ASSESSING THE STRUCTURAL VALIDITY AND MEASUREMENT INVARIANCE OF THE PSYCHOLOGICAL PROBLEMS SCALE ON A SAMPLE OF INTERNATIONAL STUDENTS

Giusy Danila Valenti, Palmira Faraci, Giuseppe Craparo

Abstract

Objective: The main purpose of the current study is to examine the factorial structure of the Psychological Problems (PP) scale, which is an inventory used for assessing anxious and depressive symptoms. No previous studies were found in which the dimensionality of the measure has been examined.

Method: In order to test the underlying factorial structure of the scale, we performed both exploratory principal axis factor analysis and maximum likelihood robust confirmatory factor-analytic analysis on a sample of 266 international students. Parallel analysis was also computed to identify the number of factors to take into account. Multigroup confirmatory factor analyses were performed to evaluate the factorial invariance of the scale across gender.

Results: Our results yield a bifactorial structure. One item is removed due to its double-loading, resulting in a 14-item scale. Both of the subscales' alpha coefficients are good, with satisfying levels of adjusted item-scale correlations. Findings from the measurement invariance indicate adequate configural, metric, scalar and strict invariance across gender.

Conclusion: We conceptualize the two scales emerged from our factor analysis studies in terms of affective and psychosomatic domains of psychological problems. Despite some limitations, our findings may add new theoretical and practical implications, which are discussed also providing directions for future research.

Key words: psychological problems, depression; anxiety, affective; psychosomatic

Giusy Danila Valenti, Palmira Faraci, Giuseppe Craparo
Faculty of Human and Social Sciences, University Kore Enna

OPEN ACCESS

Corresponding author
Palmira Faraci, PhD, Full Professor
Faculty of Human and Social Sciences – University Kore Enna
Cittadella Universitaria - 94100 Enna, Italy
E-mail: palmira.faraci@unikore.it
Phone number: 0039 0935 53653

1. Assessing anxious and depressive symptoms: A general overview

Psychology has always had its primary focus on assessing mental illness and problems related to psychological health. Among the various psychological problems, anxiety and depression are two of the main investigated topics, due to their high relevance and clinical incidence. Indeed, international literature on anxiety and depression measurement is extensive and well examined, with a number of various assessment tools, developed worldwide for both clinical and research purposes.

Most of the commonly-used instruments are aimed to evaluate the presence (and severity) of clinical symptoms, such as the well-known Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), and the Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown 1996) for assessing depression, or the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988) for measuring anxiety. Some of them have been specifically developed for particular age groups, such as the Children's Depression Inventory (CDI; Kovacs, 1992) and the State-Trait Anxiety Inventory for Children (STAIC; Spielberger, 1973), aimed at assessing depressive and anxious symptoms during childhood, or the Geriatric

Depression Scale (GDS; Yesavage, Brink, Rose, Lum, Huang, Adey et al., 1983) and the Geriatric Mental State Examination (GMSE; Copeland et al., 1976) for the elderly. Further scales have been created to be employed in very specific and clearly defined contexts, such as the Pregnancy Anxiety Scale (PAS; Levin, 1991) and the Edinburgh Postnatal Depression Scale (EPDS; Cox, Holden, & Sagovky, 1987), addressed to women during the pre- and post-natal period.

Although depressive and anxiety disorders can be considered as two distinct and well distinguishable concepts, they are frequently examined jointly, due to their common co-occurrence and comorbidity (Hirshfeld, 2001; Lamers et al., 2011). This could explain the reason why some well-established scales assess both of the disorders. The Depression Anxiety Stress Disorders (DASS-21; Lovibond & Lovibond, 1995) and the Hospital Anxiety Depression Scale (HADS; Zigmond & Snaith, 1983) are two suitable examples.

It is not surprising the existence of a broad number of measures aimed to evaluate anxious and depressive symptoms. As for almost all of the psychological constructs, the presence of different assessment tools derives not only from the specific context and the particular population they are addressed to, but also from the way the constructs are conceptualized. It is

indeed very common that psychological inventories measuring the same construct differ each other due to their theoretical framework.

After that, besides their clinical incidence, sub-clinical manifestations of depression and anxiety often occur. In everyday life, everyone has experienced at least once feelings of sadness, worthlessness, unhappiness or worry and unease about something with an uncertain outcome, and the more stressful the events are for an individual, the more likely these symptoms can appear. This means that, in some stressful situations, although anxious and depressive symptoms can easily and frequently take place, they may not necessarily imply the occurrence of the clinical disorders.

2. Assessing anxious and depressive symptoms in acculturation research

Moving from one own country of origin to a new socio-cultural context may negatively affect psychological health since cross-cultural transitions are often characterized by a large set of stressors. For instance, facing with difficulties in speaking another language and in living in a socio-cultural environment with different values, rules, and habits may decrease the levels of psychological wellness. Staying apart from family members and friends can also be a source of stress, which may lead the appearance of homesickness or loneliness. From this perspective, the effects of cross-cultural transitions and intercultural relationships on adaptive outcomes are the core themes of the acculturation field of study.

In cross-cultural psychology research there seems to be a wide agreement about the dichotomization of adaptation into two different and interrelated domains: sociocultural and psychological (Searle & Ward, 1990). The former, defined in terms of behavioural competence, concerns the ability to adjust and to achieve culturally appropriate competences of the host environment, and it is mainly predicted by cultural knowledge, cultural distance, degree of contact, and positive intergroup attitudes (Brisset et al., 2010; Ward & Kennedy, 1999); the latter refers to feelings such as satisfaction or well-being, and it is mainly influenced by personality traits, coping strategies, and social support (Brisset et al., 2010; Ward & Rana-Deuba, 2000). In order to evaluate the two distinguished dimensions of adaptation – considering their conceptual and empirical distinction – different assessment scales have been used.

The socio-cultural domain has usually been evaluated using the Socio-Cultural Adaptation Scale (SCAS; Ward & Kennedy, 1999), which is a flexible measure addressing the required abilities to face daily social situations when adjusting in new cultural contexts. The psychological domain has been conceptualized as a lack of psychological problems (i.e., distress, anxiety, depression, mood disturbances, negative emotions, loneliness or feelings of worthlessness), and presence of wellness, such as life satisfaction and high levels of self-esteem (Brisset et al., 2010; Zhang & Goodson, 2011), experienced during the intercultural contacts and transitions in new and culturally different environments. In order to evaluate the psychological domain, well established scales have been used. Generally, in almost all of the studies, three specific indicators of psychological adaptation were taken into account: (1) self-esteem, (2) life satisfaction, and (3) psychological problems (e.g., anxiety and depression). Self-esteem is usually measured by means of the well-known Rosenberg Self-Esteem Scale (SES; Rosenberg, 1965), whereas the most common assessment tool used for evaluating life satisfaction is the Satisfaction With Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985). Both of them were applied in numerous studies and with different acculturating samples (Gui, Safdar, & Berry, 2016; Savicki 2010; Stuart, Ward, & Robinson, 2016).

Regarding the evaluation of the psychological problems, an agreement about which measurements to use does not exist. For example, Fritz, Chin, and De Marinis (2008) employed the Anxiety Symptom Checklist-90 (SCL-90; Derogatis, 1977), a 15-item scale in which participants had to rate how often they had experienced feelings of nervousness and shakiness, tension and trembling, fearfulness, apprehension and dread during the previous month. Jung, Hecht, and Chapman Wadsworth (2007), and Zhang and Goodson (2011) used a shortened version of the Centre of Epidemiologic Study of Depression (CES-D; Radloff, 1977), assessing depression level by four factors: depressive affect, somatic symptoms, well-being, interpersonal relationships. Zhang, Mandl, and Wang (2010) used the Chinese version of the Zung Self-rating Depression Scale (Wang, Wang, & Ma, 1999; Zung, 1965), a self-report inventory composed of 20 items evaluating distinct components of depression (i.e., psychological, physiological, and affective domains). Cemalcimar and Falbo (2008) assessed psychological wellness with the Hudson's (1987) Generalized Contentment Scale (GCS) for evaluating general levels of mental well-being, whereas Wang and Mallinckrodt (2006) applied the 21-item Anxiety Depression Scale (DASS-21; Lovibond & Lovibond, 1995).

These examples clearly show how assessing psychological problems in cross-cultural research is vague and indefinite. Although acculturation studies are strictly related to the specific examined context and acculturating sample, using the same assessment tool for measuring psychological problems could represent a useful strategy to improve the different studies comparability.

3. The Psychological Problems (PP) scale: A new measure for evaluating anxious and depressive symptoms in cross-cultural research

In order to assess anxious and depressive symptoms during intercultural interactions and cross-cultural transitions, a new measure has recently been developed. The scale, named Psychological Problems (PP; Berry, 2017), is commonly applied in the extended literature of intercultural contacts and it is widely used as an indicator of psychological adjustment.

As already mentioned, findings from a broad number of studies have showed how cross-cultural transitions negatively affected psychological health and wellness on different acculturating samples and different socio-cultural contexts (Cemalcimar & Falbo, 2008; Fritz, Chin, & De Marinis, 2008; Zhang & Goodson, 2011). Although several well-established scales evaluating psychological problems already exist, none of them have been specifically developed for this particular field of study. In other words, the Psychological Problems scale has been created to measure people's feelings in adjusting to a new sociocultural context.

The PP scale is a brief measure composed of 15 short items describing both depressive and anxious symptoms. Its items derive from previous existing measures, aimed at evaluating mood disturbances in acculturating settings and populations (Mollica, Wyshak, Demarneffe, Khuan, & Lavelle, 1987; Robinson, Shaver, & Wrightsman, 1991). All of them, having a high level of legibility and understandability, are easy to be filled out also by

non-native English speakers. This feature, undoubtedly, represents a key strength of the measure, since it makes it adequate to be administered also to participants whose mother tongue is not English. It is worthy to clarify that the PP scale has been translated in other languages (Berry, 2017), in order to make items much easier to be understood by specific acculturating groups (e.g., Italian, Portuguese, Spanish, Finnish, and Chinese speaking people). Participants have to indicate the frequency they experience mood disturbances related to anxious and depressive symptoms, using a Likert scale ranging from 1 (Never) to 5 (All the time). Elevated scores indicate more psychological symptoms.

The measure has been psychometrically tested through several studies, especially in terms of internal consistency, reaching satisfactory results (Grad, 2017; Inguglia, Musso, & Lo Coco, 2017; Neto & Neto, 2017). Nevertheless, Berry (2017) did not report factor-analytic studies, and no further researches examining its dimensionality have been conducted. In the light of these considerations, empirical evidence concerning the underlying structure of the PP scale is still required. According to the need to establish the factorial composition of widely used measures in psychological research (Faraci & Tirrito, 2013), our study is addressed to fill this gap.

Method

Participants and Procedure

The total group of participants consisted of 266 college students (mean age ± SD: 24.23 ± 4.59 years) of both sexes of whom 131 (49.2%) were men and 135 (50.8%) were women. They were international or visiting students/scholars ($n = 213$) at University of California, Los Angeles (UCLA) or at two English language schools ($n = 53$).

They were asked to specify the length of stay in the U.S.A., which ranged from 1 month to 10 years (mean ± SD: 27 ± 25.09). Regarding the country of origin, 186 (69.9%) came from Asia, 41 (15.5%) from Europe, 29 (10.9%) from North and South America, 9 (3.4%) and 1 (.4%) from (Africa and Australia).

A subgroup of 150 participants was randomly selected to perform exploratory factor analysis. It was made up by 72 (48%) men and 78 (52%) women, with a mean age of 24.85 (SD = 5.08). Regarding the country of origin, 103 (68.7%) were from Asia, 22 (14.7%) from Europe, 19 (12.7%) from North and South America, 5 (3.3%) and 1 (.7%) from (Africa and Australia). They indicated their length of stay in U.S.A. (mean ± SD: 26 + 24.59). A large part of them ($n = 114$) was composed of UCLA international students, whereas 35 participants attended the two English schools. Subsequent analyses were conducted on the total sample.

Data at UCLA were collected individually, by informal contacts with the international students met all around the campus (libraries, restaurants, shops), whereas at the two English Language Schools one of the authors was introduced by the English teacher in the class, where the questionnaire was administered to small groups of 8-10 international students. In both cases, participants were informed about the aim of the research. Data were gathered from January to March 2019.

Instrument

The PP scale is a 15-item self-report questionnaire describing anxious and depressive symptoms. A Likert scale from 1 (Never) to 5 (All the time) is used to compute a unique and general score, considering that elevated scores have to be interpreted as higher psychological problems.

Data analyses

Data normality was evaluated by applying both Kolmogorov-Smirnov and Shapiro-Wilk tests. Bartlett's test of sphericity was inspected in order to evaluate items correlations, whereas the measure of sampling adequacy was computed to assess items shared variance. An exploratory principal axis factor analysis with promax rotation was worked out. Factor extraction was ruled by random parallel analysis, together with Scree test inspection and Kaiser criterion application. Internal consistency reliability was established by checking Cronbach's α values.

The fit of the theoretical model to the observed data was verified through a ML robust confirmatory factor analysis, checking the following goodness of fit indices: Comparative Fit Index, Non-Normed Fit Index, Standardized Root Mean Square Residual, Root Mean Square Error of Approximation, and Chi^2 to degrees of freedom ratio. The CFI and NNFI values equal or greater than .90, SRMR values equal or lower than .10, and RMSEA values equal or lower than .08 have been claimed as useful indicators of acceptable fit (Browne, 2015; Hu & Bentler, 1999).

Multigroup Confirmatory Factor Analysis (MCFAs) was run in JASP to evaluate the factorial stability across male and female participants. Configural, metric, scalar, and strict invariance across gender were tested. Firstly, the configural invariance test was used in order to evaluate the dimensional stability across sex. The configural invariance is an unconstrained model, in which parameters to be estimated are allowed to vary freely, and it represents the required prerequisite to compare all subsequent models. Metric invariance was obtained by adding constraints on the factor loadings to the base model, assessing whether the relationships between the factors and the items were invariant across the two different groups. After that, in order to test whether there was a significant difference between the two comparative models, the CFI, NNFI and χ^2 difference test was computed. When this discrepancy is non-significant, it indicates that factor loadings do not change across groups (Byrne, 2001). Strict invariance was tested by adding the residual equality constraint, indicating whether the scale shows the same pattern of error variance between groups. As previously, in order to examine model invariance, we used ΔCFI, ΔNNFI, and $\Delta\chi^2$ to compare the two nested models. The achievement of measurement invariance may allow us to claim that the PP items measure the same dimensions for both male and female participants.

Results

Item distribution

PP items descriptives and normality inspection are shown in **table 1**. Since both Kolmogorov-Smirnov and Shapiro-Wilk tests of normality were statistically significant, we decided to employ principal axis factoring method for EFA, and ML robust estimation for CFA (Perdighe et al., 2015).

Table 1. *Item analysis for PP items*

Item	*M*	*SD*	Skewness	Kurtosis	Kolmogorov-Smirnov test	Shapiro-Wilk test
Item 1	3.23	.86	-.21	-.12	.22***	.88***
Item 2	2.27	.84	-.37	-.33	.28***	.86***
Item 3	1.91	.82	.59	-.26	.24***	.83***
Item 4	1.77	.85	1.11	1.16	.26***	.79***
Item 5	1.91	.79	.64	.35	.25***	.83***
Item 6	2.47	.97	.27	-.50	.23***	.90***
Item 7	2.51	.91	.20	-.22	.22***	.90***
Item 8	2.52	.94	.30	-.27	.23***	.89***
Item 9	2.48	.82	.25	-.06	.25***	.87***
Item 10	2.61	1.04	.46	-.23	.22***	.90***
Item 11	2.42	.76	.34	.63	.27***	.84***
Item 12	2.73	.90	.06	-.14	.23***	.89***
Item 13	2.71	1.02	.40	-.38	.23***	.89***
Item 14	2.35	.98	.53	-.10	.25***	.88***
Item 15	2.10	.90	.71	.40	.27***	.85***

***$p < .001$.

Exploratory principal axis factor analysis

The appropriateness for a factor-analytic study was supported by our results from Bartlett's and Kaiser-Meyer-Olkin Tests: $\chi^2(91) = 783.702$, $p < .001$ and .88, respectively. Parallel analysis outcomes, together with the interpretation of the line plot of the factors' eigenvalues and the Kaiser's rule, suggested the retention of two dimensions, with an acceptable proportion of explained variance (52.091%). Item 1 *"I feel tired"* was removed because it loaded simultaneously on both factors (.347 on the first factor and .327 on the second factor).

Factor 1 was composed of 10 items, and Item 11 *"I feel unhappy and sad"* and item 13 *"I worry a lot of the time"* showed the greatest loadings. Factor 2 was comprised of 4 items, and Item 3 *"I feel dizzy and faint"* showed the highest loading. **Table 2** displays the entire factor solution of the PP scale. **Figure 1** depicts the scree plot. Both of the subscales' alpha coefficients were good, with satisfying levels of adjusted item-scale correlations (see **table 3**). **Table 4** shows the factor correlation matrix.
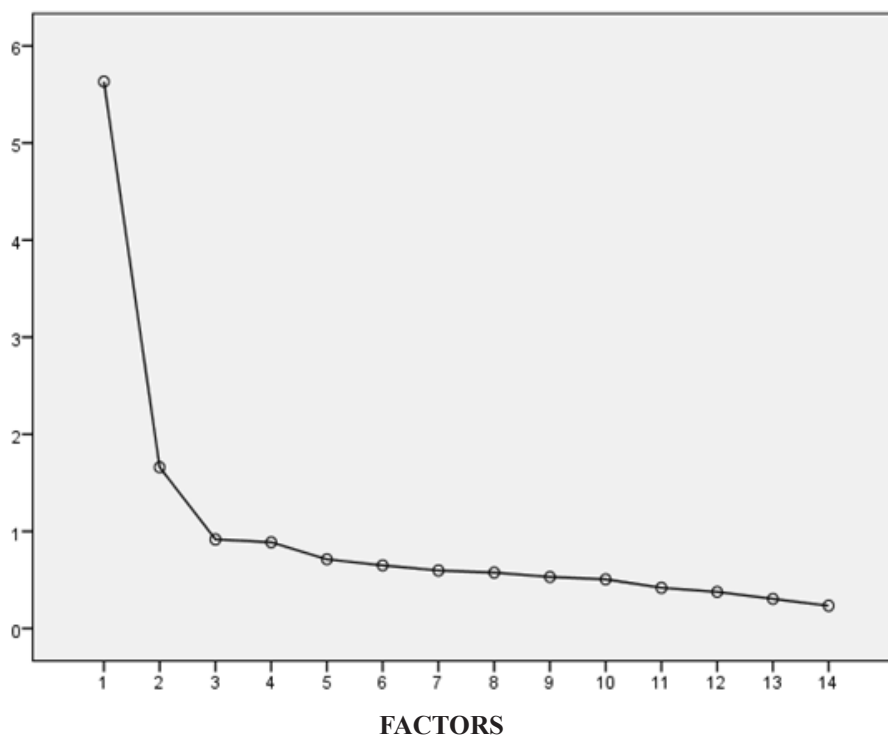
Figure 1. *Scree plot*

FACTORS

Table 2. *Factor solution of the PP items*

Item	F1	F2	Initial	Extraction
13. I worry a lot of the time	**.779**	-.027	.367	.393
11. I feel unhappy and sad	**.748**	-.100	.456	.540
12. My thoughts seem to be mixed up	**.696**	-.270	.425	.486
14. I feel lonely even with other people	**.657**	-.036	.430	.442
8. I feel restless	**.607**	.181	.549	.533
9. I feel annoyed or irritated	**.577**	-.014	.564	.559
7. I feel nervous and shaky inside	**.574**	.248	.554	.531
10. I am worried about something bad happening to me	**.559**	.013	.329	.323
6. I feel tense or keyed up	**.532**	.277	.375	.321
15. I lose interest and pleasure in things which I usually enjoy	**.390**	.798	.493	.482
3. I feel dizzy and faint	-.202	**.835**	.325	.337
4. I feel short of breath even when not exerting myself	-.032	**.716**	.596	.583
2. I feel sick in the stomach	-.059	**.659**	.404	.405
5. I feel weak all over	.229	**.504**	.319	.277
% explained variance	40.234	11.857		
Eigenvalues	5.633	1.660		

Note. *F1 = Affective domain; F2 = Psychosomatic domain.*

**Table 3.** *Item-total correlations and Cronbach's alpha*

Item	F1	F2
Item 13	.703	
Item 11	.629	
Item 12	.486	
Item 14	.592	
Item 8	.678	
Item 9	.516	
Item 7	..671	
Item 10	.500	
Item 6	.652	
Item 15	.466	
Item 3		.601
Item 4		.591
Item 2		.512
Item 5		.497
Cronbach's alpha	.87	.75

Note. *F1 = Affective domain; F2 = Psychosomatic domain.*

Table 4. *Factor correlation matrix*

	F1	F2
F1. Affective domain	–	
F2. Psychosomatic domain	.588	–

Figure 2. *Standardized solution of the empirical model*
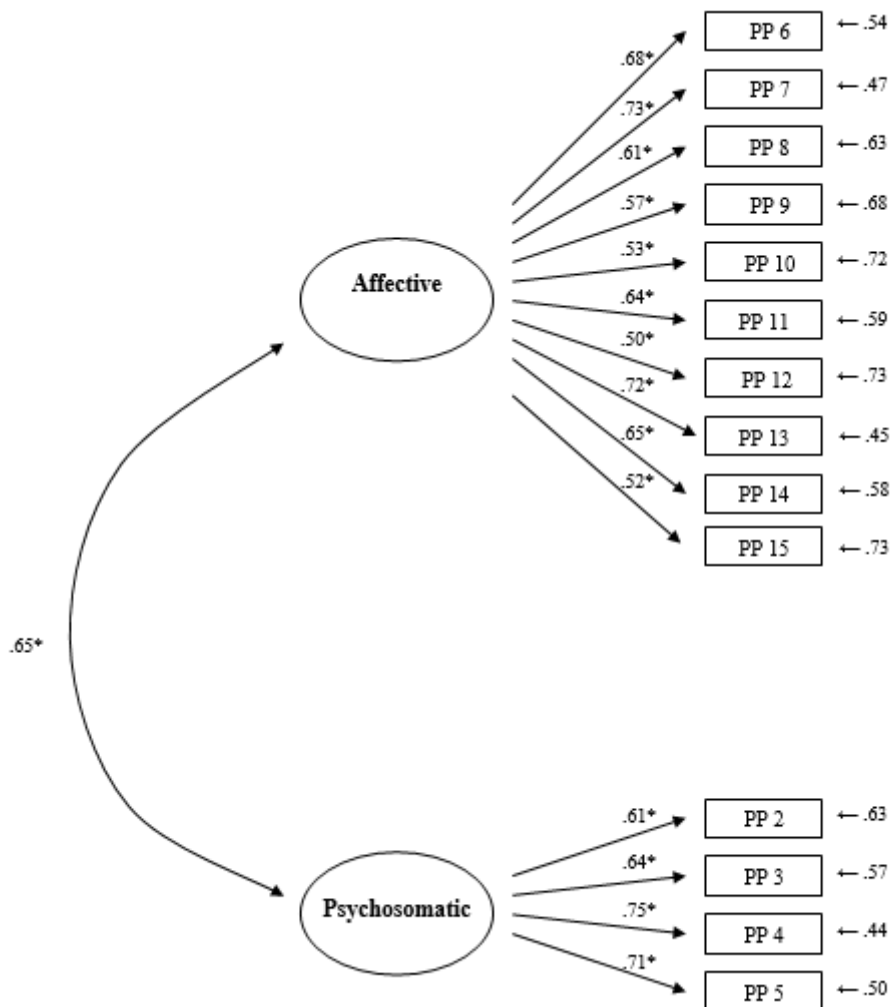
*Note. *p < .05*

ML robust confirmatory factor-analytic analysis

After inspecting the modification indices, a correlation between item 3 and item 6 and between item 10 and item 13 was added. The Maximum Likelihood robust confirmatory factor-analytic analysis revealed a plausible goodness of fit for a bi-factorial model: $\chi^2(74, N = 266) = 156.011$, $p < .001$; $\chi^2/df = 2.10$; CFI = .93; NNFI = .92; RMSEA = .066, 90% confidence interval = .051–.080; SRMR = .053. **Figure 2** shows the standardized solution of the empirical model.

Test of invariance

In order to verify whether the PP scale showed the same factorial structure across gender, measurement of invariance tests were performed. The first step consisted of evaluating the configural invariance by means of an unconstrained model in which parameters to be estimated were allowed to vary freely. Based on the inspection of

measurement errors, a correlation between item 3 and item 13 residuals was added. Though a significant χ^2 was obtained, all other fit indices were above the cut-off of acceptability. The configural invariance was then supported. When factor loadings were constrained to be equal, the χ^2 did not increase significantly, and change in CFI and in NNFI were not greater than the critical values, suggesting that the factors-items relationships are consistent across gender. After that, the scalar invariance was assessed by constraining the intercepts. As with the previous models, even if the χ^2 value was significant, all other indices showed an acceptable fit. A non-significant χ^2 difference, a discrepancy in CFI < .01, and a change in NNFI < .05 showed that the scalar invariance was satisfied. Finally, the strict invariance was examined, adding the residual equality constraint to the previous model. Also the residual invariance model was accepted, not reporting neither significant differences in χ^2 nor discrepancies in both CFI and NNFI above the critical values. Fit indices for invariance tests across gender are displayed in **table 5**.

Discussion

Assessing the occurrence of psychological problems in stressful situations is one of the major and more common topics in several branches of psychology. Among the broad number of available instruments, the Psychological Problems scale was developed to evaluate the presence of anxious and depressive symptoms during cross-cultural transitions (Berry, 2017). Despite its wide application in cross-cultural research, empirical evidence about the factorial structure of the PP scale is lacking. The current research can be seen as the first attempt aimed to evaluate the underlying factorial composition of the measure.

Our findings derived from the exploratory factor-analytic study show a bifactorial model: the first factor measures feelings, such as irritation, worry, loneliness, loss of interests, and unhappiness, whereas the second factor reflects psychosomatic symptoms related to anxiety and depression, including weakness, dizziness, and breathlessness. In other words, it seems that the first factor describes the affective domain linked to both depression and anxiety, while the second factor deals with physical problems associated with these mood disturbances. For this reason, Factor 1 may be called *Affective domain,* and Factor 2 may be named *Psychosomatic domain.*

By applying the confirmatory factor-analytic techniques and inspecting the obtained fit indices, the bifactorial structure of the scale was supported, providing evidence of a plausible goodness-of-fit to the empirical data. Moreover, results from the measurement invariance indicated that the 14-item PP scale featured adequate configural, metric, scalar, and strict invariance across gender. This means that the measure shows the same factorial structure both for male and female participants, and that PP items equally function across gender.

Interestingly, our results do not highlight a distinction between anxious and depressive symptoms. From this perspective, if one is interested in evaluating and discriminating between the two different disorders, the PP scale does not seem to satisfy this aim. Other existing measures, such as the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) or the Depression Anxiety Stress Scale (DASS; Lovibond & Lovibond, 1995) distinctively assess both psychological problems. However, they use semantically distinct items, distributed in two subscales, allowing to easily distinguish anxious symptoms from depressive ones. In this regard, we argue that a viable way to conceptualize the two revealed dimensions consists in taking into account the affective and the psychosomatic domain of psychological problems.

Some authors (Clark & Watson, 1991; Eysenck, 1970; Zurawski & Smith, 1987) have suggested the concepts of Negative Affectivity (NA), Neuroticism, and General Psychological Distress, as the common and underlying factor including a broad range of aversive emotional states related to both anxiety and depression. The existence of a general and unspecific factor may explain the reason why scores on anxiety and depression measures are often intercorrelated (Watson, 2009). Considering their large amount of shared portion of a higher order factor, one may be interested in assessing the general level of psychological problems, rather than discriminating between the two aforementioned disturbances. From this point of view, the PP scale addresses this scope. Furthermore, the identified bifactorial structure could have intriguing and useful practical implications: affective and psychosomatic domains are two well distinguished manifestations of general psychological distress, with a possible different impact on performance and general wellness. The affective domain might mostly affect emotional stability or emotional regulation, whereas the psychosomatic domain might mainly affect behavioural outcomes.

Limitations and future directions

Our results should be considered in light of some limitations. First, since we used a convenience sample of international students in the U.S.A., generalization beyond the study population cannot be properly done, as this sampling procedure does not guarantee adequate representativeness of the population. Second, due to the high level of disparity regarding the ethnic background, we could not carry out multi-sample analyses to test structural invariance nor levels of psychological problems according to the country of origin. Interestingly, some of the existing empirical studies reported statistically significant differences in psychological wellness related to international students' socio-cultural background (Fritz, Chin, & Demarinis, 2008; Wang & Hannes, 2014). The large range of participants' age and length of stay in the U.S.A. might also represent a further source of sampling bias. We also recommend that researchers, when examining the PP dimensionality, take into account possible external variables that may affect and alter responses on the questionnaire. For instance, the Covid-19 diffusion might bias responses on questionnaire. From this perspective, since our research was performed prior to the pandemic, studies conducted in the current period should be careful in interpreting results.

Table 5. *Measurement invariance analyses across gender*

	Fit indices						Change model fit statistics and fit indices			
	χ^2	*df*	CFI	NNFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	ΔCFI	ΔNNFI
Configural Invariance	234.23[a]	146	.93	.92	.069	.065	-	-	-	-
Metric Invariance	250.13[a]	158	.92	.91	.067	.080	15.9	12	-.01	-.01
Scalar Invariance	262.48[a]	170	.92	.92	.065	.077	12.35	12	.00	.01
Strict Invariance	302.32[a]	187	.91	.91	.069	.084	39.84	17	-.01	-.01

Note. CFI = Comparative Fit Index; NNFI = Non-Normed Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.
[a]*Models are at < .001 significant.*

Future research should be addressed at providing additional evidence of the factorial structure of the PP scale, including wider and more representative samples of individuals. Moreover, since the PP scale is largely used in cross-cultural research, future studies should also examine its dimensionality across different acculturating populations, such as immigrants. Likewise, since versions of the PP scale translated in other languages do exist (Grad, 2017; Inguglia, Musso, & Lo Coco, 2017; Neto & Neto 2017), it may be remarkable to evaluate if the same factorial structure could emerge in language versions other than English.

Conclusion

To summarize, the PP scale can be considered as a suitable measure when the primary goal consists in evaluating the levels of general psychological wellness. If included in a longer and more complex set of questionnaires, the PP scale is an advantageous measure for psychological well-being to take into account, because with only 14 items it allows to provide a general indicator of psychological health. Besides, though the PP scale has been created and widely applied in cross-cultural psychology research, it may be equally used in other sectors and branches of knowledge, since its items are not context-dependent.

References

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An Inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consultation and Clinical Psychology, 56,* 893-897.

Beck, A. T., Steer, R. A., & Brown, G. (1996). *BDI-II: Beck Depression Inventory,* 2nd ed. San Antonio, TX: Psychological Corporation.

Beck, A. T., Ward, C. H. Mendelson, M., Mock, N., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Berry, J. W. (2017). *Mutual intercultural relations.* Cambridge: Cambridge University.

Brisset, C., Safdar, S., Lewis, J. R., & Sabatier, C. (2010). Psychological and sociocultural adaptation of university students in France: The case of Vietnamese international students. *International Journal of Intercultural Relations, 34,* 413-426.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* NewYork, NY: Guilford Press.

Byrne, B. M. (2001). *Structural equation modelling with AMOS: Basic concepts, applications, and programming.* Hillsdale, NJ: Erlbaum.

Cemalcimar, Z., & Falbo, T. (2008). A longitudinal study of the adaptation of International students in the United States. *Journal of Cross-Cultural Psychology, 39*(6), 799-804.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316-336.

Copeland, J. R. M., Kelleher, M. J., Kellett, J. M., Gourlay, A. J., Gurland, B. J., Fleiss, J. L., & Sharpe, L. (1976). A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly: The Geriatric Mental State Schedule I: Development and reliability. *Psychological Medicine, 6,* 439–449.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburg Postnatal Depression Scale. *British Journal of Psychiatry, 150,* 782-786.

Derogatis, L. R. (1977). *SCL-90: Administration scoring and procedures manual-I and other instruments of the psychopathology rating scale series.* Baltimore, MD: Johns Hopkins University School of Medicine.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49,* 71-75.

Eysenck, H. (1970). *The structure of human personality.* London: Methuen.

Faraci, P., & Tirrito, A. (2013). Fifty years studying the Beck Depression Inventory (BDI) factorial stability without consistent results: A further contribution. *Clinical Neuropsychiatry, 10*(6), 274-279.

Fritz, M.V., Chin, D., & Demarinis, V. (2008). Stress, anxiety, acculturation, and adjustment among international and North American students. *International Journal of Intercultural Relations, 32,* 244-259.

Grad, H. (2017). Intercultural relations in Spain. In J. W. Berry (Ed.), *Mutual intercultural relations* (pp.268-289). Cambridge: Cambridge University Press.

Gui, Y., Safdar, S., & Berry, J. W. (2016). Mutual intercultural reactions among university students in Canada. *Frontiers: The Interdisciplinary Journal of Study Abroad, 27,* 17-32.

Hirshfeld, R. M. A. (2001). The comorbidity of major depression and anxiety disorders: Recognition and management in primary care. *Journal of Clinical Psychiatry, 3*(6), 244-254.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Hudson, W. W. (1987). Generalized contentment scale. In J. Fischer & K. Corcoran (Eds.), *Measures for clinical practice: A sourcebook* (pp. 165-167). New York: Free Press.

Inguglia, C., Musso, P., & Lo Coco, A. (2017). Intercultural relations in Italy. In J.W. Berry (Ed.), *Mutual intercultural relations* (pp.210-230). Cambridge: Cambridge University Press.

Julian, L. J. (2011). Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care and Research, 63*(11), 467–477.

Jung, E., Hecht, M. L., & Champman Wadsworth, B. (2007). The role of identity in international students' psychological well-being in the United States: A model of depression level, identity gaps, discrimination, and acculturation. *International Journal of Intercultural Relations, 31,* 605-624.

Kovacs, M. (1992). *Children's Depression Inventory.* North Tonawanda, NY: Multi-Health System.

Lamers, F., van Oppen, P., Comijis, H. C., Smit, C. H., Spinhoven, P., van Balkom, A. J. L., et al. (2011). Comorbidity pattens of anxiety and depressive disorders in a large cohort study: The Netherlands study of depression and anxiety (NESDA). *Journal of Clinical Psychiatry, 72*(3), 341-348.

Levin, J. S. (1991). The factor structure of the Pregnancy Anxiety Scale. *Journal of Health and Social Behavior, 32,* 368–381.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney: Psychology Foundation.

Mollica, R., Wyshak, G., DeMarneffe, D., Khuon, F., & Lavelle, J. (1987). Indochinese versions of the Hopkins Symptom Checklist-25: A screening instrument for the psychiatric care of refugees. *American Journal of Psychiatry*, *144,* 497–500.

Neto, F., & Neto, J. (2017). Intercultural relations in Portugal. In J. W. Berry (Ed.), *Mutual Intercultural Relations* (pp.249-267). Cambridge: Cambridge University Press.

Perdighe, C., Cosentino, T., Faraci, P., Gragnani, A., Saliani, A., & Mancini, F. (2015). Individual differences in guilt sensitivity: The Guilt Sensitivity Scale (GSS). *Testing, Psychometrics, Methodology in Applied Psychology, 22*(3), 349-362.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in general population. *Applied Psychological Measurement, 1,* 385–401.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.) (1991). *Measures of personality and social psychological attitudes.* San Diego, CA: Academic Press.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Savicki, V. (2010). Implications of Early Sociocultural Adaptation for Study Abroad Students. *The Interdisciplinary Journal of Study Abroad*, *19,* 205-223.

Searle, W., & Ward, C. (1990). The prediction of psychological and sociocultural adjustment during cross-cultural transitions. *Intercultural Journal of Intercultural Relations, 4,* 449-464.

Spielberger, C. D. (1973). *Manual for the State–Trait Anxiety Inventory for Children.* Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gorsush, R. L, & Lushene, R. (1970). *STAI manual.* Palo Alto, CA: Consulting Psychologist Press.

Stuart, J., Ward, C., & Robinson, L. (2016). The influence of family climate on stress and adaptation for Muslim immigrants young adults in two Western Countries. *International Perspectives in Psychology, 5*(1), 1–17.

Wang, C. -C.D.C. & Mallinckrodt, B. (2006). Acculturation, attachment, and psychological adjustment of Chinese/Taiwanese international students. *Journal of Counseling Psychology, 53*(4), 422-433.

Wang, Q., & Hannes, K. (2014). Academic and socio-cultural adjustment among Asian international students in the Flemish community of Belgium: A photovoice project. *International Journal of Intercultural Relations, 39,* 66–81.

Wang, X., Wang, X., & Ma, H. (1999). *Rating scales for mental health.* Beijing: Chinese Mental Health Journal Press.

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II (BDI-II): A comprehensive review. *Revista Brasileira de Psiquiatria, 35*(4), 416-431.

Ward, C., & Kennedy, A. (1999). The measurement of sociocultural adaptation. *International Journal of Intercultural Relations, 23*(4), 659-677.

Watson, D. (2009). Differentiating the mood and anxiety disorders: A quadripartite model. *Annual Review of Clinical Psychology, 5,* 221-247.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37-49.

Zhang, J., & Goodson, P. (2011). Acculturation and psychosocial adjustment of Chinese international students: Examining mediation and moderation effects. *International Journal of Intercultural Relations, 35,* 614-627.

Zhang, J., & Goodson, P. (2011). Predictors of international students' psychosocial adjustment to life in the United States: A systematic review. *International Journal of Intercultural Relations, 35,* 139–162.

Zhang, J., Mandl, H., & Wang, E. (2010). Personality, acculturation, and psychological adjustment of Chinese international students in Germany. *Psychological Reports, 107*(2), 511-525.

Zigmond, A. S. & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67,* 361-370.

Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry, 12,* 615-630.

Zurawski, R. M., & Smith, T. W. (1987). Assessing irrational beliefs and emotional distress: Evidence and implications of limited discriminant validity. *Journal of Counseling Psychology, 34,* 224-227.