

The Depression Anxiety and Stress Scale–21: Bifactor Statistical Indices in Support of the Total and Depression scores*

Ljiljana Mihić, Marija Volarov, Milan Oljača, and Zdenka Novović

Department of Psychology, Faculty of Philosophy, University of Novi Sad, Serbia

This study explored several, latent factor models of the Depression Anxiety and Stress Scale–21 (DASS–21) using both a sample of clinically depressed patients and a Facebook sample from Serbia. The DASS–21, the Beck Depression Inventory–II, and the State Trait Anxiety Inventory–Trait were administered to a sample of depressed individuals ($N = 296$; $M_{\text{age}} = 52.21$, $SD_{\text{age}} = 11.56$). A Facebook sample ($N = 376$; $M_{\text{age}} = 29.12$, $SD = 8.96$) completed the DASS–21 only. A bifactor model with one general distress (G) and two specific factors (Depression and Anxiety) were supported. The three factors had high omega coefficients, whereas omega hierarchical for Depression and Anxiety were low. Based on all evidence from our study, external validation, factor determinacy, and replicability, we concluded that the Serbian version of the DASS–21 assesses reliably general distress and anhedonia in both people with the clinical level of severity of distress and in general population. The Anxiety subscale can be safely used in clinical settings when one wishes to assess individuals presenting with a mixture of distress and anxious arousal. However, until further validation support is provided, we do not recommend use of this scale when the task is to estimate precisely anxious arousal only. The Stress subscale assesses general distress only. Low omega hierarchical coefficients of the Depression and Anxiety subscales could be addressed by re-selecting their items from the initial item pool, based on which the DASS–21 was created, using those with substantial loadings on both the G and their respective group factors.

Keywords: bifactor model, clinical sample, DASS–21, depression, validity

Highlights:

- The total DASS–21 score is a valid and reliable measure of general distress.
- The Depression subscale score is a valid and reliable measure of anhedonia.

Corresponding author: lmihic@ff.uns.ac.rs

* Please cite as: Mihić, Lj., Volarov, M., Oljača, M., Novović, Z. (2021). The Depression Anxiety and Stress Scale–21: Bifactor Statistical Indices in Support of the Total and Depression scores. *Psihologija*, 54(2), 155–172. doi: <https://doi.org/10.2298/PSI191227025M>

- The Anxiety subscale score is a valid and reliable measure of both general distress and physiological arousal.
- The Stress subscale seems to measure general distress only.

The Depression, Anxiety and Stress Scale-21 (DASS-21; Lovibond & Lovibond, 1995a) is a worldwide known measure of the core symptoms of “depression, physical arousal, and psychological tension and agitation” (Antony et al., 1998, p. 176). The scale has been translated into many different languages such as German, Russian, Chinese (Brailovskaia et al., 2017), and also Serbian (Jovanović et al., 2014). The Depression subscale includes items that refer to sad mood, a lack of motivation and perseverance in reaching life goals, and low self-esteem. The Anxiety subscale assesses physiological arousal and responses in fear-evoking situations, while the Stress subscale contains the items reflecting “a state of persistent arousal and tension with a low threshold for becoming upset or frustrated” (Lovibond & Lovibond, 1995b, p. 342). In the clinical samples, the subscales have moderately high correlations among each other, whereby the Stress scale has higher correlations with both the Depression and the Anxiety scales than the Depression and the Anxiety scales with one another (e.g., Antony et al., 1998; Clara et al., 2001). Lovibond and Lovibond (1995b) have pointed out that the moderately high correlations between the three subscales might be due to a common vulnerability factor which influences all three states. Their explanation was influenced by the Tripartite model of anxiety and depression (Clark & Watson, 1991). This model posits that negative affectivity (NA) is a dimension common to both depression and anxiety, but each syndrome has its specific characteristics: depression is characterized by low positive affectivity, whereas anxiety is characterized by physical hyperarousal (Clark & Watson, 1991). Hence, responses to the Depression and the Anxiety subscales seem to reflect two sources of variance: one general dimension (NA) and specific dimensions of low positive affectivity and high physiological arousal, respectively.

However, the status of the Stress scale has been more controversial. On the one hand, it was suggested that it was defined by the items tapping an affective state different from those captured by the Depression and Anxiety subscales but that it shares with them a common vulnerability such as, for example, negative affectivity or common environmental influences (Lovibond & Lovibond, 1995b). Others have also argued for its specificity, in addition to its common element of NA, in both English and non-English speaking countries (e.g., Bottesi et al., 2015; Henry & Crawford, 2005; Jovanović et al., 2014). However, the exact nature of the Stress scale has remained elusive in the literature. For example, Lovibond and Lovibond (1995b) have suggested that this scale measures the constructs similar to those originated from the stress research (e.g., appraisal and coping). On the other hand, Brown et al. (1997) have argued that this scale is a measure of general negative affect/distress. Additionally, some obtained less

conclusive evidence for existence of the Stress factor in their factor analytic studies (e.g., Szabó, 2010). Finally, the specificity of both the Stress and Anxiety factors (Chin et al., 2018) and all three DASS–21 specific factors have been questioned recently (Osman et al., 2012).

Given the proposed idea that the variance of the DASS–21 items stems from two sources simultaneously: one general dimension (NA) and the specific dimensions (low positive affectivity, high physiological arousal, and possibly stress), the most promising analytic tool to discern the issue of specificity versus non-specificity is bifactor modelling (Reise et al., 2018). Using this statistical tool in both clinical (Bottesi et al., 2015) and non-clinical samples (Henry & Crawford, 2005; Jovanović et al., 2014), various authors supported the specificity of the three DASS–21 subscales. In these studies, a bifactor model with one general (G) factor and three specific factors was superior to a correlated three-factor solution or a bifactor model with one G factor and Depression and Anxiety as specific factors while the stress items were constrained to load only on the G factor. However, Tully et al. (2009) and Szabó (2010) opted for a bifactor solution with two specific factors in adolescent samples. More recently, Chin et al. (2018) reported the results of both exploratory and confirmatory bifactor models on non-clinical samples supporting a bifactor solution in which only the revised Depression subscale had sufficient validity.

However, in many bifactor modelling studies, the conclusion regarding the specificity of all three scales was reached by focusing mainly on the overall model fit. The authors rarely reported additional indicators such as existence and the number of substantive factor loadings per specific factor or how much reliable variance remains within the scales once the reliability due to the G factor has been accounted for. However, the studies which considered these additional criteria, on non-clinical samples did not provide evidence for sufficient specificity of all three DASS–21 subscales (Osman et al., 2012), or the Anxiety and Stress scales (Chin et al., 2018). No bifactor modelling study on clinical samples, with detailed reliability and dimensionality indices, has been reported, leaving the opportunity that some inconsistent findings might have arisen from insufficient reporting practices. Also, examining a clinical sample and a sample drawn from a general population (i.e., Facebook sample) in a single study could clarify if factor structure may have varied in prior studies, in part, because of a possibility that the DASS–21 could have different factor structure across various levels of negative affectivity seen in different samples.

In contrast to the recent bifactor studies in non-clinical samples, a review of literature suggests that most of the factor analytic studies in clinical samples provided support for the correlated three-factor model (e.g., Antony et al., 1998; Clara et al., 2001). Even though bifactor modelling has been offered as a useful tool in discerning dimensionality of psychological instruments (e.g., Reise et al., 2018), it also has certain limitations such as the problems of overfitting or better fitting in comparison to second-order or correlated models even when the tested model does not have a bifactor structure (Eid et al., 2018; Watts et al., 2019). Hence, it is advisable to base decisions about the preferred model not solely

on the model fit but also comparison of external validity of competing models (Watts et al., 2019).

Besides comparing the bifactor and correlated-factor models, we believe that it would be informative to test a second-order model with three first-order factors (anhedonic depression, physiological hyperarousal, and stress), and one higher-order factor. Even though this model is statistically equivalent to the correlated-model, the second-order model can provide more interpretable solution, especially in the area when one can hypothesize the existence of one higher-order factor such as general negative affectivity (Chen et al., 2005).

Also, a comparison between the second-order and bifactor models deem important because they can inform our understanding the way the G and putative specific factors influence the common DASS-21 variance. Both models have been used to represent various psychopathological phenomena which are believed to be hierarchically organized (see Markon, 2019 for a review). However, they differ in important ways. The second-order model hypothesizes that the specific factors, which are subordinate to the G factor, mediate the effects of the G factor on the DASS-21 item variance. On the other hand, the bifactor model tests the possibility that both the G and specific factors directly and independently influence the common DASS-21 variance (Reise et al., 2018).

Apart from factor analytic studies, another line of evidence regarding construct validity of the DASS-21 explores its relations with various dimensional measures of stress, depression, and anxiety (e.g., Jovanović et al., 2014; Lovibond & Lovibond, 1995b), and also with different symptom measures (Brown et al., 1997). Based on these relations, some authors concluded that the Depression and the Anxiety scales have a certain amount of specificity whereas the Stress scale taps into the general distress (Brown et al., 1997). Given the recent finding that the most of the DASS-21 item variance stems from one general dimension (Osman et al., 2012), it remains unclear what underlies different correlation patterns between the specific DASS-21 scales and other external criteria i.e., whether the systematic variance due to the G factor, the specific factors, or both are responsible for the observed relations. Additionally, a number of validation studies conducted in clinical samples is still small. More importantly, in these studies there was no correspondence between the supported factor analytic models of the DASS-21 and the scores used for validation. For example, Bottesi and colleagues (2015) supported a bifactor model; however, inferences about validity were based on partial correlations rather than the scores implied by the supported model. Given that the bifactor model was supported, a more appropriate way to analyze data was to use general and residualized group factor scores (i.e., what remains after the G factor was accounted for) while exploring their relation with external measures.

The aim of this study was to provide a comprehensive test of several competing, latent factor models of the DASS-21 that were reported in research studies using both a sample of clinically depressed patients and a Facebook sample from Serbia. The tested models were: a) the single-factor model; b) the

correlated three-factor model; c) the second-order model with one second-order factor subsuming three specific factors; d) the bifactor model with three specific factors; and e) the bifactor model with two specific factors (Depression and Anxiety) while the stress items were allowed to load only on the G factor. The last model has never been tested in clinical samples, and some previous studies reported contradictory findings in general population (e.g., Henry & Crawford, 2005; Szabó, 2010). Using the samples, drawn from different populations, can clarify whether the DASS–21 factor structure varied across prior studies partly because of the symptom severity.

Another aim was to provide further evidence for validity of the DASS–21 scale. In this study, convergent validity of the DASS–21 scale was tested using the Beck's Depression Inventory (BDI–II) and the State-Trait Anxiety Inventory (trait form; STAI–T) in a sample of depressed individuals, but using the scores supported by the latent structure of the instrument. In light of recent criticism of bifactor modelling, especially its tendency to overfit, it would be important to compare contributions of the factors stemming from the bifactor and other competing models while predicting external criteria.

Method

Participants

A patient sample was recruited from nine psychiatric hospitals in Serbia, in a period from 2014 to 2016. 302 patients (52% females; 48% inpatients; $M_{\text{age}} = 52.21$ and $SD_{\text{age}} = 11.56$) signed consent forms to participate in the study. Most participants had high school education (41.1%). Patients were classified, according to the International Classification of Mental Disease–10, by one mental health professional based on clinical interviews, the results of clinical psychological testing, and case notes. Patients had the following primary diagnoses: depressive episode (37%), recurrent depressive disorder (38%), dysthymia (3%), unspecified affective disorder (4%), brief depressive adjustment reaction (1%), prolonged depressive adjustment reaction (1%), mixed anxiety and depressive reaction (5%), and bipolar affective disorder– current depressive episode (12%). 46% patients had co-morbid diagnoses. Psychosis and dementias were exclusionary criterion.

A non-patient sample included 376 participants (76.9% females; $M_{\text{age}} = 29.12$, $SD = 8.96$) who were recruited via Facebook in 2015, using a Google forms online survey. Participants were informed that the survey was anonymous, and if they accepted to fill out the survey they would be asked about their typical emotional, cognitive, and behavioral reactions and experiences in everyday life. Study invitation was shared in different student Facebook groups.

Measures

The Depression Anxiety and Stress Scale–21 (DASS–21)

DASS–21 (Lovibond & Lovibond, 1995a; Jovanović et al., 2014) is a self-assessment inventory for registering presence of the symptoms of depression (e.g., “I couldn't seem to experience any positive feeling at all”), anxiety (e.g., “I was aware of dryness of my mouth”), and stress (e.g., “I found it hard to wind down”) in the past two weeks. It consists of 21 items, with seven items per scale, followed by 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*). Psychometric properties of the

Serbian adaptation of the instrument were previously reported (Jovanović et al., 2014). In the present study, internal consistencies (α) of the scales were very high: $\alpha_{\text{depression}} = .91$, $\alpha_{\text{anxiety}} = .87$ and $\alpha_{\text{stress}} = .88$.

The Beck Depression Inventory–II (BDI–II; Beck et al., 1996; Mihić & Novović, 2019) is a multiple-choice, 21-item (e.g., “Concentration Difficulty”) self-report measure of severity of symptoms of depression. Validity and reliability indices of the Serbian adaptation of the BDI–II were reported previously (Mihić & Novović, 2019). Each answer is scored on a scale ranging from 0 to 3. In the present study, internal consistency was very high, $\alpha = .94$.

The State Trait Anxiety Inventory–Trait (STAI–T)

The STAI–T (Spielberger et al., 1983; Mihić & Novović, 2018) consists of 20 items measuring trait anxiety. Validity and reliability indices of the Serbian adaptation of the STAI were reported previously (Mihić & Novović, 2018). Each item is followed by a 4-point Likert scale (from 1 = *not at all* to 4 = *very much so*). Its internal consistency in the present study was very high, $\alpha = .95$.

All three questionnaires are adaptations of the original questionnaires into the Serbian language. The BDI–II and STAI–T were adapted into Serbian using the back-translation method (Mihić & Novović, 2018; Mihić & Novović, 2019). The Serbian adaptation of the DASS–21 was obtained from the official DASS website (<http://www2.psy.unsw.edu.au/Groups/Dass/Serbian/Serbian.htm>).

Data Analytic Plan

In the clinical sample, a small number of missing values (from 0.4 % to 3.9%) was replaced using the *Expectation Maximization* algorithm implemented in the SPSS software v21 (IBM corp., 2011). Two patients were identified as univariate outliers on the standardized scores, then winsorized. There were no missing values and univariate outliers in the Facebook sample.

Multivariate outliers (6 patients and 10 Facebook users) were excluded from the data sets. Mardia’s coefficients (Mardia, 1974) were 14.27 and 16.49 for the clinical and the Facebook samples, respectively. To account for multivariate nonnormality and ordinal measurement level of the DASS–21, diagonally weighted least squares (WLSMV) estimator was used. Confirmatory factor analysis (CFA) was calculated using the “lavaan” package (Rosseel, 2012) within the R environment (R Core Team, 2016). The following fit indices were used: Comparative Fit index (CFI > .95), Tucker-Lewis index (TLI > .90), the Root Mean Square Error of Approximation (RMSEA ≤ .06), and Akaike Information Criterion (AIC) (Akaike, 1987; Hu & Bentler, 1999; Tanaka, 1993).

For better understanding of the DASS–21 dimensionality the following indicators were calculated: omega for the total score and the subscales (ω and ω_s), omega hierarchical (ω_h), omega hierarchical for each subscale (ω_{hs}), H index, and Factor determinacy (FD; Dueber, 2017). ω and ω_s reflect the proportion of reliable variance attributable to all sources of common variance i.e., variance due to the G and specific factors (Reise et al. 2018). ω_h coefficient estimates how much of the total DASS–21 score variance is attributable to the G factor, whereas ω_{hs} reflects systematic variance that is left once individual variability due to the general factor was partitioned (Reise et al., 2018). The value of $\omega_h > .80$ suggests that the scale can be assumed to measure a unidimensional construct (Reise et al., 2018). Given that some authors pointed out to the limitations of ω_{hs} as reliability coefficients (Rodríguez, Reise, & Haviland, 2016), we considered additionally H index as a measure of factor replicability, with the values > .80 suggesting well-defined latent variables (Hancock & Mueller, 2001). Finally, FD, the correlation between factor scores and the factors, was considered, with

values $> .90$ suggesting a reliable measure (Gorsuch, 1983). However, there is also a recent recommendation that the FD^2 or H values above $.70$ might suggest the presence of specific factors (Rodriguez et al., 2016).

Results

Descriptive Statistics

From Table 1 it can be seen that all measures were normally distributed (Tabachnick & Fidell, 2013) in both samples. In the clinical sample, the values on the three DASS–21 scales were significantly higher than their counterparts in the Facebook sample, $F(3, 668) = 85.92, p = .000, \lambda = .72, \eta_p^2 = .28$, suggesting a higher intensity of depression, $F(1, 670) = 177.44, p = .000; \eta_p^2 = .21$, anxiety, $F(1, 670) = 80.92, p = .000; \eta_p^2 = .11$, and stress, $F(1, 670) = 25.71, p = .000; \eta_p^2 = .04$ within the former. Also, compared to the DASS–21 norms (Lovibond & Lovibond, 1995a), the level of symptoms within the clinical sample ranged from moderate to severe, whereas the participants in the Facebook sample endorsed mild levels of the symptoms.

Table 1
Descriptive Statistics for the Clinical and the Facebook Samples

Sample	Measures	M	SD	Sk	Ku
Clinical	Depression	21.00	11.74	0.04	-0.91
	Anxiety	16.10	10.11	0.44	-0.52
	Stress	21.92	10.02	-0.19	-0.71
	STAI–T	47.66	5.55	-0.19	0.69
	BDI	26.71	14.26	0.11	-0.81
Facebook	Depression	9.80	9.56	1.25	1.11
	Anxiety	9.28	9.85	1.21	0.75
	Stress	17.94	10.40	0.42	-0.49

Note. Sk = Skewness; Ku = kurtosis. The DASS–21 scores were multiplied by 2.

Confirmatory Factor Analysis (CFA)

Five CFA models were tested in both samples. Fit indices for all models are shown in Table 2. In the clinical sample, all fit indices for Models 4 and 5 were satisfying, showing that these two models fitted the data slightly better than the other models. However, the lower AIC value for Model 4 in comparison to Model 5 was in favour of the former ($\Delta AIC > 10$; Burnham & Anderson, 2004). In the Facebook sample, the best fitting model was Model 5 while Model 4 did not converge.

Table 2*Fit Indices Obtained on the Clinical and the Facebook Samples*

		WLSMV χ^2	df	RMSEA (95% CI)	CFI	SRMR	TLI	AIC
Clinical sample	M1	799.6	189	.105 (.097–.112)	.946	.077	.940	967.62
	M2	511.2	186	.077 (.069–.085)	.971	.054	.968	685.24
	M3	511.2	186	.077 (.069–.085)	.971	.054	.968	685.24
	M4	366.5	186	.063 (.054–.072)	.983	.042	.978	576.53
	M5	402.1	175	.066 (.058–.075)	.980	.045	.976	598.13
Facebook sample	M1	934.1	189	.103 (.096–.109)	.920	.077	.911	1102.17
	M2	519.2	186	.069 (.062–.076)	.964	.055	.960	693.23
	M3	519.2	186	.069 (.062–.076)	.964	.055	.960	693.23
	M4	model has not converged						
	M5	393.6	175	.058 (.050–.065)	.976	.046	.972	589.66

Note. M1 = the single factor model; M2 = the second-order three-factor model; M3 = the correlated three-factor model; M4 = the bifactor model with three specific factors (Depression, Anxiety, Stress); M5 = the bifactor model with two specific factors (Depression and Anxiety) while the Stress items were allowed to load only on the G factor. All WLSMV χ^2 values were significant at $p = .000$ level. M2 and M3 models are statistically equivalent.

Standardized factor loadings for best fitting models in both samples are shown in Table 3a. Given that it was difficult to select one best fitting model in the clinical sample based only on the fit criteria, both Models 4 and 5 were considered. As can be seen, all items had significant loadings on the G factor in all models. All items from the Depression factor had significant loadings in all models in both samples. All items that were expected to load on the Anxiety factor had significant loadings in both samples, with the exception of the A9 item in the clinical sample (“I was worried about situations in which I might panic and make a fool of myself.”). However, none of the Stress items loaded significantly on their putative specific factor in Model 4 in the clinical sample. Hence, although the fit indices were slightly in favour of Model 4 compared to Model 5 in this sample, the size of the factor loadings suggested that there was not enough evidence to extract the Stress factor.

We also considered various reliability indices (see the lower part of Table 3) obtained for Model 5 in both samples. The values of ω_h suggest that 92% and 91% of the variance for the composite DASS-21 score was accounted for by the G factor in the clinical and Facebook samples, respectively. Even though ω_h s for the Depression and the Anxiety subscales were high in both samples, their relative omega coefficients suggested that only 20% (.186/.938) and 22% (.204/.909) of the reliable variance within these scales left once the reliability due to the G factor was controlled for in the clinical sample. Similar statistics for both scales were obtained in the Facebook sample, 24% and 25%.

Finally, only the G factor satisfied both sets of recommendations for FD and H in both samples (see the data analysis section). However, the conclusions about replicability and determinacy of the Depression and the Anxiety factors varied, depending on the criteria. According to one set of recommendations (Gorsuch, 1993; Hancock & Mueller, 2001), these specific factors approach the recommended values for factor determinacy but fall short of the recommended replicability value in both samples. However, according to newer, more lenient criteria (Rodriguez et al., 2016), individual variability measured by the Depression and Anxiety subscales could have reflected true variability in their respective latent variables, whereas their stability coefficients approached .70 in both samples. It is, however, noteworthy that despite lower H values than recommended, both Depression and Anxiety subscales replicated to a large degree when one compares the solutions from our two samples (see Model 5 in Table 3).

Table 3

Standardized Factor Loadings and Dimensionality Measures for Best Models in Both Samples

	Clinical sample						Facebook sample			
	Model 4			Model 5			Model 5			
	Dep.	Anx.	Stress	Gen.	Dep.	Anx.	Gen.	Dep.	Anx.	Gen.
D3	.224**			.579**	.231**		.576**	.310**		.578**
D5	.312**			.717**	.320**		.714**	.117*		.653**
D10	.388**			.782**	.396**		.779**	.585**		.667**
D13	.268**			.857**	.276**		.854**	.352**		.769**
D16	.428**			.711**	.434**		.708**	.232**		.724**
D17	.476**			.754**	.482**		.750**	.497**		.691**
D21	.418**			.772**	.424**		.768**	.566**		.685**
A2		.522**		.460**		.522**	.457**		.235**	.581**
A4		.438**		.636**		.442**	.633**		.613**	.626**
A7		.453**		.594**		.456**	.591**		.474**	.659**
A9		-.004		.770**		.006	.768**		.178**	.690**
A15		.415**		.727**		.422**	.723**		.307**	.811**
A19		.437**		.700**		.442**	.698**		.604**	.664**
A20		.195**		.767**		.204**	.764**		.386**	.750**
S1			.077	.764**			.768**			.811**
S6			.477	.616**			.652**			.617**
S8			.309	.808**			.831**			.822**
S11			.210	.837**			.851**			.800**
S12			-.067	.928**			.920**			.825**
S14			.257	.607**			.625**			.579**
S18			-.038	.767**			.763**			.791**
ω/ω_s	.938	.909	.925	.971	.938	.909	.969	.922	.927	.967
ω_h/ω_{hs}	.179	.198	.046	.915	.186	.204	.918	.219	.236	.906
H index	.534	.575	.346	.968	.544	.580	.968	.647	.621	.960
FD	.829	.824	.721	.978	.835	.827	.978	.867	.875	.972

Note. Gen. = General factor; Dep. = Depression factor; Anx. = Anxiety factor.

* $p < .01$; ** $p < .001$.

As can be seen from the Table 3, the Stress factor was not detected in the best bifactor model in both samples. However, it was well-defined in the correlated three-factor model (see Table 4). This model had satisfactory fit indices in both samples, which were very close to the bifactor models (Table 3). As was pointed out in the introduction, model fit indicators are not sufficient to support validity of the bifactor model over the competing models. Hence, the question of its validity was addressed next.

Table 4

Standardized Factor Loadings for Models With Correlated Factors in Both Samples

	Clinical sample			Facebook sample		
	Model 3			Model 3		
	Dep.	Anx.	Stress	Dep.	Anx.	Stress
D3	.633**			.661**		
D5	.786**			.699**		
D10	.873**			.822**		
D13	.922**			.870**		
D16	.802**			.795**		
D17	.866**			.826**		
D21	.870**			.841**		
A2		.564**			.652**	
A4		.733**			.807**	
A7		.691**			.794**	
A9		.799**			.749**	
A15		.826**			.901**	
A19		.800**			.840**	
A20		.827**			.855**	
S1			.771**			.817**
S6			.650**			.622**
S8			.831**			.828**
S11			.856**			.804**
S12			.926**			.831**
S14			.621**			.584**
S18			.753**			.797**

Note. Gen. = General factor; Dep. = Depression factor; Anx. = Anxiety factor.
* $p < .01$; ** $p < .001$.

Convergent Validity of the DASS-21 Scores in the Clinical Sample

Before examining validity, several comments are in order. In the best fitting bifactor model in both clinical and Facebook samples, the G factor appeared reliable and stable unequivocally. However, additional indices for the DASS-21 Depression and Anxiety subscales were somewhat mixed, with their ω_{hs} being most troublesome. However, some concerns have been raised about the limitations of ω_{hs} as reliability coefficients (Rodriguez et al., 2016). Besides, the Depression and Anxiety factors, despite their small number of items and low to

moderate loadings, had satisfactory FD and ω_s values as well as stability in our two samples. Therefore, we decided to examine their validity, in addition to G. We created latent variables based on the best fitting model in order to investigate the differential associations with the external variables. We wanted to determine whether these latent variables can incrementally predict external criteria in a theoretically meaningful way after controlling for the G factor variance. For example, we wanted to see whether the Depression and Anxiety factors from the DASS–21 could predict the BDI–II depression and the STAI–T anxiety over and above the G factor. Finally, given that the Stress factor does not exist according to the bifactor model but it does according to the correlated-factor model, we compared the predictive power of the latent variables from both models. The G, Depression, and Anxiety factors from the bifactor model were positive and significant predictors of the BDI–II scores (see Table 5), $F(3, 292) = 77.567, p < .000$. The G and the Depression factors were significant predictors of the STAI–T anxiety scores, $F(3, 292) = 38.572, p < .000$. Similar predictions of the BDI–II, $F(3, 292) = 72.411, p < .000$, and the STAI–T, $F(3, 292) = 38.695, p < .000$, were obtained for the Depression and Anxiety factors from the correlated three-factor model. However, the Stress factor from this model was not significant predictor either of the BDI or the STAI–T.

Table 5

Factors from the Bifactor and Correlated-Factor Models as Predictors of the STAI–T and the BDI–II in Clinical Sample

DASS–21 predictors	Criteria	β	t	p	R^2	
Bifactor model					STAI–T	BDI
Depression factor	STAI–T	.14	2.34	.020	.284	.443
	BDI	.26	5.01	.000		
Anxiety factor	STAI–T	.03	.517	.605		
	BDI	.21	4.09	.000		
General factor	STAI–T	.51	10.09	.000		
	BDI	.59	13.45	.000		
Correlated model						
Depression factor	STAI–T	.44	2.76	.006	.284	.427
	BDI	.58	4.01	.000		
Anxiety factor	STAI–T	.12	.758	.449		
	BDI	.35	2.37	.018		
Stress factor	STAI–T	-.014	-.054	.957		
	BDI	-.25	-1.09	.273		

Note. $N = 296$.

Discussion

In this study we explored the structural characteristics of the DASS–21 in two samples, clinical and Facebook. We identified several methodological issues in the previous studies which dealt with the structural characteristics of

the DASS-21, precluding making clear conclusions regarding its structure. In our study conducted in Serbian, we contrasted several factor analytic models that were reported in the literature and found a strong support that all DASS-21 items measure a single underlying construct in both samples, and that its bifactor structure is similar across clinical and Facebook samples. All items had substantial factor loadings on the G factor and this factor accounted for 92% and 91% of the common variance in the clinical and Facebook samples, respectively. The items from all three subscales were strongly and similarly related to this factor, supporting the interpretation of the G factor as a tendency to experience general distress i.e., a tendency to experience a mixture of negative emotions such as depression, anxiety, and to be stress-reactive. This interpretation of the G factor is consistent with our finding that the G predicted both the BDI-II and the STAI-T scores. The BDI-II measures severity of depression whereas the STAI-T measures a composite of negative affect and anxiety (Bados et al., 2010). Hence, the general DASS-21 factor seems to represent severity of general distress. Finally, similarity between factor structures in the clinical and Facebook samples supports the idea that the structure of negative affect, measured via the DASS-21, is isomorphic across individuals regardless of the intensity of their affect.

Existence of the G factor, at least when emotional disorders are considered, is in accordance with some older theories purporting to explain comorbidity among emotional disorders, such as the tripartite model (Clark & Watson, 1991), but also with more recent theorizing such as the hierarchical taxonomy of psychopathology (e.g., Kotov et al., 2017). According to these theories, comorbidity between anxiety and depression is explained by the existence of one common vulnerability contributing to each of these conditions and their mixed states. It is usually viewed as a temperamental liability such as negative affectivity and/or neuroticism. Hence, those wishing to pursue preventative research in this area or treatment selection should feel safe to use the total Serbian DASS-21 score to identify vulnerable individuals. Evidence for the existence of two specific factors, Depression and Anxiety, in both samples was somewhat mixed and more difficult to discern. There were several indicators supporting existence of these two specific factors: a) a sufficient number of items with satisfactory loadings on their respective specific factors, b) omega coefficients (ω) for the Depression and Anxiety subscales were over .90 in both samples, suggesting that both are highly reliable multidimensional composites i.e. their unit-weighted scores are precise estimate of all sources of common variance, c) their FD scores were acceptable, d) they replicated between the two samples in our study, and e) both specific factors predicted the external criteria, controlling the contribution of the G factor – the Depression factor predicted both the BDI-II and the STAI-T whereas the Anxiety factor predicted only the BDI-II. The relation between various depression scales and the STAI-T is often reported in literature (e.g., Bados et al., 2010), demonstrating once again that the STAI-T items reflect depression and negative affect rather than

anxiety. The predictive relation between the Anxiety factor and the BDI-II seems somewhat surprising, but it could suggest that somatic aspects of anxiety, which are measured by the Anxiety scale, are better represented in some BDI items (e.g., irritability, agitation) than in the STAI-T. Also, this physiological arousal is often seen in depressed patients, either as a standalone symptom or a comorbid symptom with fear disorders characterized by physiological arousal (panic disorder, social anxiety disorder, or specific phobias). Finally, the STAI-T is supposed to assess feelings of stress, worry, and discomfort (Spielberger et al., 1983), whereas the DASS-21 Anxiety subscale taps more of those fear-related symptoms. Future research should include a more encompassing set of outcome measures to demonstrate further external validity of the specific factors, especially the Anxiety factor. In particular, one might hypothesize that this subscale might predict, independently from the Depression and G factors, a wider range of fear-related constructs such as panic and panic-like symptoms or specific phobias.

However, it seems that both the Depression and the Anxiety subscales possess sufficient reliability only when it comes to assessment of a composite variance stemming from both the general distress and their respective specific factors. Once the contribution of the general factor was left out, the reliability of the subscales did not seem remarkable (see ω_{hs}). However, as others noted, these scores represent residualized variables and not “observed scores” making their interpretation as reliability coefficients difficult (Rodríguez et al., 2016). Given these problems, others have suggested that among the bifactor model indices a greater emphasis should be given to H index and how much the specific factors are likely to replicate across studies (Watts et al., 2019). Considering all evidence, one can recommend that, in addition to the total score, the summed Depression subscale score can be used as a measure of anhedonia in both clinical and research settings. Concerning the Anxiety subscale, until further external validation is obtained, it is safe to conclude that this subscale has a potential to reliably estimate its specific content (physiological arousal). However, given that this subscale is a reliable composite of common and specific variance, it can reliably identify those individuals who are highly distressed and, at the same time, physiologically reactive in fear-evoking situations. Many individuals seen in therapeutic contexts can be characterized as such. Hence, use of the summed Anxiety subscale score can be safe in this context. However, if one wanted to use this scale in a research context (e.g., one wanted to examine the hierarchical structure of emotional disorders when it would be important to estimate precisely only physiological arousal) until further evidence regarding its external validity is obtained, one would be advised to use different instruments⁴. Finally, low ω_{hs} for the Depression and Anxiety subscales might be a result of the small number of items comprising them. One could resolve this problem by reconsidering the

4 For example, one could use the Mood and Anxiety Symptom Questionnaire (Watson & Clark, 1991) designed to measure the tripartite model constructs, including physiological arousal.

initial item pool based on which the DASS-21 was created and to select the items which have substantial loadings on both the G and their respective specific factors. Another option would be to consider addition of items that would emphasize the specific component of each subscale.

Different from some studies, including one previously conducted in Serbian (Jovanović et al., 2014), the best fitting bifactor model did not support retention of the Stress subscale. In the clinical sample, although the bifactor model suggested that this factor could be isolated, there was insufficient number of items to justify creation of the specific subscale. In the Facebook sample, the bifactor model with the three specific factors could not converge, suggesting misspecification of the model. However, the Stress factor was clearly identified in the correlated three-factor solution with satisfactory fit indices in both samples. However, it failed to predict the external criteria, over and above the Depression and Anxiety factor, supporting our interpretation that this subscale reflects general distress. It is unfortunate that that Stress subscale cannot tap tension, apart from the negative affect, given that some neurophysiological studies of anxiety disorders make a clear distinction between anxious apprehension (akin to worry and tension) and anxious arousal (panic) (Burdwood et al., 2016; Nitschke et al., 1999). The Stress subscale item content seems to reflect anxious apprehension, whereas the Anxiety subscale reflects somatic anxiety. If one wants to further revise the DASS-21 in order to make this distinction, our recommendation would be to include more items that would measure apprehension of danger. It should also be emphasized that our validation measures were not perfectly matched to demonstrate validity of this scale. Hence, further studies should aim at including the measures which would be more appropriate such as physiological measures of tension, blood pressure, or patterns of brain activity.

Limitations

The study has a certain number of limitations. We focused on self-report data only. As was pointed out in the discussion section, further validation of the DASS-21 subscale might be obtained by including different sets of physiological measures. We used convenience samples. Even though both samples exceeded 200 cases, which could be considered “large” (Kline, 2005), it would still be advisable to replicate the findings given that the ratio of cases to the number of free parameters was below 10 (Jackson, 2003). Finally, bifactor models have been criticised to perform better than correlated or second-order models even when the true model does not have bifactor structure (e.g., Watts et al., 2019). However, we believe that our results support validity of the best fitting bifactor model. For example, the G, Depression, and Anxiety factors were well represented by their respective indicators. The relations between the Depression and Anxiety factors and external correlates were similar in the bifactor and the correlated three-factor model. However, the Stress factor that was only identified in the latter did not receive validation support, justifying our preference for the bifactor model.

Conclusion

Results from our study suggest that the Serbian version of the DASS–21 assesses reliably the general distress (via the total DASS–21 score) and anhedonia (via the Depression subscale score). The Anxiety subscale can be safely used in clinical settings when an aim of a clinician is to assess individuals presenting with a mixture of distress and anxious arousal. However, until further validation support is provided, we do not recommend the use of this scale when the task is to estimate precisely the anxious arousal only. The Stress subscale assesses general distress only. For further insights into the specificity of the DASS–21 subscales and their discriminant power, samples should be expanded to include patients with anxiety disorders alone but also depression-anxiety comorbid states.

References

- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52, 317–332.
- Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, 10(2), 176–181. doi:10.1037//1040-3590.10.2.176
- Bados, A., Gomez-Benito, J., & Balaguer, G. (2010). The statit-trait anxiety inventory, trait version: Does it really measure anxiety? *Journal of Personality Assessment*, 92(6), 560–567. doi: 10.1080/00223891.2010.513295
- Beck, A., T., Steer, R., A., & Brown, G., K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX.
- Bottesi, G., Ghisi, M., Altoe, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, 60, 170–181. doi: 10.1016/j.comppsy.2015.04.005
- Brailovskaia, J., Schönfeld, P., Zhang, X. C., Bieda, A., Kochetkov, Y., & Margraf, J. (2017). A cross-cultural study in Germany, Russia, and China: Are resilient and social supported students protected against depression, anxiety, and stress. *Psychological Reports*, 121(2), 265–281. doi:10.1177/0033294117727745
- Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy*, 35(1), 79–89. doi:10.1016/s0005-7967(96)00068-x
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33, 261–304. doi: 10.1177/0049124104268644
- Chen, F. F., Sousa, K. H., & West, S. G. (2005) Teacher's Corner: Testing Measurement Invariance of Second-Order Factor Models, Structural Equation Modeling. *A Multidisciplinary Journal*, 12(3), 471–492. doi: 10.1207/s15328007sem1203_7
- Chin, E. G., Buchanan, E. M., Ebesutani, C., & Young, J. (2018). Depression, Anxiety, and Stress: How Should Clinicians Interpret the Total and Subscale Scores of the 21-Item Depression, Anxiety, and Stress Scales? *Psychological Reports*, 122(4), 1550–1575. doi:10.1177/0033294118783508
- Clara, I. P., Cox, B. J., & Enns, M. W. (2001). Confirmatory factor analysis of the Depression-Anxiety-Stress scales in depressed and anxious patients. *Journal Of Psychopathology and Behavioral Assessment*, 23(1), 61–67. doi:10.1037/pas0000207

- Clark, L.A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316–336. doi:10.1037/0021-843x.100.3.316
- Dueber, D. M. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. <https://dx.doi.org/10.13023/edp.tool.01> [Available at <http://sites.education.uky.edu/apslab/resources/>]
- Eid, M., Krumm, S., Koch, T. & Schulze, J. (2018). Bifactor models for predicting criteria by general and specific factors: Problems of nonidentifiability and alternative solutions. *Journal of Intelligence, 6*(3), 1–23. <http://dx.doi.org/10.3390/jintelligence6030042>
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking Construct Reliability within Latent Variable Systems. In R. Cudeck, S. du Toit, & D. Sorbom (Eds.), *Structural Equation Modeling: Present und Future—A Festschrift in Honor of Karl Joreskog* (pp. 195–216). Lincolnwood, IL.
- 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*(2), 227–239. doi: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: A Multidisciplinary Journal, 6*, 1–55. doi:10.1080/10705519909540118
- IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY.
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling, 10*, 128–141.
- Jovanović, V., Gavrilov-Jerković, V., Žuljević, D., & Brdarić, D. (2014). Psychometric evaluation of the depression anxiety stress scales–21 (DASS–21) in a Serbian student sample. *Psihologija, 47*(1), 93–112. doi:10.2298/PSI1401093J
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd). New York, NY.
- Kotov, R., Watson, D., Bagby, R. M., Carpenter, W. T., Clark, L. A., Forbes, M. K., ...Wright, A. G. C. (2017). The hierarchical taxonomy of psychopathology (HiTop): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology, 126*, 454–477. doi: /10.1037/abn0000258
- Lovibond, S., H., & Lovibond, P., F. (1995a). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: Psychology Foundation of Australia.
- Lovibond, P. F., & Lovibond, S. H. (1995b). 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*(3), 335–343. doi:10.1016/0005-7967(94)00075-u
- Mardia, K. (1974). Applications of Some Measures of Multivariate Skewness and Kurtosis in Testing Normality and Robustness Studies. *Sankhyā: The Indian Journal of Statistics, Series B* (1960–2002), *36*(2), 115–128. Retrieved from <http://www.jstor.org.proxy.library.ohio.edu/stable/25051892>
- Markon, K. E. (2019). Bifactor and Hierarchical Models: Specification, Inference, and Interpretation. *Annual Review of Clinical Psychology, 15*(1), 51–69. doi: 10.1146/annurev-clinpsy-050718-095522
- Mihić, Lj. & Novović, Z. (2018). *STAI– upitnik anksioznosti kao stanja i osobine ličnosti [State–Trait Anxiety Inventory for Adults]*. Beograd, Srbija.
- Mihić, Lj. & Novović, Z. (2019). *BDI–II Bekov inventar depresije II [Beck Depression Inventory–II (BDI–II)]*. Beograd, Srbija.
- Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The Depression Anxiety Stress Scales–21 (DASS–21): Further examination of dimensions,

- scale reliability, and correlates. *Journal of Clinical Psychology*, 68(12), 1322–1338. doi:10.1002/jclp.21908
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2018). Bifactor modeling and the evaluation of scale scores. In P. Irwing, T. Booth, & D. Hughes (Eds.), *The Wiley-Blackwell handbook of psychometric testing* (pp., 675–707). New York, NY.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98, 223–237.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <http://www.jstatsoft.org/v48/i02/>.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA.
- Szabó, M. (2010). The short version of the Depression Anxiety Stress Scales (DASS–21): Factor structure in a young adolescent sample. *Journal of Adolescence*, 33(1), 1–8. doi: 10.1016/j.adolescence.2009.05.014
- Tabachnick, B. G. & Fidell, L. S. (2013). *Using Multivariate Statistics*. Pearson, Boston.
- Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 10–39). Newbury Park, CA.
- Tully, P. J., Zajac, I. T., & Venning, A. J. (2009). The structure of anxiety and depression in a normative sample of younger and older Australian adolescents. *Journal of Abnormal Child Psychology*, 37(5), 717–726. doi: 10.1007/s10802-009-9306-4
- Watts, A. L., Poore, H. E., & Waldman, I. D. (2019). Riskier Tests of the Validity of the Bifactor Model of Psychopathology. *Clinical Psychological Science*, 7(6), 1285–1303. <https://doi.org/10.1177/2167702619855035>

Skala depresije, anksioznosti i stresa–21: Bifaktorski statistički indeksi govore u prilog totalnog skora i skora depresivnosti

Ljiljana Mihić, Marija Volarov, Milan Oljača i Zdenka Novović

Odsek za psihologiju, Filozofski fakultet, Univerzitet u Novom Sadu, Srbija

U ovoj studiji je ispitano nekoliko modela latentnih faktora Skale depresije, anksioznosti i stresa–21 (eng. Depression, Anxiety and Stress Scale, DASS–21) na uzorku klinički depresivnih pacijenata i na uzorku korisnika Fejsbuka iz Srbije. Skala DASS–21, Bekov inventar depresivnosti–II i Skala anksioznosti kao crte su zadate uzorku depresivnih pacijenata ($N = 296$; $M = 52.21$, $SD = 11.56$). Korisnici Fejsbuka su popunjavali samo skalu DASS–21 ($N = 376$; $M_{age} = 29.12$, $SD_{age} = 8.96$). Podaci govore u prilog bifaktorskog rešenja sa jednim opštim – faktorom distresa (G) i dva specifična faktora (depresivnost i anksioznost). Ova tri faktora su imala visoke omega koeficijente, dok su hijerahijski omega koeficijenti za depresiju i anksioznost bili niski. Na osnovu ukupnih rezultata ovog istraživanja – podataka o eksternoj validnosti, odredivosti faktora i replikabilnosti, zaključili smo da srpska verzija DASS–21 skale pouzdano procenjuje opšti distres i anhedoniju, kako kod osoba sa kliničkim nivoima distresa, tako i u opštoj populaciji. Subskala anksioznosti se može sa sigurnošću koristiti

u kliničkom kontekstu kada je potrebno uraditi psihološku procenu osoba koje prezentuju mešavinu distresa i anksiozne pobuđenosti. Međutim, dok se ne obezbede dodatni argumenti u prilog validnosti, ne preporučuje se upotreba ove skale ako je cilj precizna procena anksiozne pobuđenosti. Subskala stresa procenjuje jedino opšti distres. Problem niskih hijerarhijskih omega koeficijenata subskala depresije i anksioznosti može se potencijalno rešiti ponovnim izborom ajtema-indikatora ovih konstrukata iz inicijalne banke ajtema na osnovu koje je DASS-21 skala nastala, uzimajući one ajteme koji imaju supstantivna zasićenja i na opštem faktoru i na odgovarajućem grupnom faktoru.

Ključne reči: bifaktorski model, klinički uzorak, DASS-21, depresija, validacija

RECEIVED: 27.12.2019.

REVISION RECEIVED: 17.04.2020.

ACCEPTED: 26.04.2020.

© 2021 by authors



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution ShareAlike 4.0 International license