Exploring Factor Validity of 20-Item Toronto Alexithymia Scale (TAS-20) in Albanian Clinical and Nonclinical Samples

Aliriza Arenliu¹, Brikena Krasniqi², Kaltrina Kelmendi¹, and Shukrije Statovci³

Abstract
This study aims to examine the factor structure and validity of the Albanian TAS-20 (Toronto Alexithymia Scale) using a sample comprised of 342 students and 196 patients from a psychiatric clinic. Based on a literature review of studies of confirmatory factor analysis (CFA), three types of models were tested: first-order models with method factors and covariances, a second-order model with method factors and covariances, and nested models with method factors and covariances. The findings suggest that a three-factor correlated model with method factors was the best and most parsimonious solution for the clinical sample, exhibiting adequate levels of performance based on the goodness of fit criteria. However, regarding the student sample, the nested three-factor model with method factors and covariances demonstrated a superior fit when compared with the other tested models. Although the total scale of difficulty identifying feelings (DIF) and difficulty describing feelings (DDF) scores provided sound internal consistency, the externally oriented thinking (EOT) subscale did not. Nonetheless, as the CFA suggests the plausibility of negatively keyed items in the method factors, further interpretation of this scale is suggested. This study concludes that the TAS-20 of the Albanian language is appropriate for research purposes, and further research is needed for its application for clinical practice.

Keywords
student, clinical sample, Toronto Alexithymia Scale, factor validity, Kosovo

Introduction
Alexithymia refers to limited ability and individual differences regarding the extent to which one can access, identify, and communicate emotions. While the literature on alexithymia presents several construct definitions, this study focuses on three main models. The first is the most widely used, that is, the Toronto Alexithymia Scale. This instrument considers that individuals with alexithymia can identify their mood states but are unable to identify their emotions, a phenomenon that occurs because they cannot link their affective condition to a specific situation, memory, or expectation (Taylor et al., 1999). This same perspective considers that people with alexithymia have difficulty fantasizing (DFAN). Accordingly, the Toronto model conceptualizes difficulty identifying feelings (DIF), difficulty describing feelings (DDF), externally oriented thinking (EOT), and constrained imaginal processes, also known as DFAN. The Amsterdam model, however, offers an alternative definition to alexithymia as it adds an additional component, specifically, reduced emotional reactivity or difficulty emotionalizing (DEMO) (Vorst & Bermond, 2001), to the components defined in the Toronto model (DIF, DDF, EOT, DFAN). The third model, that of Preece et al. (2017), considers alexithymia to be comprised of DIF, DDF, and EOT, a conclusion that is based on a review of existing studies on factor structure.

Described initially by Nemiah et al. (1976), alexithymia is thought to reflect deficits in the cognitive processing of emotion regulation. However, it has also been reported to be associated with substance abuse (Cleland et al., 2005), post-traumatic stress disorder (Frewen et al., 2006), eating disorders (Nowakowski et al., 2013; Westwood et al., 2017), suicide attempts (Pompili et al., 2013), and depressive disorders (Li et al., 2015). Alexithymia is considered a transdiagnostic construct that provides valuable information with respect to emotion regulation as an important indicator and predictor of various psychopathologies.

¹University of Prishtina, Kosovo
²Labyrinth, Prishtina, Kosovo
³University Clinical Center of Kosovo, Prishtina, Kosovo

Corresponding Author:
Kaltrina Kelmendi, Department of Psychology, University of Prishtina, Nena Tereza Street, n.n., 10000 Prishtina, Kosovo.
Email: kaltrina.kelmendi@uni-pr.edu

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
The most widely used instrument to measure alexithymia is the 20-item Toronto Alexithymia Scale (TAS-20) (Bagby et al., 1994). The TAS-20 consists of three-factor scales, namely, DIF, DDF, and EOT. The DIF factor scale assesses the ability to identify feelings and differentiate those feelings from bodily sensations that accompany emotional arousal; the DDF scale assesses issues associated with describing feelings; and the EOT scale assesses externally oriented modes of thinking. This article uses the Toronto model as its basis for the measures being investigated as the TAS-20 can be administered in the necessary language, is culturally appropriate, and exhibits a good statistical fit.

There is strong evidence that a three-factor solution is generalizable and replicable across languages and cultures (Taylor et al., 2003) and that it can be replicated into many different languages, including German (Bach et al., 1996), Finnish (Joukamaa et al., 2007), Hungarian (Cserjési et al., 2007), Hindi (Pandey et al., 1996), Italian (Bressi et al., 1996), Chinese (Zhu et al., 2007), Japanese (Komaki et al., 2003), Korean (Lee et al., 1996), North American Aboriginal Population (Parker et al., 2007), Portuguese (Prazeres et al., 2000), Spanish (Martínez Sánchez, 1996), Swedish (Simonsson-Sarnecki et al., 2000), Dutch (Meganck et al., 2012), Turkish (Güleç et al., 2009), and Greek (Tsoureas et al., 2010).

Although the majority of factor analytic studies regarding the TAS-20 support the original three-factor solution, some studies indicate a variability in the factorial structure of the TAS-20 across samples. Investigators have reported alternative two-factor and four-factor models to provide a better statistical fit than the original three-factor structure of the TAS-20. Several studies have found two-factor solutions where the DIF and DDF items constitute a single factor, while the EOT items constitute a second factor in both clinical and nonclinical samples (Erni et al., 1997; Kooiman et al., 2002; Loas et al., 1996). In a study where a German translation of the TAS-20 was administered to clinical and nonclinical samples in Germany, Müller et al. (2003) compared five different models with one to four factors and found a four-factor model in which the EOT items were divided into two distinct groups of factors, namely, those associated with pragmatic thinking and a lack of subjective significance or importance of emotions, which provided a better fit to the data when compared with the one-, two-, and three-factor models. Furthermore, in a validation study where the Chinese translation of the TAS-20 was administered to a student sample (Zhu et al., 2007), the four-factor model was determined to provide a better fit than the standard three-factor model. This was not the case when the instrument was administered to a clinical sample, however. In addition, the Spanish version of the TAS-20 that was administered to a Peruvian sample exhibited low internal consistency and low mean item correlations and also failed to replicate the original three-factor structure (Loiselle & Cossette, 2001). That said, scholars such as Gignac et al. (2007) have questioned the validation of the three-factor structure of the TAS-20 (DIF, DDF, and EOT) based on inadequate and poor fit confirmatory factor analysis (CFA) indices. As a consequence, Gignac et al. (2007) have suggested testing nested factor models with one broad general alexithymia (GA) factor, up to four nested substantive factors (DIF, DDF, EOT, and EOT2), and one negatively keyed item factor. Findings from their study indicate that while the TAS-20 was better represented by a nested factor model with five factors, poor reliability with respect to the EOT scale failed to replicate the original three-factor structure (Gignac et al., 2007).

A similar study conducted by Preece et al. (2018) aimed to assess the psychometric properties in a nonclinical population and a clinical population and determined that the three-factor correlated model (DIF, DDF, and EOT) exhibited the best and most parsimonious solution, albeit it did not meet the CFA criteria of fit indices for either sample. Preece et al. (2018) argued that the fit issues were attributable to the poor factor loading of EOT and the reverse-scored item method factor. Recently, Tuliao et al. (2020) tested the TAS-20 factor structure with samples from the United States and the Philippines. In their study, they examined 18 competing factor structure solutions using nested and nonnested models. The results of their study indicated that a bifactor model with common method factor exhibited a better model fit when compared with 17 other models. However, it still did not achieve an adequate fit with the data due to negatively keyed items and poor loadings on the EOT factor and GA factor. With respect to the internal consistency of the subscales, several studies determined that EOT has relatively low Cronbach’s alpha values that range from 0.45 to 0.76 (Bressi et al., 1996; Fukunishi et al., 1997; Loas et al., 2001; Pandey et al., 1996; Picardi et al., 2011). A similar pattern was observed in various studies of clinical populations where the EOT scale had either a low internal consistency or a low correlation with clinical measures. This was observed among patients with anorexia nervosa (Torres et al., 2019), depression and anxiety (Bagby et al., 1994; Berthoz et al., 1999; De Gucht et al., 2004), insomnia (Lundh & Broman, 2006), alcohol-related problems (Thorberg et al., 2010), and substance dependence problems (Loas et al., 2001).

As factor structure is clearly an important issue with respect to the multidimensionality of the alexithymia construct, the replication of the three-factor model of the TAS-20 in other cultures and languages would provide support for the claim that alexithymia is a universal trait (Taylor et al., 2003). Conducting TAS-20 studies in other cultures will contribute to the findings on the multidimensionality of the construct and its stability or variability across samples and cultures, and thus, they will contribute to the diversity of the science. In this regard, we hypothesized the applicability of the three-factor model of alexithymia developed by Bagby et al. (1994) but also a more complex types of models, including higher-order factors or bifactor models proposed by Gignac et al. (2007) and Preece et al. (2018).
The aim of this article is to evaluate the factor validity and reliability of the authors’ Albanian translation of the TAS-20 by administering the instrument to both clinical and nonclinical samples in Kosovo. As a postwar country since 1999, Kosovo is experiencing a period of rapid social and cultural transition as well as economic hardships, all of which are additional burdens for those with unresolved war trauma. Hence, Kosovo’s civilian adults and children are exhibiting a high prevalence of symptoms associated with posttraumatic stress, anxiety, and depressive disorder (Danish Refugee Council [DRC], 2006; Schick et al., 2013). Nonetheless, Kosovars largely ignore or neglect mental health issues, often because of the stigmas attached to such issues (DRC, 2006). Furthermore, there is a lack of validated mental health instruments available in the Albanian language. Accordingly, efforts aimed at evaluating the factor validity and reliability of any Albanian translated instruments are highly encouraged and welcomed by the professionals in the field of mental health treatment and care.

Method

Subjects

A total of 196 outpatients from the psychiatric clinic at the University Clinic of Kosovo and the outpatient clinic for substance abuse treatment participated in the study. Both institutions use the International Classification of Diseases (ICD) 11th revised version as a guide for coding health information and diagnoses.

With respect to the demographic data of the clinical participants, 73% were male and 27% were female. The participants were diagnosed with the following disorders: anxiety (20.4%), depression (17.3%), and substance use (62.3%). The mean age of the patients was 35.1 (SD = 6.1 years with a range of 18–61 years). The nonclinical population consisted of public university students recruited from the University of Prishtina, faculty of philosophy (N = 342). Of the participants, 74.9% were female and 25.1% were male. The mean age of the sample was 19.45 (SD = 1.68 years with a range of 18–30 years). Approval to conduct the research was obtained from the university and the ethics committee, and written informed consent was obtained from all participants at the time of enrollment.

Measures

The TAS-20 (Bagby et al., 1994) is a 20-item self-report questionnaire that measures three dimensions of alexithymia: (a) DIF, which involves distinguishing feelings from the bodily sensations of emotions (seven items); (b) DDF (five items); and (c) EOT (eight items). The respondents are asked to indicate the extent to which they agree or disagree with each statement by recording their responses on a 5-point Likert-type scale. Total scores range from 20 to 100. The English version of the TAS-20 was translated into Albanian. Throughout this process, the translators consulted with the developers of the English version of the TAS-20 (Graeme Taylor, Personal communication, 29/01/2016) to ensure that the intended meaning of each item was captured by the Albanian translation. The procedure included separate translations performed by two authors of the study, back translation, a review of the back translation, a discussion regarding the problematic items with the authors of the original measure, and a review and approval of the final version in coordination with the authors of the original instrument.

Statistical Procedures

The CFAs were conducted using the R package lavaan developed by Rossel (2012). All other statistics were conducted in Statistical Package for Social Science (SPSS). Factor structure was assessed using a series of CFAs, specifically, the weighted least square estimator due to the nonnormal distribution of the data. The factor structure of the TAS-20 was analyzed separately for the student and clinical populations. To evaluate model fit, four fit indices were evaluated, namely, the comparative fit index (CFI), the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), and the Tucker–Lewis index, also known as the nonnormed fit index (NNFI). The following standards were used to evaluate the model: RMSEA values less than 0.05 indicate a good fit; an SRMR cutoff value less than 0.08 is recommended (Brown, 2006; Hu & Bentler, 1999); and CFI and TLI/NNFI values greater than 0.90 indicate a good fit (Bentler, 1990; Brown, 2006). Furthermore, when considering nonnormally distributed data, the standard evaluation criteria for assessing the goodness of fit may not be appropriate, and therefore, more rigorous criteria should be adopted, as suggested by Nye and Drasgow (2011). Specifically, the values of RMSEA and SRMR should be used to evaluate the fit of the Diagonal Weighted Least Square (DWLS) estimator.

Two approaches were used to evaluate the internal reliability and item-to-scale homogeneity of the TAS-20-C. First, we calculated Cronbach’s α coefficients and mean interitem correlations (MICs) for the total scale and for each factor in both the student and clinical samples. Second, as adopted by Gignac et al. (2007) and suggested by Hancock and Muller (2001), instead of using the standard coefficients of internal consistency reliability, the omega (Ω) approach as a framework that represents the true score variance to total variance was used.

Tested Models

Although the three-factor model of alexithymia developed by Bagby et al. (1994) has received the most support, other studies have used more complex models, including higher-order factor models and bifactor models (Gignac et al., 2007;
Table 1. Mean and Standard Deviations for Subscales of TAS-20.

| Populations | DIF      | DDF      | EOT      | Total scale |
|-------------|----------|----------|----------|-------------|
|             | M (SD)   | M (SD)   | M (SD)   | M (SD)      |
| Student     | 18.37 (5.37) | 14.75 (3.21) | 25.03 (3.83) | 58.16 (9.30) |
| Clinical    | 22.82 (6.22) | 16.64 (3.92) | 27.29 (5.11) | 66.76 (11.45) |

Note. TAS-20 = 20-item Toronto Alexithymia Scale; DIF = difficulty identifying feelings; DDF = difficulty describing feeling; EOT = externally oriented thinking; M= mean; SD = standard deviation.

Meganck et al., 2008; Preece et al., 2018; Tuliao et al., 2020). This study aims to understand the factor structure of the TAS-20 for Albanians by administering the instrument to a student sample and a clinical sample and examining the three types of factor solutions proposed by Gignac et al. (2007) and Preece et al. (2018), including first-order models, models with method factor, higher-order models, and nested models. The first-order models tested the one-factor model and the standard three-factor model (DIF, DDF, EOT) for the TAS-20, as reported by Bagby et al. (1994), and a four-factor model in which the EOT factor was split into pragmatic thinking (PR, three items) and lack of importance of emotions (IM, five items). Models with method factor tested the models with the addition of method factor loading on the reversed scored items, specified as orthogonal to the other first-order factors. The higher-order models tested the factor solution where the three first-order factors were specified to load on the one second-order general factor (GF). The method factor solution was then applied to the higher-order models. Finally, the nested model was tested, and the first-order GF and three first-order unique factors (DIF, DDF, EOT) were examined using the method factor.

Results

Preliminary Analysis

Prior to the analysis, the distribution of the items for normality was verified and missing values were explored. As all items were nonnormally distributed based on the Shapiro–Wilk test of normality, the DWLS approach was used in the CFA. In addition, the missing value analysis was conducted and revealed that none of the participants had missing values in excess of 5%. The analysis was based on the sample of students (N = 342) and the clinical sample (N = 196). Table 1 presents the means and standard deviations of the subscales as well as the reliability coefficients for each scale for both groups.

Factor Structure

The fit indices for the tested models for both student and clinical populations are provided separately in Table 2 for all models.

As observed in Table 2, the first-order models indicated poor parameter fit indices for the clinical sample. However, while an acceptable parameter fit was achieved in the student population for both three-factor and four-factor correlated models; the four-factor correlated method exhibited a slightly better fit for TLI with a DIFF of 0.01 and an SRMR with a DIFF of 0.01.

In the second phase, the method factor was employed, which improved the fit indices of the three-factor correlated model (Table 2), as all but one goodness of fit criteria were met, including the RMSEA, CFI, and TLI. The SRMR did not meet the goodness of fit criteria for the clinical sample as it exceeded 0.08, although it was an acceptable fit for the student sample, with a value of 0.053. An examination of the factor loadings found that four of the five items loaded more significantly on the method factor than they did on their supposed EOT factor for both samples (see Table 4). In both samples, item 15 exhibited nonsignificant loading in the method factor. However, it is noted that factor loadings in the method factor were stronger for the clinical sample (0.38–0.83) than they were for the student sample (0.15–0.53). Despite the applicability of the method factor, however, the EOT factor was strongly and negatively associated with DIF and DDF for both samples (see Table 3). Furthermore, the higher-order model of the three-factor solution did not improve the parameter fit criteria of the construct for either sample, specifically, for the three-factor higher-order model and the three-factor higher-order model with the factor method. The most parsimonious model among the higher-order models for both samples was the three-factor higher-order model with method factor and covariances. All the items were significantly loaded on the DIF and DDF factor, whereas for the EOT latent factor, items 5, 18, and 19 were not significantly loaded (Table 4). Moreover, negatively keyed items exhibited stronger loadings in the method factor when compared with their original factor loadings. Nonetheless, the factor loadings for all subscales were weaker when compared with first-order models with the method factor. As evidenced in Table 4, the factor loadings for the DIF latent factor ranged from 0.47 to 0.66 while for the DDF latent factor, the loadings ranged from 0.19 to 0.62. The EOT subscale recorded the weakest loadings, ranging...
### Table 2. Confirmatory Factor Analysis Fit Statistics and Close Fit Indexes.

|                          | χ² (df) | CFI   | TLI   | RMSEA (90% confidence) | SRMR |
|--------------------------|---------|-------|-------|-------------------------|------|
| **Clinical sample**      |         |       |       |                         |      |
| First-order models       |         |       |       |                         |      |
| One-factor model         | 335.338 (169) | 0.861 | 0.844 | 0.070 [0.060, 0.082]    | 0.094|
| Three-factor correlated model | 316.736 (164) | 0.873 | 0.852 | 0.069 [0.058, 0.080]    | 0.091|
| Four-factor correlated model | 315.436 (160) | 0.87  | 0.846 | 0.071 [0.059, 0.082]    | 0.091|
| Models with a method factor |         |       |       |                         |      |
| Three-factor correlated model + method | 797.062 (165) | 0.473 | 0.393 | 0.140 [0.130, 0.150]    | 0.146|
| Three-factor correlated model + method + covariances | 216.456 (162) | 0.917 | 0.903 | 0.056 [0.043, 0.068]    | 0.083|
| **Higher-order models**  |         |       |       |                         |      |
| Three-factor higher-order model | 373.680 (170) | 0.83  | 0.81  | 0.078 [0.068, 0.089]    | 0.99 |
| Three-factor higher-order model + method | 317.662 (162) | 0.873 | 0.853 | 0.069 [0.058, 0.080]    | 0.091|
| Three-factor higher-order model + method + covariances | 261.456 (162) | 0.917 | 0.903 | 0.056 [0.043, 0.068]    | 0.083|
| **Student sample**       |         |       |       |                         |      |
| First-order models       |         |       |       |                         |      |
| One-factor model         | 321.749 (169) | 0.905 | 0.894 | 0.052 [0.043, 0.060]    | 0.072|
| Three-factor correlated model | 287.403 (164) | 0.924 | 0.911 | 0.047 [0.038, 0.057]    | 0.068|
| Four-factor correlated model | 282.365 (160) | 0.924 | 0.91  | 0.048 [0.038, 0.057]    | 0.067|
| Models with a method factor |         |       |       |                         |      |
| Three-factor correlated model + method + covariances | 1028.809 (165) | 0.465 | 0.384 | 0.125 [0.117, 0.132]    | 0.124|
| Three-factor higher-order model + method + covariances | 25.417 (162) | 0.967 | 0.961 | 0.031 [0.019, 0.042]    | 0.058|
| **Nested higher-order models** |         |       |       |                         |      |
| Three-factor nested model | 400.788 (170) | 0.857 | 0.708 | 0.063 [0.055, 0.072]    | 0.079|
| Three-factor nested model + method | 327.268 (165) | 0.9  | 0.844 | 0.054 [0.045, 0.063]    | 0.071|
| Three-factor nested order model + method + covariances | 215.417 (162) | 0.967 | 0.971 | 0.031 [0.019, 0.042]    | 0.058|

Note. CFI = comparative fit index; TLI = Tucker–Lewis Index; RMSEA = root mean square approximation; SRMR = standardized root mean square residual; DF=degrees of freedom.

### Table 3. Estimated Factor Intercorrelations for the Most Parsimonious Models for the Clinical and Student Samples.

|                          | DIF     | DDF     | EOT     |
|--------------------------|---------|---------|---------|
| **Clinical sample**      |         |         |         |
| Three-factor correlated model + method + covariances | 0.877** | (~0.548)** | (~0.832)** |
| Three-factor higher-order model + method + covariances | 0.756** | 0.09 | 0.655** |
| **Student sample**       |         |         |         |
| Three-factor correlated model + method + covariances | 0.967** | (~0.488)** |         |

(continued)
Table 4. Standardized Factor Loadings for the Most Parsimonious Models for the Clinical and Student Samples.

| Clinical sample | Student sample |
|-----------------|----------------|
|                 | DIF            | DDF            | EOT            |
|                 | Three-factor correlated model + method + covariances | Three-factor higher-order model + method + covariances | Three-factor correlated model + method + covariances | Three-factor higher-order model + method + covariances | Nested higher-order model + method + covariances |
| DIF             | B1 0.83        | 0.587          | 0.583          | 0.412          | 0.703          |
|                 | B3 0.726       | 0.514          | 0.683          | 0.483          | 0.663          |
|                 | B6 0.665       | 0.47           | 0.731          | 0.517          | 0.692          |
|                 | B7 0.878       | 0.621          | 0.644          | 0.455          | 0.629          |
|                 | B9 0.713       | 0.504          | 0.788          | 0.558          | 0.723          |
|                 | B13 0.941      | 0.666          | 0.726          | 0.513          | 0.681          |
|                 | B14 0.669      | 0.473          | 0.77           | 0.69           | 0.79           |
| Omega           | 0.86           | 0.81           | 0.77           | 0.69           | 0.79           |
| DDF             | B2 0.887       | 0.627          | 0.633          | 0.447          | 0.802          |
|                 | B4 0.254       | 0.18           | -0.417         | -0.295         | -0.437         |
|                 | B11 0.7        | 0.495          | 0.735          | 0.52           | 0.679          |
|                 | B12 0.62       | 0.438          | 0.584          | 0.413          | 0.521          |
|                 | B17 0.729      | 0.515          | 0.65           | 0.459          | 0.601          |
| Omega           | 0.93           | 0.46           | 0.71           | 0.42           | 0.77           |
| EOT             | B5 -0.022      | 0.015          | 0.031          | -0.022         | -0.194         |
|                 | B8 -0.433      | 0.306          | -0.508         | 0.359          | 0.404          |
|                 | B10 0.29       | -0.205         | -0.175         | 0.124          | 0.27           |
|                 | B15 -0.631     | 0.446          | 0.419          | -0.296         | -0.382         |
|                 | B16 -0.68      | 0.481          | -0.566         | 0.4            | 0.516          |
|                 | B18 -0.155     | 0.109          | -0.138         | 0.097          | 0.319          |
|                 | B19 -0.05      | 0.035          | -0.212         | 0.15           | 0.37           |
|                 | B20 -0.6       | 0.424          | -0.555         | 0.392          | 0.467          |
| Omega           | NEG 0.79       | 0.39           | 0.23           | 0.17           | 0.25           |
|                 | B4 0.609       | 0.609          | 0.156          | 0.156          | 0.225          |
|                 | B10 0.379      | 0.379          | 0.353          | 0.353          | 0.287          |
|                 | B18 0.837      | 0.837          | 0.531          | 0.531          | 0.393          |
|                 | B19 0.669      | 0.669          | 1.011          | 1.011          | 0.914          |

Note. DIF = difficulty identifying feelings; DDF = difficulty describing feeling; EOT = externally oriented thinking.
**p<0.01.
Table 4. (continued)

|                      | Clinical sample | Student sample |
|----------------------|-----------------|----------------|
|                      | Three-factor correlated model + method + covariances | Three-factor higher-order model + method covariances |
|                      | Three-factor correlated model + method + covariances | Three-factor higher-order model + method covariances |
| B15                  | 0.061           | 0.061          |
| Omega                | 0.95            | 0.77           |
| G-factor             |                 |                |
| B1                   | −0.175          | 0.29           |
| B3                   | −0.152          | −0.152         |
| B6                   | −0.231          | −0.133         |
| B7                   | −0.133          | −0.133         |
| B9                   | −0.4            | −0.4           |
| B13                  | −0.342          | −0.268         |
| B14                  | 0.389           | 0.389          |
| B2                   | 0.01            |                |
| B4                   | −0.349          |                |
| B11                  | −0.349          |                |
| B12                  | −0.295          |                |
| B17                  | −0.345          |                |
| B5                   | −0.4            |                |
| B8                   | 0.129           |                |
| B10                  | 0.16            |                |
| B15                  | 0.16            |                |
| B16                  | −0.193          |                |
| B18                  | 0.334           |                |
| B19                  | 0.239           |                |
| B20                  | −0.3            |                |
| Omega                | 1.17            |                |

Note. DIF = difficulty identifying feelings; DDF = difficulty describing feeling; EOT = externally oriented thinking; NEG = negatively keyed items. Italic values are represented nonsignificantly loaded items \( p > .05 \), and bold items are loading below 0.4.

from 0.01 to 0.48. In the final phase, the nested model method was employed. However, as noticed in Table 2, with respect to the clinical sample, the nested model with method and covariance was not identified, whereas in the student sample, the best model was the nested model with method and covariance. Specifically, regarding the student sample, the factor loadings for DIF and DDF were all significantly loaded in their original factors. Similarly, all the items were significantly loaded in the EOT latent factor, with the exception of item 5. All negatively keyed items were significantly and positively loaded in the method factor, with the exception of item 15. However, factor loadings for GFs were weaker for all items and many items were nonsignificantly loaded, specifically, items 3, 6, 7, 9, 14, 4, 10, 15, and 16.

**Reliability**

Reliability is measured using the Cronbach’s alpha coefficient (\( \alpha \)). All items for the nonclinical population resulted in an \( \alpha \) of 0.671, whereas the \( \alpha \) for the clinical sample was 0.771. With respect to subscales the lowest coefficients were observed for DDF, with an \( \alpha \) of 0.28, and EOT, with an \( \alpha \) of 0.20, for the nonclinical sample, and a DDF \( \alpha \) of 0.563 and an EOT \( \alpha \) of 0.361 for the clinical sample. Furthermore, we used omega (\( \Omega \)) as an estimate of the internal reliability of the scales and models (see the bottom of Table 4). As presented in Table 4, with respect to the clinical sample, the \( \Omega \) values for the three-factor correlated model with method and covariance were extremely high, particularly the values for the specific latent factors where the DIF was 0.86, the DDF was 0.93, and the EOTs were 0.79 and 0.95. In contrast, the \( \Omega \) values for the three-factor higher-order model with the method factor and covariances were lower. With respect to the student population, the three-factor nested model with factor method and covariances revealed better omega estimates, with the exception of the EOT latent factor. Specifically, the \( \Omega \) estimates were DIF at 0.79, DDF at 0.77, EOT at 0.25, NEG at 0.53, and GF at 1.17.

**Discussion**

This study aimed to examine the psychometric properties of the TAS-20 when administered to clinical and student
samples of Kosovar Albanians. The results from this study suggest that the TAS-20, for most of the tested models, operates similarly among both samples, albeit differences with respect to the goodness-of-fit criteria were noticed in the nested models. That said, the nested models for the clinical sample were not identified due to the small number of participants. Moreover, the findings suggest that the EOT subscale with the negatively keyed items contaminates the overall fit of the scale.

Of the examined first-order correlated models, it is interesting to note that none of the models meet the goodness-of-fit criteria for the clinical sample; however, both the three-factor and four-factor correlated models exhibit approximately the same fit in the student population. These results are not consistent with previously published CFA results of the TAS-20 when administered to Canadian, American, German, Hindi, Korean, Italian, Lithuanian, Swedish, Turkish, Greek, and French samples (Bagby et al., 1994; Gignac et al., 2007; Güleç et al., 2009; Loas et al., 1996, 2001; Seo et al., 2009; Simonsson-Sarnecki et al., 2000; Tsaousis et al., 2010). With respect to clinical sample concerns, the fit issues are likely attributable to three factors, namely, (a) the poor factor loadings of several EOT items, (b) the presence of negatively keyed items, and (c) the psychiatric status of the respondents. Therefore, the inclusion of the method factor substantially improves the fit of the three-factor correlated model in both samples, and at least four of the reversed scored items loaded more heavily in the method factor than they did in their original factors. Nonetheless, while the method factor improved the fit indices of the model, it is evident that two of the five items were not significantly loaded in this factor and that some of the items exhibited extremely poor loadings. These findings are similar to the findings of Preece et al. (2018), who suggest that the EOT subscale score should not be used due to the content validity of several items and their low internal reliability.

Furthermore, the results associated with higher-order models, with three first-order factors and one global factor, and method factor and covariances revealed better fitting indices than did previously tested models. However, based on the unique omega estimates, the results suggest that the scores associated with the EOT and DDF subscales suffer from low levels of unique internal consistency reliability for both samples. A thorough analysis showed that the mean loading of the DDF latent factor for the clinical sample in the three-factor model with covariances and the method factor was 0.63, with a range of (0.25–0.88), which was higher than the mean factor loading of 0.45 with a range of (0.18–0.62) for the three-factor higher-order model with method and covariances. These findings suggest instability and an underestimation of the variance that each DDF item shared with the GF factor. A similar observation was noted in the student sample.

Conversely, the results associated with the nested factor model exhibited a slightly better fit than did the results of the corresponding higher-order model for the student sample. Specifically, a CFI difference of 0.033, TLI difference of 0.042, and SRMR difference of 0.014 for the student population. On the other hand, the nested models were not identified in the clinical sample due to the low number of respondents. In this regard, even though the nested models were more complex and there were fewer degrees of freedom, the nested factor model provided an improved solution for model fit for a student sample, a finding that is similar to Gignac et al. (2007).

Furthermore, the final nested model in this investigation tested the negatively keyed items using the method factor and incorporated the link among the three latent factors. The addition of the covariances was guided by the theoretical considerations provided by Gignac et al. (2007) and Preece et al. (2018). As evidenced, the correlation between DIF and DDF was extremely high, suggesting that individuals who have problems identifying emotions also experience difficulty describing emotions. These findings provide a valuable substrate to the debate on the impact of the cultural context in which alexithymia is studied, specifically by considering the Kosovar context, which is a nonwestern environment. Previous studies indicate that people from nonwestern countries having difficulty linguistically naming their feelings, cognitively explaining how they are feeling, and verbally discussing their emotions and feelings, preferring, instead, to avoid or ignore their feelings and emotions (Taylor et al., 2003). This is specifically true when considering that alexithymia is strongly associated with western values, which regard awareness of emotions and the ability to express one’s emotions as critical characteristics of the self, while nonwestern cultures emphasize the importance of emotional restraint and social cues that enable harmonious social relations (Dere et al., 2012). Accordingly, studies that have adopted a cross-cultural perspective on alexithymia have found cultural variations in emotional processes account for the general differences among cultures in the current research on alexithymia (Ryder et al., 2008; Zhu et al., 2007). For example, findings from the study conducted by Lo (2014) indicate that Chinese Canadians score significantly higher in EOT than do Euro Canadians and that these differences are explained by the cultural context as the Chinese culture places greater emphasis on social relationships and interpersonal harmony. In this regard, Dere et al. (2012) suggest that differences in alexithymia are the result of a cultural emphasis on emotions rather than on deficits in emotional processing. Furthermore, similar to Asian cultures, Kosovar Albanians emphasize emotional restraint and EOT. Regardless of the tremendous social and cultural changes the Kosovar society has experienced since the war in 1999, it remains a society dominated by patriarchal values (Kelmendi, 2015), which, in general, reinforce and promote power, control, rationality, and emotional restraint rather than emotional expressiveness.
Therefore, the eventual blend of DIF and DDF and a good fit may be better explained by Albanian’s cultural values rather than by the biomedical and metacognitive factors found in other cultures (Dere et al., 2012).

We note that the three-factor model with covariances and the method factor identified in this investigation provides the most parsimonious model for the clinical sample, as demonstrated by the acceptable estimates of the unique internal consistency reliability when measured by omega. As demonstrated, the latent factor DIF has an omega of 0.86, the DDF has an estimated unique variance of 0.93, the EOT is 0.79, and the NEG has an omega estimate of 0.95. However, the three-factor nested model with covariances and method factor is perhaps the most parsimonious solution for the student sample, where the unique variances estimated by omega are 0.79 for the DIF, 0.77 for the DDF, 0.25 for the EOT, 0.53 for the NEG, and 1.17 for the GF. While it is noted that the EOT exhibits the weakest internal consistency, the inclusion of the method factor with negatively keyed items improves the model fit and resolves the contamination of this subscale. Moreover, the internal consistency of this subscale reaches an acceptable range, indicating that only item 15 did not significantly load in this latent factor. One of the possible explanations for this is related to the double-barreled nature of the questions that do not follow the same logic as the other questions on this scale, which include negative statements similar to the explanation provided by Gignac et al. (2007).

These mixed findings in our study regarding the factor structure may also be related to the fact that the items on the TAS-20 may be interpreted differently by the clinical and nonclinical samples. However, it is also possible that the characteristics of the respective samples, for example, the student sample is highly homogeneous while the clinical sample was heterogeneous with respect to age, diagnosis and education, influence the fit in favor of different fitting models. Not surprisingly, mixed findings regarding the different factorial structures of the TAS-20 models in the clinical and nonclinical samples were identified by several authors (Haviland & Reise, 1996; Müller et al., 2003; Zhu et al., 2007).

Another explanation for the mixed findings obtained with the TAS-20 may be due to genuine cultural differences, especially when considering that problems in transposing psychological constructs to other cultures and other social contexts are well known and not uncommon due to the conceptual nonequivalence across cultures (Hui & Triandis, 1985). Another possible reason may be because some items may not exhibit the language sensitivity in our culture necessary to obtain a three-factor solution. However, this conclusion may be premature as failure to replicate the factors could be due to poor content validity of the EOT items and/or the impact of the method factor used in factor analysis studies (Gignac et al., 2007; Preece et al., 2018). Another cultural factor to consider, however, is the previous finding and explanation that in all cultures other than English speaking cultures, the EOT lacks internal reliability, a result that may be due to cultural differences or response bias to several of the negatively worded items on the EOT (Taylor et al., 2003).

Conclusion

The results of this study provide empirical support for the multidimensionality of alexithymia, specifically, by validating the three-factor correlated model for the clinical population and the three-factor nested model for the student population. Although the three-factor correlated model with covariances and method factor appears to be reasonably replicated in our data from the clinical sample and student sample, that is, with acceptable goodness of fit via a few minor freed parameters, the failure to appropriately confirm the three-factor nested model with covariances and method factor in the clinical sample can promote erroneous clinical reasoning without further clinical exploration. Accordingly, the questionnaire of the TAS-20 in Albanian requires further research to understand the psychometric features of the instrument.

Hence, we conclude that the TAS-20 in the Albanian language is appropriate for research purposes. Furthermore, we support the growing evidence that the TAS-20 exhibits clinical relevance in the study of alexithymia. However, we recommend caution with respect to the application of the scale in Albanian as a total measure of alexithymia, unless the TAS-20 is used in clinical research or clinical practice in combination with other instruments. Further investigation, specifically among clinical samples, is highly encouraged. Finally, the findings from this study highlight the importance of understanding and clarifying the differences between cultures when examining the psychological concepts, while also being cognizant of the concept of universality when studying such theories.

Limitations

One limitation of our study was that the participants in the nonclinical sample consisted of only one age cohort and females were overrepresented. Furthermore, due to the small number of respondents in the clinical sample, the factorial invariance between samples could not be computed. The results of this study are based on university undergraduates who are a relatively homogeneous group. It would be desirable to further examine the factor structure of the TAS-20 in Albanian, with samples comprised of different age groups, different socioeconomic backgrounds, and varied geographical regions and countries where the Albanian language is spoken in different dialects. In the clinical sample, as males were overrepresented, gender balance was not achieved in this sample. Furthermore, the clinical sample was overrepresented by
individuals with substance abuse issues, which means the heterogeneity of the sample with respect to psychopathology was also not well balanced. However, after refining the measurement of alexithymia on the TAS-20 in Albanian, other clinical samples should be included.

**Declaration of Conflicting Interests**
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**
The author(s) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iDs**
Aliriza Arenliu https://orcid.org/0000-0002-7016-4188
Kaltrina Kelmendi https://orcid.org/0000-0002-6619-1415

**References**
Bach, M., Bach, D., & Zwaan, M. D. (1996). Independence of alexithymia and somatization: A factor analytic study. *Psychosomatics, 37*, 451–458.
Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale: I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research, 38*, 23–32.
Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin, 107*, 238–246.
Berthoz, S., Consoli, S., Perez-Diaz, F., & Jouvent, R. (1999). Alexithymia and anxiety: Compounded relationships? A psychometric study. *European Psychiatry, 14*(7), 372–378.
Bressi, C., Taylor, G., Parker, J., Bressi, S., Brambilla, V., Aguglia, E., . . . Invernizzi, G. (1996). Cross validation of the factor structure of the 20-item Toronto Alexithymia Scale: An Italian multicenter study. *Journal of Psychosomatic Research, 41*, 551–559.
Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press.
Cleland, C., Magura, S., Foote, J., Rosenblum, A., & Kosanke, N. (2005). Psychometric properties of the Toronto Alexithymia Scale (TAS-20) for substance users. *Journal of Psychosomatic Research, 58*, 299–306.
Cserjési, R., Luminet, O., & Lénárd, L. (2007). Reliability and factor validity of the Hungarian translation of the Toronto Alexithymia Scale in undergraduate student samples. *Magyar Pszichológiai Szemle, 62*, 355–368.
Danish Refugee Council. (2006). *Long-term sequels of War, social functioning and mental health in Kosovo*. https://www.ecoi.net/en/file/local/1308565/432_1170782377_final-quantitative-research-report.pdf
De Gucht, V., Fontaine, J., & Fischler, B. (2004). Temporal stability and differential relationships with neuroticism and extraversion of the three subscales of the 20-item Toronto Alexithymia Scale in clinical and nonclinical samples. *Journal of Psychosomatic Research, 57*(1), 25–33.
Dere, J., Falk, C. F., & Ryder, A. G. (2012). Unpacking cultural differences in alexithymia: The role of cultural values among Euro-Canadian and Chinese-Canadian students. *Journal of Cross-Cultural Psychology, 43*, 1297–1312.
Erni, T., Lötscher, B., & Modestin, J. (1997). Two-factor solution of the 20-item Toronto Alexithymia Scale confirmed. *Psychopathology, 30*, 335–340.
Frewen, P. A., Pain, C., Dozois, D. J., & Lanius, R. A. (2006). Alexithymia in PTSD. *Annals of the New York Academy of Sciences, 1071*, 397–400.
Fukunishi, I., Nakagawa, T., Nakamura, H., Kikuchi, M., & Takubo, M. (1997). Is alexithymia a culture-bound construct? Validity and reliability of the Japanese versions of the 20-item Toronto Alexithymia Scale and modified Beth Israel Hospital Psychosomatic Questionnaire. *Psychological Reports, 80*(3), 787–799.
Gignac, G. E., Palmer, B. R., & Stough, C. (2007). A confirmatory factor analytic investigation of the TAS–20: Corroboration of a five-factor model and suggestions for improvement. *Journal of Personality Assessment, 89*(3), 247–257.
Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variables models. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation modeling: Present and future—A Festschrift in honor of Karl Jöreskog* (pp. 195–216). Lincolnwood, IL: ScientificSoftware International, In.
Güleç, H., Köse, S., Güleç, M. Y., Citak, S., Evren, C., Borckardt, J., & Sayar, K. (2009). Reliability and factorial validity of the Turkish version of the 20-item Toronto Alexithymia Scale (TAS-20). *Bulletin of Clinical Psychopharmacology, 19*, 214–220.
Haviland, M. G., & Reise, S. P. (1996). Structure of the twenty-item Toronto Alexithymia Scale. *Journal of Personality Assessment, 66*, 116–125.
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16*, 131–152.
Joukamaa, M., Taanila, A., Miettunen, J., Kivimäki, M., Koskinen, M., & Veijola, J. (2007). Epidemiology of alexithymia among adolescents. *Journal of Psychosomatic Research, 63*, 373–376.
Kelmendi, K. (2015). Domestic violence against women in Kosovo: A qualitative study of women’s experiences. *Journal of interpersonal violence, 30*(4), 680–702.
Komaki, G., Maeda, M., Arimura, T., Nakata, A., Shinoda, H., & Ogata, I., et al. (2003). The reliability and factorial validity of the Japanese version of the 20-item Toronto Alexithymia Scale. *Journal of Psychosomatic Research, 55*, 143.
Kooiman, C. G., Spin hopping, P., & Trijsburg, R. W. (2002). The assessment of alexithymia: A critical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20. *Journal of Psychosomatic Research, 53*, 1083–1090.
Lee, Y. H., Rim, H. D., & Lee, J. Y. (1996). Development and validation of a Korean version of the 20-item Toronto Alexithymia Scale (TAS-20K). *Journal of Korean Neuropsychiatry Association, 35*, 888–899.
Li, S., Zhang, B., Guo, Y., & Zhang, J. (2015). The association between alexithymia as assessed by the 20-item Toronto Alexithymia Scale and psychological distress in a Chinese sample. *Journal of Psychiatry Research, 68*, 10–16.
Li, S., Zhang, B., Guo, Y., & Zhang, J. (2015). The association between alexithymia as assessed by the 20-item Toronto Alexithymia Scale and psychological distress in a Chinese sample. *Journal of Psychiatry Research, 68*, 10–16.
Alexithymia Scale and depression: A meta-analysis. *Psychiatry Research*, 227, 1–9.

Lo, C. (2014). Cultural values and alexithymia. *SAGE Open*, 4(4), 215824401.

Loas, G., Corcos, M., Stephan, P., Pellet, J., Bizouard, P., Venisse, J. L., ... Jeanmet, P. (2001). Factorial structure of the 20-item Toronto Alexithymia Scale: Confirmatory factorial analyses in nonclinical and clinical samples. *Journal of Psychosomatic Research*, 50, 255–261.

Loas, G., Otmani, O., Verrier, A., Fremaux, D., & Marchand, M. P. (1996). Factor analysis of the French version of the 20-item Toronto Alexithymia Scale. *Psychopathology*, 29, 139–144.

Loiselle, C. G., & Cossette, S. (2001). Cross-cultural validation of the Toronto Alexithymia Scale (TAS-20) in US and Peruvian populations. *Transcultural Psychiatry*, 38, 348–362.

Lundh, L. G., & Broman, J. E. (2006). Alexithymia and insomnia. *Personality and Individual Differences*, 40(8), 1615–1624.

Martínez Sánchez, F. (1996). The Spanish version of the Toronto Alexithymia Scale (TAS-20). *Clínica y Salud*, 7(1), 19–32.

Meganck, R., Markey, S., & Vanheule, S. (2012). Reliability and factor structure in an adolescent sample of the Dutch 20-item Toronto Alexithymia Scale. *Psychological Reports*, 111, 393–404.

Meganck, R., Vanheule, S., & Desmet, M. (2008). Factorial validity and measurement invariance of the 20-item Toronto Alexithymia Scale in clinical and nonclinical samples. *Assessment*, 15(1), 36–47.

Müller, J., Bühner, M., & Ellgring, H. (2003). Is there a reliable factorial structure in the 20-item Toronto Alexithymia Scale? A comparison of factor models in clinical and normal adult samples. *Journal of Psychosomatic Research*, 55, 561–568.

Nemiah, J. C., Freyberger, H., & Sifneos, P. E. (1976). Alexithymia: A view of the psychosomatic process. In O. W. Hill (Ed.), *Modern trends in psychosomatic medicine* (Vol. 3, pp. 430–439). Butterworths.

Nowakowski, M. E., McFarlane, T., & Cassin, S. (2013). Alexithymia and eating disorders: A critical review of the literature. *Journal of Eating Disorders*, 1, 1–21.

Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods*, 14(3), 548–570.

Pandey, R., Mandal, M. K., Taylor, G. J., & Parker, J. D. (1996). Cross-cultural alexithymia: Development and validation of a Hindi translation of the 20-item Toronto Alexithymia Scale. *Journal of Clinical Psychology*, 52, 173–176.

Parker, J. D. A., Shanghessy, P. A., Wood, L. M., Majeski, S. A., & Eastabrook, J. M. (2005). Cross-cultural alexithymia validity of the 20-item Toronto Alexithymia Scale in North American aboriginal populations. *Journal of Psychosomatic Research*, 58, 81–83.

Picardi, A., Fagnani, C., Gigantesco, A., Toccarcel, V., Lega, I., & Stazi, M. A. (2011). Genetic influences on alexithymia and their relationship with depressive symptoms. *Journal of Psychosomatic Research*, 71(4), 256–263.

Pompili, M., Shrivastava, A., Serafini, G., Innamorati, M., Millesi, M., Erbuto, D., ... Girardi, P. (2013). Bereavement after the suicide of a significant other. *Indian Journal of Psychiatry*, 55(3), 256–263.
Vorst, H. C., & Bermond, B. (2001). Validity and reliability of the Bermond–Vorst alexithymia questionnaire. *Personality and Individual Differences, 30*(3), 413–434.

Westwood, H., Kerr-Gaffney, J., Stahl, D., & Tchanturia, K. (2017). Alexithymia in eating disorders: Systematic review and meta-analyses of studies using the Toronto Alexithymia Scale. *Journal of Psychosomatic Research, 99*, 66–81.

Zhu, X., Yi, J., Yao, S., Ryder, A., Taylor, G., & Bagby, M. (2007). Cross-cultural validation of a Chinese translation of the 20-item Toronto Alexithymia Scale. *Comprehensive Psychiatry, 48*, 489.