Validation of a Greek adaptation of the 20-item Toronto Alexithymia Scale

Ioannis Tsaousis,† Graeme Taylor, Lena Quilty, Stelios Georgiades, Marios Stavrogiannopoulos, R. Michael Bagby

**Background:** The purpose of the current investigation was (1) to test whether the 3-factor structure of the 20-item Toronto Alexithymia Scale (TAS-20) corresponding to the theoretical conceptualization of the alexithymia construct could be recovered in a Greek translation of the scale (the TAS-20-G), (2) to assess if a 3-factor structure provides a better fit to the TAS-20-G compared with the recently proposed alternative factor structures, and (3) to evaluate the internal reliability of the TAS-20-G.

**Methods:** The English version of the TAS-20 was translated into Greek and then back-translated and modified until cross-language equivalence was established. The Greek version was then administered to 340 university students. Confirmatory factor analyses were conducted, and 4 different factor structure models were compared. Internal consistency and item-to-scale homogeneity of the TAS-20-G and its factor scales were also evaluated.

**Results:** The 3-factor model provided a good fit to the data and proved superior to alternative 1-, 2-, and 4-factor models. Apart from a coefficient \( \alpha \) below the recommended range for the externally oriented thinking factor, the TAS-20-G and its factor scales demonstrated adequate internal consistency and homogeneity.

**Conclusion:** The TAS-20-G is a valid and reliable measure of alexithymia in university students and may be suitable for investigations of alexithymia in other Greek-speaking population samples.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The term “alexithymia” has its linguistic roots in the Greek language; coined by Sifneos [1], this synthetic Greek term literally means having no words for emotions. The construct of alexithymia was formulated by Nemiah and Sifneos [2,3] to encompass difficulties some people have in identifying and describing their own feelings and emotions as well as an impoverished fantasy life and an externally oriented mode of thinking. Rather than being a psychiatric diagnosis or a categorical phenomenon, alexithymia is conceptualized as a dimensional personality trait [4]; a taxometric investigation has provided support for this dimensional view of the construct [5].

A number of investigators have reported associations between alexithymia and a variety of medical and psychiatric disorders, including substance use disorders, eating disorders, posttraumatic stress disorder, functional gastrointestinal disorders, and essential hypertension [4,6]. There is also accumulating evidence that alexithymia influences patients’ responses to some medical and psychotherapeutic treatments and may even predict treatment response. Porcelli et al [7], for example, reported that alexithymia was a significant predictor of treatment outcome for patients with functional gastrointestinal disorder. Clinical researchers have also found alexithymia to be associated with a less favorable

---

* Corresponding author. Tel.: +30 210 8062110; fax: +30 210 6143435. E-mail address: tsousis@psy.soc.uoc.gr (I. Tsaousis).
response to psychodynamic psychotherapy [8] and a higher level of residual symptoms than reported by nonalexithymic patients even when the psychotherapy achieves significant symptom reduction [9].

Bagby and colleagues developed the 20-item Toronto Alexithymia Scale (TAS-20) [10,11], which has become the most widely and frequently used instrument to assess the alexithymia construct [12]. Based on factor analytic work [13], the 20 items that compose the TAS-20 are organized into 3 factor scales—Difficulty Identifying Feelings (DIF), Difficulty Describing Feelings (DDF), and Externally Oriented Thinking (EOT). The DIF factor scale assesses difficulties identifying feelings and distinguishing them from the somatic sensations that accompany emotional arousal; DDF assesses difficulties describing feelings to other people; and EOT assesses EOT. There is evidence that this factor also indirectly assesses fantasy and other imaginal processes [11,14,15].

Numerous investigators have examined the latent structure of the TAS-20. Using confirmatory factor analysis (CFA), the Toronto team of researchers demonstrated that the 3-factor structure recovered initially in the derivation sample [10] was replicable in both clinical and nonclinical English-speaking populations [16,17]. This 3-factor model has also been replicated with translations of the TAS-20 administered to nonclinical and clinical samples in many different languages including Chinese [18], Danish [19], Dutch [19], Finnish [20], French [21], German [16,22], Hebrew [19], Hungarian [23], Hindi [24], Farsi [25], Italian [26], Japanese [19], Korean [27], Lithuanian [28], Norwegian [19], Portuguese [28], Spanish [29], and Swedish [30].

Although the 3-factor model, identified initially in the English-speaking derivation sample, has been replicated in a variety of samples speaking different languages from diverse cultures (eg, Asian, European, and Middle-Eastern), some investigators have reported that alternative 2- and 4-factor models provide better statistical fit than the original 3-factor structure. For example, Loas et al [31] administered a French translation of the TAS-20 to a university student sample and obtained a 2-factor solution, with the DIF and the DDF items constituting a single factor and the EOT items comprising a second factor. A similar 2-factor solution was obtained by Erni et al [32], who had a sample of medical students complete a German translation of the TAS-20. Both of these studies, however, used exploratory factor analysis rather than CFA, to explore the factor structure of the translated versions of the TAS-20. When CFA was applied to the French data, a 3-factor model provided a better fit to the data than did the 2-factor model [33]. More recently, Müller et al [34] compared 5 different models with 1 to 4 factors in clinical and nonclinical (ie, university students) samples in Germany and found that a 4-factor model, in which the EOT items were divided into 2 distinct factors—“Pragmatic Thinking” and “Lack of subjective significance or importance of emotions”—provided a better fit to the data compared with 1-, 2- and 3-factor models. When the same models were compared in the validation study of the Chinese translation of the TAS-20, the 4-factor model provided a better fit than did the standard 3-factor model in a student sample, but not in a clinical sample [18].

In a review of 18 different language versions of the TAS-20, Taylor et al [19] refer to an unpublished Greek version of the scale that was translated by T. Anagnostopoulou and evaluated with a sample that comprised 104 normal adults and 96 asthmatic patients. Although results from CFA revealed that the 3-factor structure provided a “marginally” acceptable fit ($\chi^2/df$ ratio, 1.68; Goodness-of-Fit Index [GFI], 0.88; Adjusted Goodness of Fit Index [AGFI], 0.85; Standardized Root Mean Square Residual [SRMR], 0.07), the goodness-of-fit indices did not include an incremental fit index, which is now regarded as an essential index in assessing goodness-of-fit. Moreover, the estimates of internal reliability for the DDF and EOT factor scales were below the generally acceptable standard. The mixed clinical and nonclinical sample used in that study and some disagreements over the way some items had been translated into Greek prompted us to develop a new Greek translation of the TAS-20 (the TAS-20-G) and to examine whether the theoretical structure that underlies the factor structure of the original TAS-20 could be recovered in this Greek adaptation of the scale. Furthermore, we were interested in investigating whether alternative factor structures, such as a 2-factor model [31,32] or a 4-factor model [34], represent the latent structure of the TAS-20-G items better than the original 3-factor structure in a sample of Greek-speaking university students. Finally, we were interested in evaluating the internal reliability of this latest Greek translation of the scale.

2. Methods

2.1. Participants

The sample consisted of 340 undergraduate and postgraduate psychology students (91 men, 249 women) from the University of Crete. The mean (SD) age of the sample was 22.90 (6.35) years (range, 18-56 years).

2.2. Instrument

The TAS-20 is a 20-item self-report instrument with each item rated on a 5-point Likert scale ranging from 1 (strongly agree) to 5 (strongly disagree); 5 of the items are negatively keyed. Total scores range between 20 and 100, with higher scores indicating higher degrees of alexithymia.

The English version of the TAS-20 was translated into Greek following the rules suggested by Brislin [35]. First, a Greek-English bilingual psychologist translated the original English version of the scale into Greek. Next, from this translation, a different Greek-English bilingual psychologist translated it back into English. Finally, discrepancies emerging from this back-translated version were discussed,
and adjustments were made by the study authors to form the TAS-20-G.

2.3. Procedure

The participants completed the TAS-20-G as a partial fulfillment for a psychologic assessment course. Before administration of the questionnaire, all participants provided written consent after they were informed that their test results would be kept confidential and that they could withdraw from the study at any time.

2.4. Statistical analysis

A series of CFA models were specified and estimated using AMOS 6.0 (Analysis of Moment Structures) software [36]. In CFA, an a priori model is fit to the data. Because the expected value of the noncentrality parameter is unknown, several fit indices were estimated (see Griffin [37]). Six fit indices were examined for each of the 4 models tested: the χ²/df, the GFI, the Root Mean Square Error of Approximation (RMSEA), and the SRMR indices were selected as measures of absolute fit; the Comparative Fit Index (CFI) and the Non-Normed Fit Index (NNFI; also known as the Tucker-Lewis Index) were selected as measures of incremental fit. Generally, a χ²/df value less than 3.0 [38,39] and GFI, CFI, and NNFI values greater than 0.90 [40,41] indicate a model with a good fit; CFI and NNFI values close to or greater than 0.95 indicate an excellent fit [42]. For the SRMR, a cutoff value close to 0.08 or less is recommended; RMSEA values less than 0.05 indicate good fit, and values as high as 0.08 represent reasonable errors of approximation in the population [41,42]. Finally, the Akaike Information Criterion (AIC) [43] was selected as a criterion of model selection. Because there is no generally accepted cutoff value for the AIC criterion, lower values are interpreted to indicate a better model fit.

With regard to the estimation procedures used to assess each model, several methods are available. The popular maximum likelihood and generalized least squares methods are not robust in the presence of multivariate nonnormality. In these cases, the more robust unweighted least squares or the asymptotic distribution-free methods are preferred [42]. To test multivariate normality of the data, Mardia’s [44] multivariate skewness and kurtosis statistics were calculated via the SPSS (SPSS Inc., Chicago, IL) macro developed by DeCarlo [45].

The 4 models we compared were as follows:

(a) Model I is a 1-factor model.
(b) Model II specifies an oblique 2-factor model in which factor 1 is DIF and DDF, and factor 2 is EOT, as suggested by Erni et al [32].
(c) Model III represents the standard oblique 3-factor theoretical model as described by the developers of the TAS-20 [10].
(d) Model IV specifies an oblique 4-factor model with DIF and DDF as 2 separate factors and with EOT divided into 2 distinct factors—“Pragmatic Thinking” and “Lack of subjective significance or importance of emotions”—as suggested by Müller et al. [34].

We calculated Cronbach α coefficients and mean interitem correlations (MICs) for the total scale and for each factor scale to evaluate internal reliability and item-to-item homogeneity. An acceptable standard of higher than .70 was set for α and an optimal range of 0.20 to 0.40 for the mean interitem r [39,46]. Possible sex differences on the TAS-20-G total scale and the 3-factor scales were examined using t tests.

3. Results

3.1. Confirmatory factor analysis

Because most of the skewness and kurtosis values were significant (P < .05), and the total Mardia’s coefficient was above the minimum accepted value of 1.96 (ie, 78.03), the hypothesis of multivariate normality was rejected and the unweighted least square estimation method was applied.

Fit indices for each model are displayed in Table 1. The 1-factor model provides a poor fit to the data as, apart from the GFI, the fit indices do not meet the criterion levels. In contrast, the 2-, 3- and 4-factor models provide a good fit to the data by most indices. Model III (the standard model) provided the best fit to the data, as demonstrated by the AIC index (AIC, 469.93). This model is displayed in Fig. 1.

3.2. Means and SDs for the TAS-20-G and its factor scales

The means and SDs for the TAS-20-G total scale and the 3-factor scales are shown for the total sample and also separately for men and women in Table 2. There were significant sex differences in mean TAS-20 total scores (t<sub>249</sub> = 2.70, P = .007, d = 0.33), as well as in the DDF (t<sub>249</sub> = 2.26, P = .024, d = 0.28) and EOT (t<sub>249</sub> = 3.66, P < .001, d = 0.46) factor scales. In all 3 instances, men had higher scores than did women. There was no significant sex difference for the DIF factor scale (t<sub>249</sub> = 0.47, ns). Correlations between age and TAS-20 total score, DIF, DDF, and EOT were nonsignificant.

3.3. Reliability of the TAS-20-G

As shown in Table 2, the Cronbach α coefficients were .79 for the TAS-20-G total scale, .74 for the DIF factor scale, and .79 for the DDF factor scale, indicating adequate internal consistency. For the EOT factor scale, the coefficient α was smaller at .58 and therefore below the recommended standard. The MIC coefficients for the TAS-20-G total scale and the 3-factor scales are also shown in Table 2. Although the MICs for the TAS-20-G total scale and the DDF and EOT factor scales are slightly outside the optimal recommended range of 0.20 to 0.40, they are between 0.10
4. Discussion

The results of the current investigation provide empirical evidence for the validity of the 3-factor structure of the TAS-20-G and for the superiority of this 3-factor model versus alternative models reported in the literature with other language versions of the TAS-20 (eg, 1-, 2-, and 4-factor models). We also found support for the reliability of the total TAS-20-G and its DIF and DDF factor scales.

Results from the nonparametric CFAs indicated that the 2-, 3-, and 4-factor models all showed at least a good fit; the 1-factor model, in contrast, was a poor fit across almost all of the fit indices; thus, alexithymia is best conceptualized as a multidimensional construct, at least as measured by the TAS-20-G. Based on the AIC statistic, we also can conclude that the “original” 3-factor model extracted in the derivation sample and that maps onto and reflects the theoretical model of the alexithymia construct [4] was superior to the 2- and 4-factor models. The results from this study are in line with numerous previous investigations using translated versions of the TAS-20 in which the original 3-factor structure was found to be valid and replicable, possibly suggesting a universal structure for the construct of alexithymia [19,47].

The internal reliability coefficients obtained for the TAS-20-G in the current investigation are close in magnitude to those obtained in the original TAS-20 studies [10,17] and to those reported in many other studies with different language versions of the scale [18,19,26,30,33]. The lower internal consistency coefficient found for the EOT factor scale relative to the other factor scales is consistent with the results from several other studies [18,26,30,34,48], which have led to suggestions for revision of the items that constitute this particular factor scale. Pervious studies have suggested that the EOT factor could be decomposed into 2 latent components (ie, pragmatic thinking and lack of importance of emotions) [34]. As Bagby et al [47] have argued, the low reliability of the EOT factor might be due to the fact that items assessing a lack of importance of emotions may not adequately reflect the EOT facet of the alexithymia construct, as initially conceptualized. The items comprising “pragmatic thinking” seem to more closely reflect pensée

<table>
<thead>
<tr>
<th>Model</th>
<th>χ²</th>
<th>df</th>
<th>P</th>
<th>χ²/df</th>
<th>GFI</th>
<th>CFI</th>
<th>NNFI</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I (1-factor)</td>
<td>554.05</td>
<td>170</td>
<td>&lt;.01</td>
<td>3.26</td>
<td>0.93</td>
<td>0.88</td>
<td>0.88</td>
<td>0.10</td>
<td>0.09</td>
<td>634.05</td>
</tr>
<tr>
<td>Model II (2-factors)</td>
<td>437.20</td>
<td>169</td>
<td>&lt;.01</td>
<td>2.59</td>
<td>0.94</td>
<td>0.92</td>
<td>0.92</td>
<td>0.09</td>
<td>0.08</td>
<td>521.20</td>
</tr>
<tr>
<td>Model III (3-factors)</td>
<td>383.93</td>
<td>167</td>
<td>&lt;.01</td>
<td>2.30</td>
<td>0.95</td>
<td>0.93</td>
<td>0.93</td>
<td>0.08</td>
<td>0.07</td>
<td>469.93</td>
</tr>
<tr>
<td>Model IV (4-factors)</td>
<td>381.43</td>
<td>164</td>
<td>&lt;.01</td>
<td>2.33</td>
<td>0.95</td>
<td>0.93</td>
<td>0.93</td>
<td>0.08</td>
<td>0.07</td>
<td>473.43</td>
</tr>
</tbody>
</table>

Fig. 1. Hypothesized model of the factor structure of the TAS-20-G.

Table 1

Goodness-of-fit indices of the hypothesized latent-factor models of the TAS-20-G

Table 2

Descriptive statistics, MICs, and Cronbach α coefficients for the TAS-20-G
opératoire (operative thinking), which Nemiah and Sifneos [3] considered a central feature of the construct. Because only 3 items assess pragmatic thinking (items 5, 8, and 20), revision of the EOT scale would require the evaluation and addition of several new items to improve the assessment of this component of the construct.

One could also argue that the low internal reliability of the EOT factor precludes the validity of the 3-factor model of the TAS-20-G, or for that matter, other language versions of the TAS-20, and instead could suggest a 2-factor structure (DIF and DDF) by excluding the EOT factor. However, it is important to emphasize that the factor structure of a set of indicators and the internal consistency of the factor scales consisting of the summation of a set of indicators within those factors are distinct (albeit related) issues. Externally oriented thinking is a salient component of the alexithymia construct, and the results of the CFA clearly indicate relative superiority of the 3-factor structure. Furthermore, the results from the interitem coefficient (MIC), which is a more appropriate statistic than $\alpha$ coefficient for the investigation of the cohesiveness of a scale, showed that the EOT factor can be considered a homogeneous scale.

It is important to emphasize that the low internal reliability of the EOT factor does not preclude the validity of the 3-factor model of the TAS-20-G, or for that matter, other language versions of the TAS-20. The factor structure of a set of indicators and the internal consistency of the factor scales consisting of the summation of a set of indicators within those factors are distinct (albeit related) issues. The results suggest that in addition to the total scale score, only 2 reliable factor scales (DIF and DDF) can be derived from the scale in its current form; however, this does not imply a need to remove EOT from the content domain and create a 2-factor (DIF and DDF) structure. Externally oriented thinking is a salient component of the alexithymia construct and the results of the CFA clearly indicate relative superiority of the 3-factor structure.

An important limitation of the present study is that the results are based on a sample of university students. It is possible that the factor structure of the TAS-20 may differ between clinical and nonclinical samples. For example, in evaluating a German version of the TAS-20, Müller et al [34] found that patient data showed a better fit to tested models than data from normal adults. In contrast, Zhu et al [18] found that Chinese student data provided a better fit to tested models than Chinese patient data; although 3- and 4-factor models showed a good fit for both patients and students, as we noted earlier, the 4-factor model showed a significant improvement over the standard 3-factor model in the Chinese student sample. Thus, it would be desirable to further examine the factor structure of the TAS-20-G in a Greek community sample that is diverse in terms of age, educational level, and socioeconomic background, as well as in some clinical samples. Another limitation of our study is the sex imbalance of the sample, with only 26.8% of participants being men.

Further studies are needed to evaluate the convergent, discriminant, and concurrent validity of the TAS-20-G. The results from the current investigation, however, suggest that this Greek translation of the TAS-20 is a sufficiently valid and reliable instrument for use in both clinical and research contexts.

References


