural invariance (model 1), a baseline model for each group was construed such that for both men and women the following criteria were met:  $TLI \geq 0.90$ ,  $CFI \geq 0.90$ , and  $RMSEA \leq 0.08$ , (2) weak factorial invariance (model 2), factor loadings were equal across groups, (3) strong factorial invariance (model 3), the factor loadings and intercepts of variables were equal across groups and (4) strict invariance (model 4), the factor loadings, intercepts of variables and error variances were equal across groups, while strict invariance is not necessary for most researches (Widaman and Reise, 1997). Measurement invariance is considered established when two of following are satisfied: the  $\chi^2$  difference test resulted in a  $p$ -value  $> 0.05$ , the change of  $CFI < 0.01$ , the change of  $TLI < 0.01$  and the change of  $RMSEA < 0.015$  (Cheung and Rensvold, 2002; Chen, 2007; Ferro and Boyle, 2013).

Cronbach's alphas ( $\alpha$ ) and mean inter-item correlations ( $M_{IC}$ ) for the full neuroticism scale and for each of the facets were calculated to evaluate internal reliability. Generally speaking, a minimum standard of 0.70 is set for Cronbach's  $\alpha$  coefficients, but an  $\alpha$  of 0.60 is also considered acceptable (DeVellis, 1991). An optimal range of 0.10–0.40 was set for the  $M_{IC}$  (Briggs and Venue, 1986; Nunnally, 1994). The Pearson correlation

**TABLE 1** | The model fit indices of the six-facet model, the single-factor model and the six-facet bi-factor model in the two groups.

Models	Undergraduate sample (N = 5,494)			Clinical sample (N = 551)		
	Six-facet	Single-factor	Bi-factor	Six-facet	Single-factor	Bi-factor
S-B $\chi^2$	3080.5594	4055.9917	3077.7788	603.0128	804.0263	627.8021
Df	237	252	228	237	252	228
TLI	0.919	0.899	0.915	0.921	0.887	0.911
CFI	0.933	0.910	0.932	0.935	0.900	0.929
RMSEA	0.044	0.049	0.045	0.047	0.056	0.050
SRMR	0.033	0.038	0.034	0.041	0.046	0.041
BIC	265683.798	266528.180	265756.601	31802.734	31909.091	31884.332

S-B $\chi^2$ , Chi-Square Test of Model Fit; Df, degrees of freedom; GFI, The Goodness-of-fit Index; TLI, Tucker-Lewis Index; CFI, Comparative Fit Index; RMSEA, Root-Mean-Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; BIC, Bayesian information criterion; Six-facet, The six-facet model; Single-factor, The single-factor model; Bi-factor, The bi-factor model (six-facet model with one general factor, each item loading on both one facet and the general factor).

**TABLE 2** | The factor loading of each parcel in the six-facet model.

	Undergraduate sample	Clinical sample		Undergraduate sample	Clinical sample
P1	0.522	0.596	P13	0.476	0.447
P2	0.690	0.733	P14	0.593	0.478
P3	0.533	0.650	P15	0.526	0.579
P4	0.588	0.670	P16	0.413	0.464
P5	0.560	0.576	P17	0.448	0.507
P6	0.444	0.407	P18	0.426	0.410
P7	0.446	0.536	P19	0.563	0.512
P8	0.707	0.750	P20	0.679	0.777
P9	0.536	0.756	P21	0.659	0.705
P10	0.665	0.761	P22	0.703	0.765
P11	0.575	0.682	P23	0.692	0.706
P12	0.700	0.778	P24	0.663	0.695

P1~P4 belong to anxiety facet; P5~P8 belong to angry-hostility facet; P9~P12 belong to depression facet; P13~P16 self-consciousness facet; P17~P20 belong to impulsiveness facet; P21~P24 belong to vulnerability facet.

coefficient ( $r$ ) was used to evaluate test-retest reliability and was set with a minimum standard of 0.70 (Anastasi and Urbina, 1997). The relationships between the total scale and the six facets were also examined by Pearson's  $r$ . Independent-samples  $t$ -test was used to compare the differences in scores of neuroticism subscale between the undergraduate sample and the clinical sample, and the gender differences in scores of neuroticism subscale.

To examine whether depression and anxiety were predicted by demographic variables and neuroticism, we performed multiple linear regression both in the undergraduate sample and the clinical sample, with the CES-D total score and SAS total score as dependent variable, respectively.

## RESULTS

### Descriptive Statistics

The total scores of neuroticism subscale of the NEO-PI ranged from 11 to 155 (Mean = 72.12;  $SD$  = 20.99) in the undergraduate sample. While in the clinical sample, the total scores of neuroticism subscale ranged from 19 to 175 (Mean = 106.84;  $SD$  = 28.36).

### The Goodness of Fit Indices for Neuroticism Subscale

Both in the undergraduate sample and the clinical sample, the fit indices of the six-facet model and the bi-factor model (six-facet model with one general factor) all reached acceptable standards, but the single-factor model didn't fit quite well (Table 1). However, the six-facet model had significant improvement over the bi-factor model in the clinical sample,  $\chi^2$  (diff) = 24.79,  $p < 0.001$ , but not in the undergraduate sample  $\chi^2$  (diff) = 2.78,  $p > 0.05$ . The standardized factor loading of the six-facet model were presented in Table 2.

### Measurement Invariance Across Gender for the Undergraduate Sample

As the six-facet model fitted the data best, we choose the six-facet model to estimate the measurement invariance across gender. As shown in Table 3, the baseline models were considered optimal in representing the data both for the male (TLI = 0.917; CFI = 0.931; RMSEA = 0.041) and the female undergraduates (TLI = 0.915; CFI = 0.930; RMSEA = 0.047), providing evidence for configural invariance. Furthermore, the changes in TLI and RMSEA ( $\Delta$ TLI < 0.010,  $\Delta$ RMSEA < 0.015) supported weak

**TABLE 3** | Measurement invariance of the neuroticism subscale across gender in the undergraduate sample.

Model	S-B $\chi^2$	Df	TLI	CFI	RMSEA	SRMR	$\Delta$ TLI	$\Delta$ CFI	$\Delta$ RMSEA
Model <sub>male</sub>	1548.9178	95	0.917	0.931	0.041	0.033			
Model <sub>female</sub>	1781.5871	95	0.915	0.930	0.047	0.037			
Model1	3426.8285	184	0.914	0.928	0.044	0.036			
Model2	3429.6562	169	0.917	0.928	0.043	0.036	0.003	0.000	-0.001
Model3	3818.7461	154	0.909	0.918	0.045	0.040	-0.008	-0.010	0.002
Model4	4353.0741	130	0.899	0.905	0.048	0.043	-0.010	-0.013	0.003

Model 1, configural equivalence; Model 2, metric equivalence; Model 3, strong equivalence; Model 4, strict equivalence.

**TABLE 4** | The Cronbach's  $\alpha$ , mean inter-item correlation and test-retest reliability for the neuroticism subscale and its six facets.

	Undergraduate sample			Clinical sample	
	Cronbach's $\alpha$ ( <i>N</i> = 5,494)	$M_{IC}$ ( <i>N</i> = 5,494)	Test-retest coefficient ( <i>N</i> = 865)	Cronbach's $\alpha$ ( <i>N</i> = 551)	$M_{IC}$ ( <i>N</i> = 551)
Neuroticism	0.91	0.18	0.71	0.93	0.21
Anxiety	0.67	0.21	0.60	0.74	0.26
Angry-hostility	0.71	0.24	0.62	0.69	0.21
Depression	0.72	0.25	0.64	0.83	0.38
Self-consciousness	0.58	0.15	0.55	0.54	0.13
Impulsiveness	0.64	0.19	0.52	0.61	0.17
Vulnerability	0.77	0.30	0.62	0.79	0.32

$M_{IC}$ , mean inter-item correlation.

equivalence and strong equivalence of the neuroticism subscale. In other words, the neuroticism subscale demonstrated good measurement equivalence across gender.

## Reliability

In the undergraduate sample, the Cronbach's  $\alpha$  coefficient was 0.91 for the neuroticism subscale, and ranged from 0.58 (self-consciousness) to 0.77 (vulnerability) for the six facets. In the clinical sample, the Cronbach's  $\alpha$  coefficient was 0.93 for the neuroticism subscale, and ranged from 0.54 (self-consciousness) to 0.83 (depression) for its six facets (Table 4).

The  $M_{IC}$  of neuroticism subscale was 0.18 in the undergraduate sample and 0.21 in the clinical sample. In the undergraduate sample, the  $M_{IC}$  of the six facets ranged from 0.15 (self-consciousness) to 0.30 (vulnerability), while the clinical sample, the  $M_{IC}$  of six facets ranged from 0.13 (self-consciousness) to 0.38 (depression), Table 4.

In the undergraduate sample, the test-retest coefficient for the neuroticism subscale was 0.71, and the test-retest coefficients for the six facets ranged from 0.52 (impulsiveness) to 0.64 (depression), see Table 4. Test-retest reliability was not evaluated in the clinical sample.

## Intercorrelation Among the Total Neuroticism Subscale and Its Six Facets

The correlation coefficients among the total score of neuroticism subscale and its six facets ranged from 0.44 to 0.85 in the undergraduate sample, and from 0.47 to 0.87 in the clinical

sample (Table 5). All correlation coefficients were positive and statistically significant ( $p < 0.001$ ).

## Gender Differences and Group Differences in Neuroticism

In the student subjects, females scored significantly higher than males on the total score of neuroticism subscale and scores of six facets (Total score:  $t = 17.031$ ,  $p < 0.001$ , Cohen's  $d = 0.46$ ; Angry-Hostility score:  $t = 11.775$ ,  $p < 0.001$ , Cohen's  $d = 0.32$ ; Anxiety score:  $t = 13.868$ ,  $p < 0.001$ , Cohen's  $d = 0.37$ ; Depression score:  $t = 11.197$ ,  $p < 0.001$ , Cohen's  $d = 0.30$ ; Self-Consciousness score:  $t = 9.688$ ,  $p < 0.001$ , Cohen's  $d = 0.26$ ; Impulsiveness score:  $t = 12.547$ ,  $p < 0.001$ , Cohen's  $d = 0.34$ ; and Vulnerability score:  $t = 21.358$ ,  $p < 0.001$ , Cohen's  $d = 0.58$ ). There were no significant differences in scores of neuroticism subscale and its six facets between males and females in the clinical sample (Table 6). Compared with the undergraduate group, the clinical group got significantly higher scores on total neuroticism ( $t = 35.70$ ,  $p < 0.001$ , Cohen's  $d = 1.39$ ) and all six facets ( $t: 19.06 \sim 33.96$ , all  $p < 0.001$ , Cohen's  $d: 0.80 \sim 1.31$ ).

## Predictive Validity of Demographic Variables and the Six Facets of Neuroticism Subscale for Depression and Anxiety

Results of multiple regression analyses with the CES-D total score as dependent variable were presented in Table 7. For

**TABLE 5** | Intercorrelations among total neuroticism subscale and its six facets.

	Undergraduate sample						Clinical sample					
	V	I	S	D	A <sup>1</sup>	A <sup>2</sup>	V	I	S	D	A <sup>1</sup>	A <sup>2</sup>
I	0.56**						0.59**					
S	0.56**	0.37**					0.58**	0.47**				
D	0.67**	0.49**	0.60**				0.68**	0.54**	0.60**			
A <sup>1</sup>	0.55**	0.52**	0.44**	0.54**			0.59**	0.54**	0.54**	0.64**		
A <sup>2</sup>	0.68**	0.48**	0.56**	0.68**	0.56**		0.70**	0.54**	0.61**	0.77**	0.62**	
N	0.85**	0.72**	0.74**	0.84**	0.76**	0.83**	0.85**	0.75**	0.77**	0.87**	0.81**	0.87**

V, Vulnerability; I, Impulsiveness; S, Self-consciousness; D, Depression; A<sup>1</sup>, Angry-Hostility; A<sup>2</sup>, Anxiety; N, Neuroticism. \*\*Correlation is significant at the 0.01 level (2-tailed).

**TABLE 6** | Scores on total neuroticism subscale and each of the six facets separated by gender (Means ± SD, ranges).

Scale	Undergraduate sample				Clinical sample			
	Total (N = 5,494)	Female (N = 2,616)	Male (N = 2,878)	Cohen's d	Total (N = 551)	Female (N = 270)	Male (N = 281)	Cohen's d
N	72.12 ± 20.99 <sup>a</sup> (11~155)	77.05 ± 21.68 <sup>b</sup> (11~155)	67.64 ± 19.28 (12~146)	0.46	106.84 ± 28.36 (19~175)	106.02 ± 27.34 (19~175)	107.70 ± 29.40 (30~174)	–
A <sup>1</sup>	10.49 ± 4.60 <sup>a</sup> (0~28)	11.25 ± 4.88 <sup>b</sup> (0~28)	9.80 ± 4.22 (0~26)	0.32	16.66 ± 5.57 (0~32)	17.05 ± 5.71 (3~32)	16.29 ± 5.42 (2~30)	–
A <sup>2</sup>	12.57 ± 4.41 <sup>a</sup> (0~30)	13.42 ± 4.56 <sup>b</sup> (0~30)	11.80 ± 4.41 (0~26)	0.37	19.13 ± 5.85 (1~32)	19.42 ± 5.93 (1~32)	18.85 ± 5.77 (1~32)	–
D	11.67 ± 4.71 <sup>a</sup> (0~32)	12.41 ± 4.84 <sup>b</sup> (0~32)	11.00 ± 4.50 (0~27)	0.30	18.24 ± 6.68 (0~32)	18.08 ± 7.16 (0~32)	18.39 ± 6.18 (0~32)	–
S	13.14 ± 4.12 <sup>a</sup> (0~29)	13.70 ± 4.28 <sup>b</sup> (0~29)	12.63 ± 3.91 (0~28)	0.26	17.92 ± 5.27 (3~47)	18.07 ± 5.30 (5~47)	17.78 ± 5.24 (7~32)	–
I	13.20 ± 4.33 <sup>a</sup> (0~28)	13.96 ± 4.46 <sup>b</sup> (1~28)	12.51 ± 4.09 (0~27)	0.34	16.95 ± 5.06 (0~29)	16.75 ± 5.35 (2~29)	17.13 ± 4.77 (8~29)	–
V	11.05 ± 4.36 <sup>a</sup> (0~31)	12.32 ± 4.39 <sup>b</sup> (0~31)	9.90 ± 3.99 (0~23)	0.58	17.94 ± 6.06 (0~32)	18.32 ± 6.04 (0~32)	17.57 ± 6.07 (0~32)	–

N, neuroticism; A<sup>1</sup>, Angry–Hostility; A<sup>2</sup>, Anxiety; D, Depression; S, Self-Consciousness; I, Impulsiveness; V, Vulnerability.

<sup>a</sup>Clinical sample scored significantly higher than undergraduate sample,  $p < 0.05$ ; <sup>b</sup>Women scored significantly higher than men,  $p < 0.05$ .

the undergraduate sample, demographic variables (gender and age), anxiety facet, angry-hostility facet, depression facet, self-consciousness facet and vulnerability facet significantly predicted the CES-D total score ( $p < 0.05$ ), except impulsiveness facet ( $p > 0.05$ ). For the clinical sample, anxiety facet, depression facet and vulnerability facet significantly predicted the CES-D total score ( $p < 0.05$ ), but none of other three facets of the neuroticism subscale significantly predicted the CES-D total score ( $p > 0.05$ ).

As **Table 8** showed, for the undergraduate sample, demographic variables (gender and age), anxiety facet, angry-hostility facet, depression facet, impulsiveness facet and vulnerability facet all significantly predicted the SAS total score ( $p < 0.05$ ), except self-consciousness facet ( $p > 0.05$ ). For the clinical sample, anxiety facet, angry-hostility facet and vulnerability facet significantly predicted the SAS total score ( $p < 0.05$ ), while none of other three facets of the neuroticism subscale significantly predicted the SAS total score ( $p > 0.05$ ).

## DISCUSSION

The neuroticism subscale of the NEO-PI is widely used around the world and is considered a valid and reliable measurement for the personality trait of neuroticism (Costa and McCrae, 1992a; Nunnally, 1994). The present study evaluated the reliability and validity of the neuroticism subscale in two Chinese samples (an undergraduate sample and a clinical sample), and examined its measurement invariance across gender in the undergraduate sample. Three hypothetical models were tested: single-factor model, six-factor model and bi-factor model (six-factor model with one general factor). The model indices of the six-factor model and the bi-factor model for both undergraduate sample and clinical sample all met the fit standards, supporting the neuroticism subscale's construct validity. Since the six-factor model fitted the data best, the six-factor model was chosen in the following measurement invariance analysis. In the undergraduate sample, the model indices for both the males and females indicated a good

**TABLE 7** | Multiple regression analysis with the CES-D total score as the dependent variable.

	Undergraduate sample (N = 5,494)				Clinical sample (N = 551)			
	$\beta$	95%CI	t	p	$\beta$	95%CI	t	p
Gender	-0.03	-0.68~-0.06	-2.30	0.02	0.02	-1.17~1.90	0.47	0.64
Age	-0.02	-0.32~-0.02	-2.25	0.02	-0.02	-0.05~0.03	-0.64	0.52
Anxiety	0.18	0.25~0.36	11.52	<0.01	0.23	0.19~0.65	3.62	<0.01
Angry-hostility	0.09	0.10~0.19	6.53	<0.01	0.08	-0.04~0.36	1.55	0.12
Depression	0.31	0.43~0.53	18.98	<0.01	0.20	0.11~0.51	3.10	<0.01
Self-consciousness	0.03	0.01~0.10	2.34	0.02	-0.03	-0.27~0.14	-0.65	0.52
Impulsiveness	0.02	-0.01~0.08	1.50	0.13	-0.04	-0.29~0.11	-0.87	0.39
Vulnerability	0.13	0.17~0.28	7.97	<0.01	0.20	0.16~0.55	3.52	<0.01
Constant		61.04~649.47	2.37	0.02		-25.12~129.86	1.32	0.19
		$F = 483.42; p < 0.01; R^2 = 0.41$				$F = 31.97; p < 0.01; R^2 = 0.33$		

$\beta$ , standardized  $\beta$ ; 95%CI, 95% confidence interval for  $\beta$ .

**TABLE 8** | Multiple regression analysis with the SAS total score as the dependent variable.

	Undergraduate sample (N = 5,494)				Clinical sample (N = 551)			
	$\beta$	95%CI	t	p	$\beta$	95%CI	t	p
Gender	-0.03	-0.70~-0.02	-2.10	0.04	-0.01	-1.75~1.41	-0.22	0.83
Age	-0.04	-0.43~-0.11	-3.25	<0.01	-0.07	-0.08~0.00	-1.86	0.06
Anxiety	0.22	0.30~0.42	12.48	<0.01	0.23	0.17~0.64	3.41	<0.01
Angry-hostility	0.16	0.20~0.29	10.43	<0.01	0.20	0.17~0.58	3.60	<0.01
Depression	0.15	0.17~0.28	8.21	<0.01	0.00	-0.20~0.20	-0.01	1.00
Self-consciousness	0.01	-0.04~0.07	0.54	0.59	0.01	-0.20~0.22	0.11	0.92
Impulsiveness	0.11	0.13~0.23	7.29	<0.01	0.06	-0.09~0.32	1.08	0.28
Vulnerability	0.05	0.03~0.15	2.93	<0.01	0.16	0.08~0.49	2.76	<0.01
Constant		235.77~878.04	3.40	<0.01		17.00~176.16	2.39	0.02
		$F = 300.84; p < 0.01; R^2 = 0.30$				$F = 26.78; p < 0.01; R^2 = 0.29$		

$\beta$ , standardized  $\beta$ ; 95%CI, 95% confidence interval for  $\beta$ .

fit of the neuroticism subscale's theoretical structure. In addition, the Chinese version of the neuroticism subscale also demonstrated acceptable reliability and factorial validity in this study.

The Cronbach's  $\alpha$  coefficients for the total neuroticism subscale, in both the undergraduate sample and the clinical sample, reached accepted standards ( $\alpha > 0.90$ ). Among the six facets, 5 of them were above 0.60 in the two samples, while the self-consciousness facet was with a value of  $< 0.60$ . These results were consistent with those of the original English version, which reported the internal reliability of all facets ranged from 0.58 to 0.81 (Costa and McCrae, 1992a). Hence, the  $\alpha$  values found in this study were deemed acceptable. Although the  $\alpha$  of self-consciousness facets in this study were similar to those of the original version (Costa and McCrae, 1992a), the internal consistency of self-consciousness facet should be reexamined in the future, and researchers should be cautious in interpreting this facet. In both the undergraduate sample and the clinical sample, all the mean inter-item coefficients were above the lowest accepted level (Briggs and Venue, 1986), which also supports good

internal consistency of the neuroticism subscale of the NEO-PI.

The test-retest reliability coefficient for the total neuroticism subscale was high and supports the notion of neuroticism as a stable personality trait. These findings were in agreement with the original version of NEO-PI (Anastasi and Urbina, 1997). The 2-month stability coefficients of the six facets ranged from 0.52 to 0.64 in this study, while a previous study found that the 6-month test-retest reliability of the neuroticism subscale of NEO-PI ranged from 0.66 to 0.92 (Costa and McCrae, 1988). Our results were a little lower than those of the previous study, which might be due to natural fluctuations in the state of the six facets of neuroticism. Further studies are needed to examine the stability of neuroticism subscale in both general population and clinical population, with larger samples and longer interval.

The correlations among the six facets were all significantly positive in both two groups (all  $p < 0.01$ ), which indicated that the six facets not only possessed relative independence, but also related to each other. The six facets of the neuroticism subscale measure different aspects of neuroticism, while all reflect the

neuroticism trait. Anyway, these results provided further support for the validity of the neuroticism subscale.

The scores on total neuroticism subscale and all six facets were significantly higher in the clinical sample than the undergraduate sample. These results were consistent with previous findings (Kotov et al., 2010; Ormel et al., 2013), for example, Kotov et al. found that individuals diagnosed with mental disorders also scored high in neuroticism (Kotov et al., 2010). Previous studies also supported that neuroticism is correlated with a wide variety of mental disorders, and individuals with high neuroticism tend to have more mental disorders (Kotov et al., 2010; Ormel et al., 2013). It has previously been stated that neuroticism is one of the most important risk factors in behavioral public health (Lahey, 2009), and that the economic costs of high neuroticism are estimated to exceed those of all common mental disorders combined (Cuijpers et al., 2010). Additionally, it has been shown that neuroticism could predict the onset of common mental disorders even after controlling for most psychiatric confounding variables (Ormel et al., 2013). All these findings suggest the importance of neuroticism in screening individuals with high-risk mental disorders, and in the implementation of early prevention programs to individuals with high-level neuroticism.

The current study found that most of facets could significantly predict depression and anxiety symptoms both in the undergraduate sample and clinical sample, especially the anxiety facet. Previous research has shown a strong link between personality and psychiatric illness, especially in the relationship between neuroticism and depressive symptoms (Luciano, 2015). Individual differences in personality traits, particularly in neuroticism, are known risk factors in the onset and development of depression (Saklofske and Janzen, 1995; Kendler et al., 2004). Research in England indicated that neuroticism predisposes individuals to depression via the relationship between ruminative thinking and low mood (Thorsten and Tobias, 2010). Additionally, certain personality traits may play an important role in how individuals respond to life-events. Individuals with high level of neuroticism are particularly vulnerable to the effect of life events on anxiety (Veen et al., 2016). A study, investigating the relationship between neuroticism and symptoms of anxiety and depression in three patient groups (generalized anxiety disorder; major depressive disorder; mixed anxiety-depressive disorder), revealed that neuroticism might increase the risk of anxious and depressive symptoms, as evidenced by increased worrying or brooding (Merino et al., 2016). While in this study, several facets (anxiety, depression and vulnerability) of the neuroticism subscale could significantly predict the depression level, and three facets (anxiety, angry-hostility and vulnerability) could significantly predict the anxiety level. Therefore, our results, in conjunction with the prior researches discussed above, supported that neuroticism has a great impact on symptoms of depression and anxiety both in general population and in those diagnosed with mental disorders.

It should be stressed that the validity of comparing groups is dependent on measurement invariance. The examination of measurement invariance is a necessary step prior to performing any comparisons across groups (Meredith, 1993; Byrne, 2008). The measurement invariance across gender in the current study supported weak equivalence (equal factor loadings) and strong equivalence (equal factor loadings /intercepts), which is sufficient for meaningful comparisons between groups. No strict measurement invariance was get in our results, while as Baumgartner suggested, strict measurement invariance is not required for substantive analyses (Baumgartner and Steenkamp, 1998). In general, the results of measurement invariance across gender in this study supported the validity of the significant gender differences in neuroticism scores, which was consistent with previous findings (Hirsh, 2011). Our results showed that women scored significantly higher than men on both total neuroticism and all six facets in the undergraduate sample, while in the clinical sample, there were no significant gender differences in total score of neuroticism or scores of its six facets. In the current study, the high level of neuroticism and relatively small size of the clinical sample may contribute to the lack of gender differences.

## CONCLUSION

The current study provides further evidence that the neuroticism subscale of the NEO-PI is a reliable and valid measurement tool for assessing neuroticism in Chinese population. Additionally, its measurement invariance across gender supported that the gender differences in neuroticism found by the Chinese version of the neuroticism subscale of the NEO-PI were reliable and valid.

Our research was restricted by specific Chinese samples, and the generalizability of results in this study to other countries remains to be determined. Studies with longitudinal research design and with larger clinical sample sizes are warranted.

## AUTHOR CONTRIBUTIONS

JY and SY: conceived and designed the study; MZ, XL, YiL, YuL and XZ: collected the data; CX, MZ, XL, and YiL: analyzed the data; CX and JY: wrote the paper.

## FUNDING

This work was supported by the National Natural Science Foundation [grant number: 81370034].

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01454/full#supplementary-material>



## REFERENCES

- Anastasi, A., and Urbina, S. (1997). *Psychological Testing*. Upper Saddle River NJ: Prentice-Hall.
- Banzhaf, A., Ritter, K., Merkl, A., Schulte-Herbrüggen, O., Lammers, C. H., and Roepke, S. (2012). Gender differences in a clinical sample of patients with borderline personality disorder. *J. Pers. Disord.* 26, 368–380. doi: 10.1521/pedi.2012.26.3.368
- Baumgartner, H., and Steenkamp, J. B. (1998). Multi-group latent variable models varying numbers of items and factors with cross-national and longitudinal applications. *Market Lett.* 9, 21–35. doi: 10.1023/A:1007911903032
- Briggs, S. R., and Venue, C. J. (1986). The role of factor analysis in the development and evaluation of personality scales. *J. Pers.* 54, 106–148. doi: 10.1111/j.1467-6494.1986.tb00391.x
- Browne, M. W., and Cudeck, R. (1993). “Alternative ways of assessing model fit,” in *Testing Structural Equation Models* (Newbury Park, CA), 136–162.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: awalk through the process. *Psicothema* 20, 872–882.
- Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equal. Model* 14, 464–504. doi: 10.1080/10705510701301834
- Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equal. Model.* 9, 233–255. doi: 10.1207/S15328007SEM0902\_5
- Chou, C. P., Bentler, P. M., and Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: a Monte Carlo study. *Br. J. Math Stat. Psychol.* 44, 347–357. doi: 10.1111/j.2044-8317.1991.tb00966.x
- Costa, P. T., and McCrae, R. R. (1988). Personality in adulthood: a six-year longitudinal study of self-reports and spouse ratings on the NEO personality inventory. *J. Pers. Soc. Psychol.* 54, 853–863. doi: 10.1037/0022-3514.54.5.853
- Costa, P. T., and McCrae, R. R. (1992a). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL.
- Costa, P. T., and McCrae, R. R. (1992b). Four ways five factors are basic. *Pers. Individ. Differ.* 13, 653–665. doi: 10.1016/0191-8869(92)90236-1
- Costa, P. T., and McCrae, R. R. (1997). Stability and change in personality assessment: the revised NEO personality inventory in the year 2000. *J. Pers. Assess.* 68, 86–94. doi: 10.1207/s15327752jpa6801\_7
- Costa, P. T., Terracciano, A., and McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *J. Pers. Soc. Psychol.* 81, 322–331. doi: 10.1037/0022-3514.81.2.322
- Cuijpers, P., Smit, F., Penninx, B. W., De, G. R., Ten, H. M., and Beekman, A. T. (2010). Economic costs of neuroticism: a population-based study. *Arch. Gen. Psychiatry*, 67, 1086–1093. doi: 10.1001/archgenpsychiatry.2010.130
- Dai, X., and Yao, S. (2004). Reliability and validity of the NEO-PI-R in Mainland China. *Chin. Ment. Health J.* 18, 171–174.
- Dai, X., Yao, S., and Cai, T. (1999). Cross-cultural personality assessment in psychiatric populations: the NEO-PI-R in the people's republic of China. *Psychol. Assess.* 11, 359–368. doi: 10.1037/1040-3590.11.3.359
- Daniel, K. M., Avron, S., and Avron, S. (2009). Do health behaviors explain the effect of neuroticism on mortality? longitudinal findings from the va normative aging study. *J. Res. Pers.* 43, 653–659. doi: 10.1016/j.jrp.2009.03.016
- DeVellis, R. F. (1991). *Scale Development: Theory and Applications*. Newbury Park, CA: SAGE Publication.
- Egger, J. I., De Mey, H. R., Derksen, J. J., and Staak, C. P. (2003). Cross-cultural replication of the five-factor model and comparison of the NEO-PI-R and MMPI-2 PSY-5 scales in a Dutch psychiatric sample. *Psychol. Assess.* 15, 81–88. doi: 10.1037/1040-3590.15.1.81
- Eysenck, H. E., and Michael, W. (1985). *Personality and Individual Differences: A Natural Science Approach*. New York, NY: Plenum Press.
- Eysenck, S., Eysenck, H., and Barrett, P. (1985). A revised version of the psychoticism scale. *Pers. Individ. Differ.* 6, 21–29. doi: 10.1016/0191-8869(85)90026-1
- Ferro, M. A., and Boyle, M. H. (2013). Longitudinal invariance of measurement and structure of global self-concept: a population-based study examining trajectories among adolescents with and without chronic illness. *J. Pediatr. Psychol.* 38, 425–437. doi: 10.1093/jpepsy/jss112
- Floyd, F. J., Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychol. Assess.* 7, 286–299.
- Foot, M., and Koszycki, D. (2010). Gender differences in anxiety-related traits in patients with panic disorder. *Depress. Anxiety.* 20, 123–130. doi: 10.1002/da.20031
- Grossman, M., and Wood, W. (1993). Sex differences in intensity of emotional experience: a social role interpretation. *J. Pers. Soc. Psychol.* 65, 1010–1022. doi: 10.1037/0022-3514.65.5.1010
- Hirsh, Y. (2011). Gender differences in personality across the ten aspects of the big five. *Front. Psychol.* 2:178. doi: 10.3389/fpsyg.2011.00178
- Hu, L., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol. Methods* 3, 424–453. doi: 10.1037/1082-989X.3.4.424
- Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model* 6, 1–55. doi: 10.1080/10705519909540118
- John, O., and Kentle, R. (1991). *The Big Five Inventory—Versions 4a and 5a*. Berkeley, CA: University of California.
- Joreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Ets Res. Bull.* 36, 409–426. doi: 10.1007/BF02291366
- Kendler, K. S., Kuhn, J., and Prescott, C. A. (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *Am. J. Psychiatry* 161, 631–636. doi: 10.1176/appi.ajp.161.4.631
- Kerkhof, B. (2003). Neuroticism in the elderly. The utility of the shortened DPQ-scales. *Tijdschrift Voor Gerontologie En Geriatrie* 34, 118–124.
- Kishton, J. M., and Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: an empirical example. *Edu. Psychol. Meas.* 54, 757–765. doi: 10.1177/0013164494054003022
- Konishi, Y., Tani, H., Otowa, T., Sasaki, T., Motomura, E., and Fujita, A. (2014). Gender-specific association between the comt val158met polymorphism and openness to experience in panic disorder patients. *Neuropsychobiology* 69, 165–174. doi: 10.1159/000360737
- Kotov, R., Schmidt, F., and Watson, D. (2010). Linking “big” personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychol. Bull.* 136, 768–821. doi: 10.1037/a0020327
- Lahey, B. B. (2009). Public health significance of neuroticism. *Am. Psychol.* 64, 241–256. doi: 10.1037/a0015309
- Little, T. D. (1997). Mean and covariance structures (MACS) analysis of cross-cultural data: practical and theoretical issues. *Multi. Behav. Res.* 32, 53–76. doi: 10.1207/s15327906mbr3201\_3
- Luchetti, M., Sutin, A. R., Delitala, A., Stephan, Y., Fiorillo, E., and Marongiu, M. (2018). Personality traits and facets linked with self-reported alcohol consumption and biomarkers of liver health. *Addict. Behav.* 82, 135–141. doi: 10.1016/j.addbeh.2018.02.034
- Luciano, M. (2015). Genome-wide association study identifies novel locus for neuroticism and shows polygenic association with major depressive disorder. *JAMA Psychiatry* 72, 642–650. doi: 10.1001/jamapsychiatry.2015.0554
- Luo, Y., He, C., and Luo, F. (2006). A study on anxiety status and influence factors of medical postgraduate students with self-rating anxiety scale. *Mod. Prev. Med.* 33, 2304–2306.
- Lynn, R., and Martin, T. (1997). Gender differences in extraversion, neuroticism, and psychoticism in 37 countries. *J. Soc. Psychol.* 137, 369–373. doi: 10.1080/00224549709595447
- Malouff, J. M., Thorsteinsson, E. B., and Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: a meta-analysis. *J. Psychopathol. Behav.* 27, 101–114. doi: 10.1007/s10862-005-5384-y
- Malouff, J. M., Thorsteinsson, E. B., and Schutte, N. S. (2006). The five-factor model of personality and smoking: a meta-analysis. *J. Drug Educ.* 36, 47–58. doi: 10.2190/9EP8-17P8-EKG7-66AD
- Matthews, G. (2000). *Stress, Health and Medical Psychology*. Cambridge: Cambridge University Press.
- McCrae, R. R., and Costa, P. T. (1997). Personality trait structure as a human universal. *Am. Psychol.* 52, 509–516. doi: 10.1037/0003-066X.52.5.509
- McCrae, R. R., and Costa, P. T. (2008). *The Five-Factor Theory of Personality*. New York, NY: The Guilford Press.
- Meredith, W. (1993). Measurement invariance, factor-analysis and factorial invariance. *Psychometrika* 58, 525–543. doi: 10.1007/BF02294825

- Merino, H., Senra, C., and Ferreiro, F. (2016). Are worry and rumination specific pathways linking neuroticism and symptoms of anxiety and depression in patients with generalized anxiety disorder, major depressive disorder and mixed anxiety-depressive disorder? *PLoS ONE* 11:e0156169. doi: 10.1371/journal.pone.0156169
- Nunnally, J. C. (1994). *Psychometric Theory*. New York, NY: McGraw Hill.
- Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., and Hankin, B. (2013). Neuroticism and common mental disorders: meaning and utility of a complex relationship. *Clin. Psychol. Rev.* 33, 686–697. doi: 10.1016/j.cpr.2013.04.003
- Ortet, G., Ibanez, M. I., Moya, J., Villa, H., Viruela, A., and Mezquita, L. (2012). Assessing the five factors of personality in adolescents: the junior version of the Spanish NEO-PI-R. *Assess* 19, 114–130. doi: 10.1177/1073191111410166
- Pervin, L. (1999). *Handbook of Personality: Theory and Research*. New York, NY: The Guilford Press.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psych. Meas.* 1, 385–401. doi: 10.1177/014662167700100306
- Reise, S. P., Moore, T. M., and Haviland, M. G. (2010). Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. *J. Pers. Assess.* 92, 544–559. doi: 10.1080/00223891.2010.496477
- Saklofske, D. H., and Janzen, B. L. (1995). Neuroticism, depression, and depression proneness. *Pers. Individ. Differ.* 18, 27–31. doi: 10.1016/0191-8869(94)00128-F
- Terracciano, A. (2003). The Italian version of the NEO PI-R: conceptual and empirical support for the use of targeted rotation. *Pers. Individ. Differ.* 35, 1859–1872. doi: 10.1016/S0191-8869(03)00035-7
- Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Pers. Soc. Psychol. Rev.* 15, 28–50. doi: 10.1177/1088868310366253
- Thorsten, B., and Tobias, C. (2010). Cognitive reactivity mediates the relationship between neuroticism and depression. *Behav. Res. Ther.* 48, 275–281. doi: 10.1016/j.brat.2009.12.005
- Veen, D., Dijk, S., Comijs, H. C., Zelst, W., Schoevers, R. A., and Voshaar, R. (2016). The importance of personality and life-events in anxious depression: from trait to state anxiety. *Aging Ment. Health* 21, 1177–1183. doi: 10.1080/13607863.2016.1202894
- Wang, M., Armour, C., Wu, Y., Ren, F., Zhu, X., and Yao, S. (2013). Factor structure of the CES-D and measurement invariance across gender in mainland chinese adolescents. *J. Clin. Psychol.* 69, 966–979. doi: 10.1002/jclp.21978
- Wang, W., and Tang, W. (2011). Confirmatory factor analysis of the zung self-rating anxiety scale. *China J. Health Psychol.* 19, 781–782.
- Widaman, K. F., and Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: applications in the substance use domain. *APA* 281–324. doi: 10.1037/10222-009
- Widiger, T. (2001). *Official Classification Systems*. New York, NY: Guilford Press.
- Xu, J., and Potenza, M. (2012). White matter integrity and five-factor personality measures in healthy adults. *Neuroimage* 59, 800–807. doi: 10.1016/j.neuroimage.2011.07.040
- Zung, W. W. (1965). A self-rating depression scale. *Arch Gen Psychiat.* 12, 63–70. doi: 10.1001/archpsyc.1965.01720310065008

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Xi, Zhong, Lei, Liu, Ling, Zhu, Yao and Yi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.