

Psychometric properties and measurement invariance of Depression, Anxiety and Stress Scales (DASS-21) across cultures

Akhtar Bibi, Muyu Lin, Xiao Chi Zhang, and Jürgen Margraf

Mental Health Research and Treatment Center, Faculty of Psychology, Ruhr-Universität Bochum, Bochum, Germany

Cross-culturally measurement invariant instruments are useful for the assessment of mental symptoms across cultures. The current study aimed to investigate the psychometric properties and measurement invariance of the Depression, Anxiety and Stress Scales (DASS-21) (DASS) across Pakistan and Germany. German participants were recruited through an online survey ($N = 1323$), while Pakistani participants were recruited through online survey or paper-and-pencil survey ($N = 1841$). The DASS-21 showed good reliability, construct and structure validity in both countries. The DASS-21 indicated partial weak (stress subscale) and partial strong measurement invariance (depression and anxiety subscales) between both countries. Latent mean comparison of depression and anxiety symptoms between Pakistani and German students indicated that Pakistani university students experience more symptoms of depression and anxiety. Results suggest that the DASS-21 could be used in Pakistan and Germany, but caution should be taken when making direct comparisons between the two countries.

Keywords: Measurement invariance; Cross-cultural comparison; DASS-21; Pakistan; Germany.

Mental disorders affect the lives of people of all ages across countries. Among mental disorders, depressive disorders account for 40.5% of Disability Adjusted-Life Years (DALYs), while anxiety disorders account for 14.6% of DALYs. The rapid increase in anxiety and mood disorders is a serious, worldwide health challenge. Depression and stress are the most common cause of mental health problems among university students. Depression, anxiety and stress have been shown to be strongly and negatively linked to positive mental well-being (Bibi, Lin & Margraf, 2020), life satisfaction (Marum, Clench-Aas, Nes, & Raanaas, 2014), resilience (Teismann et al., 2018), and social support (Lin, Hirschfeld, & Margraf, 2018). Therefore, it is vital to apply well-established and cross-culturally invariant measurements when assessing the symptoms of depression, anxiety and stress in culturally diverse student samples.

One of the well-established tools for assessing symptoms of depression, anxiety and stress is the Depression, Anxiety and Stress Scales (DASS) developed by Lovibond and Lovibond (1995). The DASS-21, with its 21 items (7 items for each subscale) and three dimensions with similar psychometric properties, is based on the tripartite model of depression, anxiety and stress. The scale has several benefits; it is brief and can be applied in clinical as well as general populations.

So far, the DASS-21 is available in 42 languages and is widely accessible to clinicians and researchers. Generally, instruments and interventions developed in the United States or Europe are directly transferred and applied to other cultures as proximate translations without cultural adjustments (F. M. Cheung, 2012). Such a process is appropriate if the scale's features and constructs are applicable across cultures. However, this requirement is seldom fulfilled. For instance, a study of

Correspondence should be addressed to Akhtar Bibi, Mental Health Research and Treatment Center, Faculty of Psychology, Ruhr-Universität Bochum, Massenbergsstraße 9-13, 44787 Bochum, Germany. (E-mail: akhtar.bibi@rub.de).

The study was funded by the Alexander von Humboldt Professorship awarded to Jürgen Margraf by the Alexander von Humboldt-Foundation. Akhtar Bibi was supported by HEC/DAAD (911591439). The funders had no role in the design of the study, collection, analysis and interpretation of data, or in the writing of the manuscript.

AB and JM conceived the study and contributed to study design. AB collected the data and wrote the first manuscript draft. ML helped in revising the article. XCH helped in analysis of the data. All authors contributed to interpretation of data and manuscript revisions, and all authors read and approved the final manuscript.

measurement invariance of the DASS-21 in Asian countries Malaysia, Singapore, Indonesia, Taiwan, Thailand and Sri Lanka showed that the stress scale does not fit the Asian cultural context (Oei, Sawang, Goh, & Mukhtar, 2013).

In Asian cultures, patients experience depression predominantly in terms of somatic symptoms (Kirmayer, Robbins, Dworkind, & Yaffe, 1993). Research from numerous countries, such as China, India, Mexico and Africa, reported that patients in Eastern countries are more prone to “somatize” depression, whereas patients in developed countries “psychologize” depressive symptoms (Mumford, Devereux, Maddy, & Johnston, 1991). Similarly, Bibi, Masroor, and Iqbal (2013) reported higher somatisation of depression and conversion disorder symptoms in the Pakistani population. These results indicate that culture can affect expression, manifestation, symptomatology and interpretation of depression, anxiety and stress (Kirmayer et al., 1993).

Therefore, psychological measurements should be equivalent across cultures to avoid culture-blind errors. In order to assess whether an instrument is applicable across cultures, three constrained levels of invariance are generally tested (Byrne, 2008). First, configural invariance is tested to check whether the factor structure of a measure is the same across cultures without placing additional equality restraints on the model parameters. Second, to test metric invariance, all factor loadings are restrained to be equal across cultures to investigate whether items contribute to the latent construct equally across groups. Third, scalar invariance implies that not only factor loadings but also intercepts of items are the same across groups. If factor structure, factor loadings and intercepts are equal across groups, latent means can be meaningfully compared across the groups. In models with ordered categorical variables, thresholds rather than intercepts are constrained.

Scholten, Velten, Bieda, Zhang, and Margraf (2017) investigated measurement invariance for the DASS-21 across United Kingdom, United States, Russia and Poland and established strong measurement invariance. Studies that revised the DASS-21 for Asian countries mainly focused on psychometric properties, translation, and validation or factor structure of the DASS-21 (Aslam & Kamal, 2017; Oei et al., 2013). However, social, political and cultural factors vary between Western and Eastern countries and should be taken into account. Therefore, testing measurement invariance would show whether the DASS-21 is cross-culturally applicable.

The present study aimed to examine the psychometric properties and measurement invariance of the DASS-21 in Pakistan and Germany, representative for an Eastern and a Western country, respectively. Pakistan fosters collectivistic values, which emphasises group goals compared to individual concerns. Germany, on the other hand, fosters

TABLE 1
Demographic variables (Pakistan and Germany)

	Pakistan	Germany
<i>N</i>	1841	1323
Gender, <i>N</i> (%)		
Female	1097(59.6)	908 (68.6)
Male	744 (40.4)	314 (31.40)
Age		
<i>M</i> (<i>SD</i>)	22.84(3.01)	23.96 (4.64)
Range	18–44	16–64
School years, <i>N</i> (%)		
1 year	701(39.83)	1148(86.8)
2 year	434(24.66)	33(2.5)
3 year	318(18.07)	24(1.8)
4 year	307(17.44)	51(3.9)
5 year	—	33(2.5)
6 year	—	28(2.1)
Above 6	—	6(0.5)
Missing	4.4%	0

Notes: *M* = Mean; *SD* = Standard deviation.

a typical individualistic cultural context, focusing on individual’s rights and interests with great personal freedom of choice. Therefore, measurement invariance testing is an essential precondition for adopting the DASS-21 into the Pakistani cultural framework. If it could be shown that the DASS-21 met the precondition of scalar measurement invariance, latent mean differences could be evaluated for depression and anxiety symptoms. The current study contributes to existing literature on the cross-cultural applicability of the DASS-21 and improves the chances for cross-cultural studies in both countries.

METHODOLOGY

Participants and procedures

Data of the current study were collected within the Bochum Optimism and Mental Health (BOOM) project, which is large-scale longitudinal cross-cultural study, examining positive and negative factors of psychological health and psychological illness. Data were obtained from samples consisting of 1841 Pakistani and 1323 German participants. German participants were recruited through an online survey, while the Pakistani sample was recruited through online survey and paper–pencil method. Participants provided informed consent before participating in the survey. They were assured anonymity and confidentiality of data. Table 1 shows the demographic characteristics of the participants of both countries. The survey was conducted in Germany between 2015 and 2017 and in Pakistan between 2016 and 2017. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki

Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all participants.

Measures

Depression, Anxiety and Stress Scales

DASS-21: English original version (Lovibond & Lovibond, 1995); German version Henry & Crawford, 2005: Depression, anxiety and stress symptoms of both samples were evaluated by using the freely available English version of the DASS-21 (seven items for each) for the Pakistani sample and the German version of the DASS-21 for the German sample. Participants responded on a 4-point Likert-scale from 0 (*did not apply to me at all*) to 3 (*applied to me very much or most of the time*). In our study, total Cronbach's alpha was 0.89 for the Pakistani sample and 0.93 for the German sample.

Positive Mental Health Scale

PMH-scale: German original version (Lukat, Margraf, Lutz, van der Veld, & Becker, 2016): The nine-item PMH-Scale was used to measure the positive features of psychological well-being and health. Participants rated the statements such as "All in all, I am satisfied with my life." on a 4-point Likert-scale ranging from 0 (*I disagree*) to 3 (*I agree*). The English version of the PMH-scale was translated from the original language version by using the translation-backtranslation-modification procedure (Berry, 1989). In our sample, Cronbach's alpha reliability of the PMH-scale was 0.87 for the Pakistani sample and 0.91 for the German sample.

Resilience

RS-11: German original version (Schumacher, Lepert, Gunzelmann, Strauß, & Brähler, 2005); English version (Brailovskaia, Schönfeld, Kochetkov, & Margraf, 2017): The 11-item Resilience Scale was applied to measure students' resilience. They responded to 11 items on a 7-point Likert-scale ranging from 1 (*disagree*) to 7 (*agree*) (e.g. "I feel that I can handle many things at a time."). Cronbach's alpha was 0.82 for the Pakistani sample and 0.87 for the German sample.

Satisfaction with Life Scale

SWLS: English original version (Diener, Emmons, Larsen, & Griffin, 1985); German version (Glaesmer, Grande, Braehler, & Roth, 2011): The five-item SWLS was used to measure the satisfaction with life. Participants rated the questions on 7-point Likert-scale ranging from 1 (*Strongly agree*) to 7 (*strongly disagree*). Cronbach's

alpha was 0.72 for the Pakistani sample and 0.80 for the German sample.

Perceived Social Support Questionnaire

F-SozU K-14: German original version (Fydrich, Sommer, Tydecks, & Brähler, 2009); English version: (Brailovskaia et al., 2017): The 14-item questionnaire (F-SozU K-14) was used to evaluate participants' perceived social support. They responded questions on 5-point Likert-scale ranging from 1 (*not true*) to 5 (*true*). Cronbach's alpha was 0.86 for the Pakistani sample and 0.94 for the German sample.

DATA ANALYSIS

Data was checked for missing values, and cases with incomplete data were deleted. Standard descriptive statistics were calculated for the demographic variables. Cronbach's alphas were calculated as an indicator of internal consistency. For individual items we calculated means, standard deviations, skewness, kurtosis, "alpha if item deleted" and corrected item-total correlations. A three-factor model was used to test structure validity and thereafter measurement invariance of the three-factor model of the DASS as a whole. Furthermore, the three one-factor models for each subscale were tested.

In order to investigate construct validity, correlations between depression, anxiety and stress scales and positive mental health scales (PMH, F-SozU K-6, SWLS and RS-11) were calculated. Single-group CFAs were used to investigate the proposed factor structure of each subscale of the DASS-21.

TESTING MEASUREMENT INVARIANCE

We followed Jöreskog's (1971) multigroup confirmatory factor analysis (CFA) approach to test the measurement invariance of the DASS-21 between Pakistan and Germany using R-software (Lavaan package). Muthén and Asparouhov (2002) further specified the approach for ordered categorical variables. We considered the ordinal nature of our data and used weighted least squares mean and variance adjusted estimator (WLSMV). WLSMV is suggested for measuring thresholds when there are fewer than five response categories. Delta parameterisation was also applied, as this is suggested for ordered-categorical data.

Measurement invariance based on multigroup CFA was conducted between both countries. Since the Chi-square test is sensitive to larger sample sizes and can lead to high rates of rejection, we used a combination of different goodness-of-fit indices to assess our model fit: root comparative fit index (CFI; Bentler, 1990), mean

squared error of approximation (RMSEA; Steiger, 1980) and weight root mean square residual (WRMR). To evaluate model fit, we followed the recommendation of Hu and Bentler (1999). They suggest the cutoff criteria of CFI > .95, RMSEA < .08 for a good model fit. A change of $\Delta\text{CFI} \geq .01$ indicates a significant decrease in model fit and hence non-invariance.

Measurement invariance testing was primarily conducted in four steps. First, single-group CFAs were investigated for each country. Second, configural invariance with multi-group CFAs without equality constraints was examined. Third, the model was tested for weak (metric) measurement invariance with multi-group CFAs restricting all factor loadings to be equal across both samples. Fourth, the model was tested for strong (scalar) measurement invariance by restricting all thresholds to be equal across both samples.

In every step, each model was accepted only when the values of CFI, RMSEA and WRMR were within the cutoff criteria. When full invariance was not found, partial weak and partial strong measurement invariance was tested by setting free specific factor loadings or thresholds (Byrne, Shavelson, & Muthén, 1989). Following Byrne et al. (1989), we continued partial measurement invariance until at least two items were invariant.

RESULTS

Sample characteristics

Demographic characteristics of both samples are presented in Table 1.

Item characteristics

Item characteristics of the DASS-21 in English and German, such as means, standard deviations, skewness, corrected item total scores, factor loadings, alpha-if-item-deleted and internal consistency are demonstrated in Table 2. Internal consistency in both populations was good. Corrected item total score correlation ranged in both populations from 0.21 (Item 2 in Pakistan) to 0.78 (Item 10 in Germany); values below 0.8 show that all items measure the same construct without multi-collinearity. All scales showed normal univariate distribution in the Pakistani sample, but the German sample showed slightly left-tailed distributions.

Three factor model

Results of the single group CFAs showed that the three-factor model of the DASS-21 was appropriate in both countries (Table 3). The values of CFI, RMSEA and WRMR were within the desired range, suggesting a good model fit (Hu & Bentler, 1999). Further model

specification was not needed, as the baseline model fitted well across both countries. Item loadings are shown in Table 3.

Construct validity

All correlations between the DASS-21 (DASS-Depression, DASS-anxiety and DASS-stress) and self-reported inventories of positive mental health (PMH, F-SozU K-14, SWLS and RS-11) were statistically significant and in the expected direction, indicating that lower symptoms of depression, anxiety and stress are associated with higher positive mental health, social support, satisfaction with life and resilience (Table 4).

Multi-Group CFA of the DASS-21

Depression, Anxiety and Stress Scale (DASS-21)

Table 5 shows the model testing results for measurement invariance for the DASS-21. The configural model showed an acceptable global fit. Next, testing of weak measurement invariance indicated poor model fit, as ΔCFI exceeded .01. Thus, partial weak measurement invariance was established by setting the factor loadings of Item 9, Item 14, Item 6, Item 8 and Item 12. However, testing of strong measurement invariance indicated poor model fit, thus strong measurement invariance was not established.

Depression

Table 5 shows the model testing results for measurement invariance for the DASS-21 depression subscales. The configural model showed an acceptable global fit, suggesting that the configural invariance assumption was supported. When equality restraints of factor loadings were applied, ΔCFI did not exceed .01. Therefore, strong measurement invariance was examined, which showed poor model fit, with ΔCFI exceeding .01. When the threshold of Item 13/t1, t2 were allowed to be free, the global model fit improved, and ΔCFI value was less than .01. Thus, partial strong measurement invariance was established.

Anxiety

Table 5 shows the model testing results for measurement invariance for the anxiety subscale of the DASS-21. The global fit of the configural model was acceptable. Weak measurement invariance was established, with ΔCFI being less than .01. The model was then tested for strong measurement invariance by constraining the thresholds to be equal across both countries, ΔCFI exceeding .01. Thus, full strong measurement invariance

TABLE 2

Mean, standard deviation (*SDs*), skewness, kurtosis, correlations, corrected item-total score correlation (r_{it}) and factor loading (loading) of the DASS-21 (depression, anxiety and stress scales) in Pakistani and German student sample

<i>Items</i>	<i>Mean</i>	<i>SD</i>	<i>Skewness</i>	<i>Kurtosis</i>	r_{it}	<i>Loading</i>	<i>Alpha-if-item-deleted</i>
Pakistan—Depression Subscale							
Item 3	1.13	0.93	0.39	−0.73	.40	.49	.77
Item 5	1.16	0.86	0.32	−0.59	.39	.48	.77
Item 10	1.04	0.94	0.48	−0.77	.54	.68	.74
Item 13	1.16	0.94	0.39	−0.77	.52	.64	.75
Item 16	1.12	0.95	0.43	−0.77	.54	.69	.74
Item 17	1.08	0.98	0.47	−0.85	.58	.74	.73
Item 21	1.06	1.05	0.56	−0.94	.54	.70	.74
Cronbach's alpha	.78						
Germany—Depression Subscale							
Item 3	0.64	0.84	1.15	0.45	.74	.86	.88
Item 5	1.12	0.91	0.40	−0.68	.53	.62	.90
Item 10	0.67	0.94	1.23	0.40	.78	.90	.88
Item 13	1.09	0.99	0.56	−0.72	.74	.84	.88
Item 16	0.64	0.88	1.21	0.47	.73	.84	.88
Item 17	0.72	1.00	1.15	0.05	.71	.82	.89
Item 21	0.50	0.89	1.68	1.65	.72	.86	.88
Cronbach's alpha	.90						
Pakistan—Anxiety Subscale							
Item 2	1.41	0.95	0.11	−0.91	.21	.27	.74
Item 4	0.92	0.96	0.68	−0.63	.39	.52	.70
Item 7	1.11	0.96	0.40	−0.88	.45	.58	.68
Item 9	1.35	0.98	0.23	−0.95	.49	.64	.67
Item 15	1.14	0.95	0.43	−0.75	.51	.69	.67
Item 19	1.32	1.02	0.19	−1.08	.48	.62	.68
Item 20	1.15	1.00	0.39	−0.95	.49	.66	.67
Cronbach's alpha	.72						
Germany—Anxiety Subscale							
Item 2	0.71	0.86	1.05	0.26	.43	.53	.80
Item 4	0.40	0.76	1.89	2.79	.53	.70	.79
Item 7	0.55	0.87	1.47	1.13	.52	.66	.79
Item 9	1.03	1.03	0.60	−0.87	.50	.64	.79
Item 15	0.51	0.85	1.62	−1.63	.64	.82	.77
Item 19	0.71	0.95	1.11	0.06	.58	.72	.78
Item 20	0.63	0.90	1.27	0.54	.62	.78	.77
Cronbach's alpha	.81						
Pakistan—Stress Subscale							
Item 1	1.13	7.91	0.44	−1.02	.28	.36	.75
Item 6	1.17	0.92	0.36	−0.72	.44	.55	.72
Item 8	1.21	0.93	0.33	−0.74	.50	.64	.71
Item 11	1.17	0.91	0.33	−0.73	.51	.66	.70
Item 12	1.32	0.95	0.25	−0.85	.51	.66	.70
Item 14	1.28	0.94	0.22	−0.85	.50	.64	.71
Item 18	1.28	0.97	0.21	−0.96	.47	.59	.71
Cronbach's alpha	.75						
Germany—Stress Subscale							
Item 1	1.02	0.89	0.61	−0.32	.67	.78	.84
Item 6	1.06	0.95	0.49	−0.76	.58	.68	.85
Item 8	1.25	1.01	0.30	−1.00	.66	.76	.84
Item 11	1.19	0.98	0.36	−0.90	.71	.83	.83
Item 12	1.20	0.99	0.39	−0.90	.70	.82	.83
Item 14	0.56	0.77	1.30	1.09	.44	.54	.87
Item 18	1.13	1.03	0.43	−1.02	.66	.76	.84
Cronbach's alpha	.86						

TABLE 3
Single group CFA results and cross-cultural measurement invariance tests results of the DASS-21 scale

Models	Sample	N	Chi ² /df	RMSEA	90% C.I.	CFI	WRMR	ΔCFI
Three-factor model—DASS-21								
Single group CFA—DASS-21								
	Pakistan Student	1841	1326.898/189	0.057	0.054–0.054	0.945	1.79	
	Germany	1323	2415.069/189	0.094	0.091–0.098	0.922	2.564	
Multiple group CFA—DASS-21								
	Configural		3802.761/378	0.076	0.074–0.078	0.931	3.127	
	Weak		4622.935/398	0.082	0.080–0.084	0.915	4.005	0.016
	Partial weak((λ ₉ , λ ₁₄ , λ ₆ , λ ₈ , λ ₁₂ free)		4220.808/47	0.078	0.076–.080	0.923	3.794	0.008
	Strong Student		7204.004/438	0.099	0.097–0.101	0.863	4.523	0.06

Notes: German sample $N = 1323$, Pakistan sample $N = 1841$; RMSEA = root mean square error of analysis; 90% C.I. = 90% confidence interval of RMSEA; CFI = Comparative fit index; SRMR = standardised root mean square residual.

TABLE 4
Means, standard deviation (SDs), and correlation of depression, anxiety and stress other psychomeasurements

	Pakistan				Germany					
	depression	Anxiety	Stress	Mean	SD	depression	Anxiety	Stress	Mean	SD
N1				1840					1323	
Depression	1			7.75	4.36	1			5.39	5.10
Anxiety	.720**	1		8.40	4.17	.609**	1		5.54	4.25
Stress	.730**	.730**	1	8.56	4.04	.686**	.706**	1	7.41	4.91
PMH-9	-.179**	-.076**	-.061**	17.42	6.32	-.713	-.542	-.614	17.61	5.89
SWLS	-.177**	-.074**	-.100**	20.48	6.36	-.589**	-.432**	-.432**	24.29	6.18
Social Support	-.155**	-.045	-.033	49.97	11.04	-.500**	-.396**	-.332**	58.44	10.52
Resilience	-.143**	-.071**	-.053**	55.66	12.94	-.562**	-.412**	-.413**	57.72	9.80

* $p < .05$. ** $p < .01$. *** $p < .001$.

was not established. After allowing the thresholds of Item 9/t1 and Item 2/t1 to be freely estimated, partial strong measurement invariance was established.

Stress

Table 5 displays model testing results for measurement invariance for the stress subscale of the DASS-21. In testing measurement invariance, configural invariance assumption for the stress subscale was supported; the global fit of this model remained good. Next, factor loadings were constrained to be equal, which resulted in a ΔCFI larger than the proposed cutoff. Expected parameter change statistics suggested that factor loadings of Item 14, Item 6, Item 8 and Item 12 were non-invariant across both samples. After releasing the factor loadings for these items, partial weak measurement was established. In testing for strong measurement invariance, the decrease in the ΔCFI was larger than .01. After releasing the equality constraints of threshold for Item 14/t1, t2, partial strong measurement was also established.

Latent means comparisons

Latent mean differences were compared for the depression and anxiety subscales. Pakistan was taken as a

reference group to assess the significance of latent mean differences between Pakistani and German university students. The latent means of depression ($z = -12.68$, $p < .001$, $d = .678$) and anxiety ($z = -12.11$, $p < .001$, $d = .372$) were higher in the Pakistani sample than in the German sample. A latent mean comparison was not calculated for the stress scale as the precondition for comparing latent means was not met.

DISCUSSION

The current study investigated psychometric properties and measurement invariance of the DASS-21 across Pakistan and Germany. Generally, the DASS-21 showed good psychometric properties in both samples. For the sum score of the DASS-21, the two countries are not directly comparable. But for the individual depression and anxiety subscales, it is possible to compare the latent means. Results of single group CFAs showed the same factor structure for all three subscales. Multigroup CFAs for depression and anxiety supported the postulation of same structure and equivalent loadings, however, did not support the assumption of equivalent factor loadings and threshold for the stress scales. Latent means were compared for depression and anxiety subscales of the

TABLE 5
Single group CFA results and cross-cultural measurement invariance tests results of the depression, anxiety and stress

Models	Chi ² /df	RMSEA	90% C.I.	CFI	WRMR	ΔCFI
One-factor model—Depression						
Single group CFA—Depression						
Pakistan	83.366/14	0.052	0.041–0.063	0.987	1.059	
Germany	91.610/14	0.065	0.053–0.078	0.995	0.936	
Multiple group CFA—Depression						
Configural	171.592/28	0.057	0.049–0.065	0.993	1.413	
Weak	392.215/34	0.082	0.074–0.089	0.983	2.403	0.01
Strong	883.607/47	0.106	0.100–0.112	0.960	3.41	0.023
Partial strong ($\tau_{13/t1}$, $\tau_{13/t2}$ free)	595.541/45	0.088	0.082–0.094	0.974	2.805	0.009
One-factor model—Anxiety						
Single group CFA—Anxiety						
Pakistan	75.957/14	0.049	0.039–0.060	0.983	1.058	
Germany	133.599/14	0.080	0.068–0.093	0.971	1.333	
Multiple group CFA—Anxiety						
Configural	206.282/28	0.063	0.055–0.072	0.977	1.701	
Weak	279.001/34	0.068	0.060–0.075	0.969	2.135	0.008
Strong	564.834/47	0.083	0.077–0.090	0.934	3.037	0.035
Partial strong ($\tau_{9/t1}$ free)	436.107/46	0.073	0.067–0.080	0.950	2.677	0.019
Partial strong ($\tau_{9/t1}$, $\tau_{2/t1}$ free)	326.161/45	0.063	0.057–0.069	0.964	2.286	0.005
One-factor model—Stress						
Single group CFA—Stress						
Pakistan	56.277/14	0.041	0.030–0.052	0.990	0.875	
Germany	158.057/14	0.088	0.076–0.101	0.983	1.350	
Multiple group CFA—Stress						
Configural	205.302/28	0.063	0.055–0.072	0.986	1.609	
Weak	653.934/34	0.107	0.100–0.115	0.952	3.191	0.034
Partial weak (λ_{14} free)	408.181/33	0.085	0.078–0.092	0.971	2.449	0.015
Partial weak (λ_{14} , λ_6 free)	391.784/32	0.084	0.077–0.092	0.972	2.359	0.014
Partial weak (λ_{14} , λ_6 , λ_8 free)	358.706/31	0.082	0.074–0.090	0.975	2.234	0.011
Partial weak (λ_{14} , λ_6 , λ_8 , λ_{12} free)	342.126/30	0.081	0.074–0.089	0.976	2.156	0.01
Strong	1671.107/47	0.148	0.142–0.154	0.875	4.930	0.101
Partial strong ($\tau_{14/t1}$, $\tau_{14/t2}$, free)	415.139/41	0.076	0.070–0.083	0.971	2.353	0.005

Notes: German sample $N = 1323$, Pakistan sample $N = 1841$; RMSEA = root mean square error of analysis; 90% C.I. = 90% confidence interval of RMSEA; CFI = Comparative fit index; SRMR = standardised root mean square residual. $\tau_{14/t1}$, $\tau_{14/t2}$ = intercept of item 14.

DASS-21, indicating higher levels of depression and anxiety symptoms in Pakistani participants as compared to Germans.

Psychometric properties of the DASS-21 indicated that it is a reliable instrument that can be applied for screening of depression and anxiety symptoms in both countries. Moreover, each sample indicated negative associations between depression, anxiety and stress symptoms and positive mental health inventories, such as positive mental health, social support and satisfaction with life and resilience. However, correlations between the DASS-21 and positive scales were considerably stronger in the German sample (ranging from $-.33$ to $-.71$) than in the Pakistani sample where correlations were close to zero (ranging from $-.03$ to $-.18$) and not significant. Positive mental health and negative mental health are two dimensions of mental health. They were more strongly associated in the German sample than in the Pakistani sample. It could be that even when Pakistani people report being anxious and depressed, at the same time they can still be generally

satisfied and happy with their lives. Future studies should further investigate this association.

Full configural and weak measurement invariance was established for the depression and anxiety subscales. However, investigation of strong measurement invariance showed that the thresholds of Item 13/t1, 13/t2 (*I felt downhearted and blue*), Items 9/t1 (*I was worried about situations in which I might panic and make a fool of myself*) and Item 2/t1 (*I was aware of dryness of my mouth*) were not invariant. At the trait level, Pakistani students agreed to Item 13 less frequently than German participants. It could be that since depression was lower in Germany, German participants answered with “not applied” more frequently to Item 13. Moreover, Pakistani students had a higher frequency of responding with “not applied” to Item 9 and a lower frequency of responding with “not applied” to Item 2 than German students. It could be that German students are less tolerant of situations where they are not in control. Furthermore, as reported in previous studies, Pakistani people may

have more somatic symptoms when experiencing internal problems (Bibi et al., 2013), and thus they are more likely to report symptoms such as “dryness of mouth” (Item 2).

Testing measurement invariance for the stress scales showed full configural measurement invariance. Partial weak measurement was established by freely estimating the factor loadings for Item 14 (*I was intolerant of anything that kept me from getting on with what I was doing*), Item 6 (*I tended to over-react to situations*), Item 8 (*I found it difficult to relax*) and Item 12 (*I felt that I was using a lot of nervous energy*), suggesting that these items contribute differently to the latent construct. Partial strong measurement invariance was achieved by freely estimating the threshold of Items 14/t1, 14/t2 (*I was intolerant of anything that kept me from getting on with what I was doing*). The threshold of Item 14/t1, 14/t2 was higher in the German sample than in the Pakistani sample, which indicates that Germans are more likely to endorse these items than Pakistani participants. Thus, we recommend that future studies applying cross-cultural comparisons on this subscale adopt latent mean comparisons instead of relying on observed scores.

There are several possible explanations for the different factor loadings between our Pakistani and German samples.

First, the constructs might be understood differently. For example, the term “nervous energy” (Item 12) might have different connotations in the two samples. Second, specific items may make a greater contribution to the total score in one culture, such as Items 6, 8, 12 of the stress scale. Therefore, the factor loadings of these items were higher in the German sample than the Pakistani sample. Third, participants responded differently to extreme items (for instance, with respect to Item 14, people in Pakistan feel more annoyed than stressed in an unusual situation; Item 13, participants would understand concepts such as “down hearted and blue” differently). Fourth, differences in thresholds could emerge because of social norms or a social desirability effect (e.g. Item 6 and Item 8, participants might not want to admit their reactions to different situations or might have difficulty relaxing). Some participants might have tended to respond to some items more strongly in spite of the same factor mean, or some participants may have had different reference points when they responded about themselves (Chen, 2008). The measurement invariance results of the stress subscale indicate that the stress scale is not suitable for cross-cultural comparisons between the two samples. With respect to the stress scale, our results were consistent with the findings of Oei et al. (2013) who revised the stress scale after showing that the scale was not directly applicable to the Asian population or may be interpreted slightly differently.

The findings of the latent means comparisons clearly indicated higher levels of depression and anxiety symptoms among Pakistani students. Effect sizes were medium

to large for the symptoms of depression and anxiety, respectively. The higher levels of depression and anxiety symptoms were consistent with the higher prevalence rate of depression and anxiety among university students in Pakistan (Bibi, Blackwell, & Margraf, 2019). There are various possible explanations for the higher prevalence of these symptoms. One reason could be the poor economic situation and unstable political conditions in Pakistan compared to the developed economy of Germany. Symptoms of depression and anxiety are also associated with a declining economy. Pakistan is a developing country where 40% of the population lives in poverty. Overall, academic challenges, unemployment, high family expectations, political oppression and poor living conditions contribute to higher levels of depression, anxiety and stress symptoms among Pakistani students. The possible explanations for the mean differences in symptoms across both countries should be considered in a broader context.

There are some limitations to this investigation. First, Pakistan and Germany differ not only in their level of economic and social development (Fragile States Index, 2018), but also in terms of religion. Most Pakistani live strictly according to Islamic disciplines, while Germans have much freedom and variety with regard to their religious practices. Second, data were collected using convenient sampling techniques, which may not truly represent the general population in both countries. In spite of the limitation of true representativeness, the sample was large enough to generalise our results. Third, the survey method was not the same across countries. The German sample was collected through an online survey, whereas the Pakistani sample was collected through paper–pencil as well as an online survey, which can cause measurement bias and thus influence the accuracy of measurement invariance. Third, measurement invariance is a statistical method to test the equivalency of instruments across cultures; however, it is not easy to differentiate genuine cultural dissimilarities from biases and errors in measurement including data assembly and social norms (Byrne, 2008). Hence, measurement invariance testing is a first step toward culturally invariant measurement instruments. Qualitative methods such as focused group discussions and cognitive interviewing could be conducted to investigate possible differences in constructs and items in detail.

In nutshell, the DASS-21 generally indicated good psychometric properties within both samples. It was shown to be a reliable, valid and economical scale to assess symptoms of depression and anxiety in both the Pakistani and German cultural context. Measurement invariance testing indicated partial weak and partial strong measurement invariance, permitting latent mean comparisons between participants from Pakistan and Germany. However, direct comparisons between

samples from the two countries should be conducted with caution.

Manuscript received May 2019
Revised manuscript accepted March 2020

REFERENCES

- Aslam, N., & Kamal, A. (2017). Translation, validation and effectiveness of depression, anxiety and stress (Dass-21) in assessing the psychological distress among flood affected individuals. *Journal of Pakistan Psychiatric Society, 14*, 16–20.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Berry, J. W. (1989). *Introduction to methodology*. In H. Triandis & J. W. Berry (Eds.), *Handbook of cross-cultural psychology*. (Vol. 2, p. 1e28). Boston, MA: Allyn & Bacon.
- Bibi, A., Blackwell, S. E., & Margraf, J. (2019). Mental health, suicidal ideation, and experience of bullying among university students in Pakistan. *Journal of Health Psychology, 1*–12.
- Bibi, A., Masroor, U., & Iqbal, N. (2013). Dysfunctional attitudes and demographic correlates of patients with conversion disorder; an exploratory study. *Journal of Pakistan Psychiatric Society, 10*(1), 25–29.
- Bibi, A., Lin, M., & Margraf, J. (2020). Salutogenic constructs across Pakistan and Germany: A cross sectional study. *International Journal of Clinical and Health Psychology, 20*(1), 1–9.
- Brailovskaia, J., Schönfeld, P., Kochetkov, Y., & Margraf, J. (2017). What does migration mean to us? USA and Russia: Relationship between migration, resilience, social support, happiness, life satisfaction, depression, anxiety and stress. *Current Psychology, 38*, 421–431.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema, 20*, 872–882.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. <https://doi.org/10.1037/0033-2909.105.3.456>
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*, 1005–1018. <https://doi.org/10.1037/a0013193>
- Cheung, F. M. (2012). Mainstreaming culture in psychology. *American Psychologist, 67*, 721–730. <https://doi.org/10.1037/a0029876>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
- Fragile States Index. 2018. Retrieved from <http://fundforpeace.org/fsi/data/>
- Fydrich, T., Sommer, G., Tydecks, S., & Brähler, E. (2009). Fragebogen zur sozialen unterstützung (F-SozU): Ntraormierung der Kurzform (K-14). *Zeitschrift für Medizinische Psychologie, 18*(1), 43–48.
- Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the satisfaction with life scale (SWLS): Psychometric properties, validity, and population-based norms. *European Journal of Psychological Assessment, 27*, 127–132. <https://doi.org/10.1027/1015-5759/a000058>
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the depression anxiety stress scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*, 227–239. <https://doi.org/10.1348/014466505X29657>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. <https://doi.org/10.1080/10705519909540118>
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*(4), 409–426. <https://doi.org/10.1007/BF02291366>
- Kirmayer, L. J., Robbins, J. M., Dworkind, M., & Yaffe, M. J. (1993). Somatization and the recognition of depression and anxiety in primary care. *American Journal of Psychiatry, 150*, 734–741.
- Lin, M., Hirschfeld, G., & Margraf, J. (2018). Brief form of the perceived social support questionnaire (F-SozU K-6): Validation, norms, and cross-cultural measurement invariance in the USA, Germany, Russia, and China. *Psychological Assessment, 31*(5), 609–621. <https://doi.org/10.1037/pas0000686>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behaviour Research and Therapy, 33*, 335–343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Lukat, J., Margraf, J., Lutz, R., van der Veld, W. M., & Becker, E. S. (2016). Psychometric properties of the positive mental health scale (PMH-scale). *BMC Psychology, 4*(1), 8. <https://doi.org/10.1186/s40359-016-0111-x>
- Marum, G., Clench-Aas, J., Nes, R. B., & Raanaas, R. K. (2014). The relationship between negative life events, psychological distress and life satisfaction: A population-based study. *Quality of Life Research, 23*(2), 601–611.
- Mumford, D. B., Devereux, T. A., Maddy, P. J., & Johnston, J. V. (1991). Factors leading to the reporting of 'functional' somatic symptoms by general practice attenders. *The British Journal of General Practice, 41*(352), 454–458 Retrieved from <https://bjgp.org/content/41/352/454.short>
- Muthén, B., & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. *Mplus Web Notes, 4*(5), 1–22.
- Oei, T. P. S., Sawang, S., Goh, Y. W., & Mukhtar, F. (2013). Using the depression anxiety stress scale 21 (DASS-21) across cultures. *International Journal of Psychology, 48*, 1018–1029. <https://doi.org/10.1080/00207594.2012.755535>
- Scholten, S., Velten, J., Bieda, A., Zhang, X. C., & Margraf, J. (2017). Testing measurement invariance of the depression, anxiety, and stress scales (DASS-21) across four countries. *Psychological Assessment, 29*(11), 1376–1390.
- Schumacher, J., Leppert, K., Gunzelmann, T., Strauss, B., & Brähler, E. (2005). Die Resilienzskala-Ein Fragebogen zur Erfassung der psychischen Widerstandsfähigkeit als Persönlichkeitsmerkmal [The Resilience Scale-A questionnaire for the

assessment of psychosocial stress-resistance as a personality factor]. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie*, 53, 16–39.

Steiger, J. H.. (1980). Statistically based tests for the number of common factors. In *The annual meeting of the Psychometric Society*. Iowa City, IA.

Teismann, T., Forkmann, T., Brailovskaia, J., Siegmann, P., Glaesmer, H., & Margraf, J. (2018). Positive mental health moderates the association between depression and suicide ideation: A longitudinal study. *International Journal of Clinical and Health Psychology*, 18, 1–7. <https://doi.org/10.1016/j.ijchp.2017.08.001>