ELSEVIER

Journal of Psychosomatic Research 53 (2002) 1083–1090

Journal of Psychosomatic Research

The assessment of alexithymia
A critical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20

C.G. Kooiman[a,*], P. Spinhoven[a,b], R.W. Trijsburg[c]

[a]*Department of Psychiatry, Leiden University Medical Center, B1P, 2300 RC Leiden, The Netherlands*
[b]*Division of Clinical and Health Psychology, Leiden University, Leiden, The Netherlands*
[c]*Department of Medical Psychology and Psychotherapy, Erasmus University, Rotterdam, The Netherlands*

Received 5 March 2001; accepted 21 November 2001

Abstract

Objective: The objectives were to give an overview of studies on the validity of the Toronto Alexithymia Scale (TAS-20) and to present data regarding the validity of the TAS-20. **Methods:** The literature on the psychometric properties of the TAS-20 was reviewed and a study was conducted of its psychometric properties in a sample of students and a sample of psychiatric outpatients using a statistical method allowing identification of a stable factor structure. **Results:** The review revealed that the majority of studies on the TAS-20 were conducted with nonpatient samples. The factorial validity and reliability of the dimensions 'identifying feelings' (DIF) and 'describing feelings' (DDF) could be replicated in many of these studies. However, in practically all studies the dimension 'externally oriented thinking' (EOT) appears to be unreliable. In addition, the presupposed fantasy aspect of the alexithymia construct is not included in the TAS-20. Although many studies were conducted on the construct validity of the TAS-20, no studies have been published on its criterion validity. Some studies show a different factor structure to exist in patient samples. This was confirmed in our own study in which the dimensions 'identifying feelings' and 'describing feelings' collapsed into one single subscale. As in other studies, the reliability of the dimension 'EOT' was low. **Conclusion:** The TAS-20 has some important shortcomings with respect to validity and reliability. For the assessment of alexithymia in empirical research, it is recommended to use the TAS-20 in combination with other instruments. We do not recommend the TAS-20 to be used in clinical practice.

Keywords: Alexithymia; Personality tests; Psychometrics; Psychology; Affective symptoms

Introduction

The concept of alexithymia arose in psychosomatic practice after 1960. At that time, patients were examined for possible psychological causes of certain physical conditions (e.g., Ref. [1]). When interviewing patients about their emotional functioning, clinicians noted that so-called psychosomatic patients frequently had difficulties in expressing their feelings. Sifneos [2] introduced the term 'alexithymia' to describe this phenomenon. Initially, it was thought that alexithymia predisposed the patient to physical diseases such as hyperthyroidism. More recently, it has been suggested that alexithymia may be a predisposing factor for various psychiatric problems such as somatically unexplained physical symptoms, eating disorders, and substance dependence [3].

Since the introduction of the concept of alexithymia, its definition has been refined [4]. Currently, one speaks of 'alexithymia' when the criteria listed in Table 1 are fulfilled. Various instruments are used to measure alexithymia. The Beth Israel Psychosomatic Questionnaire (BIQ) [3], the Alexithymia Provoked Response Questionnaire (APRQ) [5], and the Karolinska Psychodynamic Profile (KAPP) [6] are assessment scales to be used with an interview. The Scored Archetypal Test (SAT9) [7] and the Rorschach [8] are used as projective instruments to measure the degree of alexithymia. The Levels of Emotional Awareness Scale (LEAS) [9] is a self-report instru-

* Corresponding author. Tel.: +31-71-526-3785; fax: +31-71-524-8156.
E-mail address: cgkooiman@lumc.nl (C.G. Kooiman).

0022-3999/02/$ – see front matter
PII: S0022-3999(02)00348-3

Table 1
Descriptive definition of alexithymia

- Difficulties in distinguishing between different emotions and insufficient realisation that some physical sensations may be the manifestation of emotions.
- Difficulties in verbally expressing emotions.
- Limited imagination and fantasy life.
- Thought is directed at external reality and hardly or not at inner experience.

ment to elicit written emotional responses in response to evocative interpersonal vignettes. An alexithymia prototype has been defined with the California Q-set (CAQ) [10] with which the degree of alexithymia can be determined as assigned by the patient himself or by an observer. The Schalling–Sifneos Personality Scale (SSPS) [11], the MMPI alexithymia subscale [12], the Amsterdam Alexithymia Scale (AAS) [13], and the Toronto Alexithymia Scale (TAS-20) [14,15] are self-report questionnaires. Lastly, the Observer Alexithymia Scale (OAS) [16] is a self-report questionnaire to be completed by a subject's acquaintance or relative. However, many of these instruments have not been sufficiently investigated or show psychometric shortcomings [3,17].

In empirical studies, almost exclusively, the TAS-20 is used to assess alexithymia because of its supposed good reliability and construct validity [14,18]. The original version of the TAS-20 [14,15] has been translated into several European and Asian languages [3,19–27] indicating its worldwide use. We also notice that clinicians tend to assess alexithymia in their patients with the TAS-20 only. However, we have reason to believe that the TAS-20 has some shortcomings that may lead to a different evaluation of the instrument. We will first discuss the development of the TAS-20 and its psychometric qualities as reported in the literature. Subsequently, we present the results of a psychometric study with the Dutch translation of the TAS-20.

The TAS was developed by Taylor et al. [28]. Initially, the authors identified five different aspects of alexithymia: (1) difficulties in describing feelings; (2) difficulties in distinguishing between emotions and physical sensations; (3) lack of introspection; (4) social overadaptation; and (5) limited fantasy life and difficulties remembering dreams. On the basis of these aspects, Taylor et al. [28] formulated 41 items, of which 8 items were extracted from the SSPS [11], 4 items from the Interoceptive Awareness Subscale of the Eating Disorder Inventory [29], and 4 items from the Need for Cognition Scale [30]. Half of the items were phrased negatively, to compensate answering tendencies. The items were scored on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). In a psychometric study with students as subjects, 15 items had to be dropped due to low item–total or item–factor correlations. The resulting scale (TAS-26) consisted of 26 items representing four factors: (1) difficulties in identifying feelings and distinguishing between emotional and physical sensations (DIF), (2) difficulties in describing feelings (DDF), (3) diminished daydreaming (supposed to represent a limited fantasy life), and (4) externally oriented thinking (EOT).

In subsequent studies, the TAS-26 was found to have some psychometric shortcomings. Low item–total correlations and high correlations with social desirability were found for a large number of items measuring diminished daydreaming [14]. Furthermore, 'diminished daydreaming' correlated negatively with DDF [14] and EOT [31]. In a first attempt to improve the questionnaire, the Taylor et al. [32] reconstructed the TAS-26 into the TAS-R with 23 items. However, the TAS-R proved to have an instable factor structure with DIF and DDF items tending to collapse into one single component. Thus, the authors made a second attempt to improve the TAS-26. In this second revision, the items of the TAS-26 were supplemented with 17 new items (7 relating to fantasising, 5 relating to EOT, and 5 relating to DDF). Items with a correlation of $r \geq .20$ with social desirability [33], with a low item–total correlation, or a low item–factor correlation ($r < .20$) were to be excluded from the new questionnaire [14].

From a study with students as subjects, it appeared that only 3 of the 12 fantasy items were in accordance with these criteria. Therefore, the authors decided to exclude all items relating to fantasy. Due to low item–total or item–factor correlations, also seven of the other items were dropped [14]. Following factor analysis of the scores for the remaining items, another four items were excluded because their loading on the extracted factors was too low. The remaining 20 items were included in the TAS-20. The factor structure and internal consistency of the TAS-20 were replicated in a student sample and a sample of psychiatric outpatients. In this study [14], the TAS-20 was found to consist of three factors: DIF, DDF, and EOT). Except for the EOT factor, the internal consistency (Cronbach's α and average inter-item correlation) of the individual factors was good.

Using confirmatory factor analysis, the factor structure of the TAS-20 has been studied with German, Italian, Hindi, French, Swedish, and English versions of the instrument in a total of eight populations of adult nonpatients (mostly student populations) and in three clinical populations [19,20,22,23,26,34]. However, although all authors conclude that the original three-factor structure could be replicated satisfactorily, in all studies about half or more of the EOT items had very low factor loadings. The factor EOT seems also problematic from another perspective. In most studies [19,20,22,23,34–38], applying the factor structure described by Bagby et al. [14], the internal consistency of the EOT factor tends to be unsatisfactory low (range: .45 < Cronbach's α < .76).

Bagby et al. [39] reported that they replicated with exploratory factor analysis the three-factor structure of the TAS-20 in the Dutch general population. But again, four of the eight items of the factor EOT had low factor loadings

with a correspondingly low internal consistency for EOT (Cronbach's $\alpha = .54$). Erni et al. [27] carried out an exploratory factor analysis in a student population using the German version of the TAS-20. They found a two-factor solution for the scale in which the dimensions DIF and DDF formed one factor and EOT the other. Bach et al. [35] also studied the German translation of the TAS-20 with exploratory factor analysis. The factor structure they found in adult nonpatients corresponded closely to that described by Bagby et al. [14]. However, this finding could not be replicated in a group of psychosomatic inpatients. Using confirmative factor analysis of the English TAS-20 in a group of medical students and a group of hospitalised addiction patients, Haviland and Reise [40] were unable to identify the factor structure described by Bagby et al. [14]. When applying exploratory factor analysis, a three-factor structure was found in the student sample that was reasonably similar to the one described by Bagby et al. [14]. However, this factor structure was not retrieved in the patient sample in which the DIF and DDF dimensions again formed one single factor.

In most studies, the individual factors (DIF, DDF, and EOT) were found to correlate positively with one another [14,19,20,22,23,34,40], which has been interpreted as these factors being interrelated aspects of the alexithymia construct [14]. In particular, the DIF and DDF factors were found to correlate strongly ($r = .43 - .80$). The correlations of these two factors with EOT are smaller and more variable ($r = -.06 - .51$ and $r = -.03 - .59$, respectively). This is probably partly due to the low internal consistency of the EOT scale.

The test–retest reliability of the TAS-20 in English [14], German [19], Italian [20], and Hindi [22] appears to be good ($r = .71 - .86$). In three of these studies, the test–retest reliability was carried out in nonpatient samples [14,20,22] and in three of them there was a short time lag (1–3 weeks) [14,19,20]. In one study [22], the interval was 3 months.

As the psychometric properties of the TAS-20 have been studied chiefly in nonpatient samples, we conducted a study on the factor structure and internal consistency of the Dutch translation of the TAS-20 in a student as well as in a patient sample. The test–retest reliability was investigated with a time interval of 3 months in a patient sample to add some clinical relevant information to the existing literature. Moreover, because the factor structure of the TAS-20 may be dependent on the sample studied, we used a statistical method allowing the identification of a factor structure valid for different samples.

Method

Study groups

Psychology students at the University of Amsterdam and psychiatric outpatients at the Leiden University Medical Center participated in this study. The students were asked to complete the TAS-20 in the context of a study obligation. Five hundred sixty-two (86%) of the 654 students participated in the study and 519 (79%) students filled out the TAS-20 completely. The mean age of the student sample was 21.4 years (S.D. = 3.78) and 347 (67%) students were female. Age and gender did not differ between the nonparticipating and the participating students.

The patient group consisted of consecutive new patients aged between 18 and 65 years of age who were referred to the psychiatric outpatients' clinic over the course of 1 year and 4 months. Patients were excluded from the study if suffering from a neurological disorder or an organic psychosyndrome, mental retardation, a psychotic disorder, sensory disturbances, or insufficient command of the Dutch language. At the start of the study, the patients were asked informed consent. The patients completed the questionnaire after the first psychiatric interview. At retest 3 months later, the questionnaires were sent to the patients' home address with a return envelope. The patients were asked to complete the questionnaire without assistance from others.

After exclusion of 68 patients who met one or more exclusion criteria, the patient group consisted of 212 consecutive patients. Forty-five (21%) patients chose not to take part in the study (nonparticipants). Of the 167 participating patients, 159 filled out the TAS-20 completely.

The mean age of the patient sample was 39.4 years (S.D. = 11.8) and 93 (59%) patients were female. Twenty percent of the patients lived as a single. Twenty percent of the patients completed primary school, 40% lower secondary education, 25% middle secondary education, and 15% higher secondary education. Thirty-one percent of the patients had a blue-collar job, 54% a white-collar job, and 15% was unemployed. The main psychiatric diagnosis was somatoform disorder (29%), mood disorder (20%), anxiety disorder (13%), adjustment disorder (10%), any other mental disorder (Axis I, 18%), or no mental disorder (Axis I, 11%). There were more women ($\chi^2 = 5.85$, $P < .05$), and more unemployed and fewer patients with blue-collar jobs ($\chi^2 = 6.07$, $P < .05$) amongst the nonparticipants. The participants and nonparticipants did not differ in other demographic or clinical characteristics.

Instruments

The TAS-20 [14,15] was translated into Dutch with the permission of the authors and subsequently translated back into English by a native speaker. Following two rounds of translation, the definitive Dutch translation of the TAS-20 was established in consultation between Bagby and the third author (R.W.T.).

The psychiatric diagnosis in the patient group was established according to the Axis I criteria of the DSM-IV [41]. The treating psychiatrist recorded both demographic data [42,43] and the nature of any somatic condition.

Statistical analyses

Factor analysis was performed by an exploratory principal components analysis (PCA) on the correlation matrix of the 20 items of the TAS-20 with 1.0 in main diagonal. First, the appropriateness of a three-factor solution was determined on the basis of the sizes of the eigenvalues and the scree test [44] in all four samples separately: male students, female students, male patients, and female patients. The percentage of explained variance was subsequently compared to the amount of variance accounted for by the three dimensions originally proposed by Bagby by means of a (confirmatory) Multiple Group Method Analysis (MGM) [45], i.e., a matrix with binary elements only, such that each variable has a weight of one on one component (e.g., EOT), and zero on the other components (e.g., DIF and DDF). MGM analysis was followed by a varimax rotation of factor loadings, followed by an oblique one. This rotation procedure was followed because it yields optimal results in those cases where it is unlikely that an orthogonal component structure will be detected [46]. The appropriateness of the resulting factor solution was indicated by the following: (a) the amount of total variance accounted for by the a priori matrix in comparison to the amount of total variance accounted for by the exploratory three-factor PCAs in the four samples separately and (b) the number of subscales well accounted for, as evidenced by the substantial correlation ($>.40$) of each item with the presupposed corresponding component in each subsample.

In case no evidence for a three-factor solution would be found, the dimensional structure of the TAS-20 was further evaluated with a (exploratory) simultaneous component analysis (SCA) [47] in order to explore the optimal component weights for the four different samples. In SCA, component weights are defined as in PCA and all groups get the same weights. The component weights are calculated in such a way that the resulting components optimally describe the variables in all groups. The appropriateness of the resulting factor solution is evaluated in a similar way as in confirmatory PCA, i.e., by comparing the amount of variance explained by the components of the SCA with the variance explained by separate exploratory PCAs and by evaluating the correlation of items with the corresponding components.

Table 2
Eigenvalues in separate exploratory PCA

	Female students ($n=347$)	Male students ($n=172$)	Female patients ($n=93$)	Male patients ($n=66$)
Component 1	4.7	5.3	5.1	5.0
Component 2	2.6	2.2	2.2	2.8
Component 3	1.5	1.4	2.0	1.5
Component 4	1.2	1.3	1.3	1.5
Component 5	1.1	1.2	1.2	1.3
Component 6	0.9	1.0	1.1	1.1

Table 3
Explained variances of the MGM analysis and the separate PCA with three components

Sample	n	MGM-explained variance (%)	Separate PCA-explained variance (%)	d%
Female students	347	34.5	44.4	9.9
Male students	172	36.0	44.6	8.6
Female patients	93	39.2	46.3	7.1
Male patients	66	36.2	47.0	10.8

Scales were constructed by including items of the TAS-20 with a loading of at least .40 on one of the components and a difference in loading of at least .10 with the other component(s) in minimally three out of the four samples. The internal consistency of the scales was calculated with Cronbach's α. Finally, items were deleted from a scale if without the item concerned the Cronbach's α of the scale increased with .05 or more in minimally three out of the four separate populations.

Pearson's Product–Moment Correlation Coefficient was used to estimate in psychiatric outpatients the test–retest reliability of the alexithymia dimensions. A reliability of $r \geq .70$ is considered satisfactory.

Finally, one-way ANOVA with Tukey HSD post hoc analysis was applied to test the differences in means between the four groups.

Results

The factor structure of the TAS-20

On the basis of exploratory PCAs in the separate subgroups no evidence for a three-factor solution as supposed by Bagby was found. In all four subgroups, five components emerged with eigenvalues greater than one. The scree test for the plot of eigenvalues showed in three of the four subgroups a distinct break between the steep slope of the first two factors and the gradual trailing of the remaining factors (Table 2). Only in the group of female patients some support for a three-factor solution emerged. Moreover, the results of the confirmatory MGM showed that the three-factor structure is not an adequate fit of the data because of two reasons. First, the three-factor structure did not adequately account for the variance. The differences in explained variance between the amounts of total variance accounted for by the a priori matrix in comparison to the amount of total variance accounted for by the separate exploratory three-factor PCAs was large, varying in the four subgroups from 7.1% to 10.8% (Table 3). Secondly, inspection of the loadings indicated that in all four subgroups items of the original TAS-20 subscales for DIF and DDF loaded inconsistently on two components, while most of the items of the original subscale EOT consistently loaded on one separate component only.

Table 4
Explained variances of the SCA and the separate PCA with two components

Sample	*n*	SCA-explained variance (%)	Separate PCA-explained variance (%)	*d*%
Female students	347	36.8	36.9	0.1
Male students	172	37.2	37.3	0.1
Female patients	93	35.7	36.4	0.7
Male patients	66	38.8	39.4	0.6

Taken together, these results clearly indicated that a two- instead of a three-factor solution might be more adequate in accounting for the variance in TAS-20 scores. SCA with a two-component solution confirmed this assumption. First, the variances explained by the two components in the SCA were very comparable with the variance explained by the separate two-factor PCAs in each of the four subgroups. The differences in explained variances were less than 0.75%, indicating that the components in the SCA were explaining the variance of the variables nearly as well as the factors from the separate PCAs (Table 4). Secondly, most of the items had a substantial correlation greater than .40 with the corresponding component in each subgroup. Only three EOT items had insufficient factor loadings of less than .40 and one DDF item had a sufficient factor loading but a difference in loading on the two components of less than .10.

The loadings of the TAS-20 items on the two components are presented in Table 5. Eleven items of the original TAS-20 subscales DIF and DDF loaded on the first subscale (Difficulties in Identifying and Describing Feel-

Table 5
Factor loadings (DIDF/EOT), internal consistency (Cronbach's α), and mean (S.D.) scores of the TAS-20 on two components in the four samples

	Original scale	Female students ($n=347$)	Male students ($n=172$)	Female patients ($n=93$)	Male patients ($n=66$)
DIDF		DIDF/EOT	DIDF/EOT	DIDF/EOT	DIDF/EOT
1. I am often confused about what emotion I am feeling.	DIF	.67/−.00	.78/−.07	.82/.00	.77/.01
2. It is difficult for me to find the right words for my feelings.	DDF	.65/−.45	.70/−.40	.74/−.49	.79/−.30
3. I have physical sensations that even doctors do not understand.	DIF	.49/.07	.44/.02	.61/−.02	.47/−.30
4. I am able to describe my feelings easily.[a]	DDF	−.58/.46	−.62/.46	−.31/.53	−.46/.33
6. When I am upset, I do not know if I am sad, frightened, or angry.	DIF	.64/−.06	.72/−.26	.70/−.07	.59/−.02
7. I am often puzzled by sensations in my body.	DIF	.65/.01	.68/−.16	.69/−.23	.57/.18
9. I have feelings that I cannot quite identify.	DIF	.76/−.08	.77/−.11	.71/−.10	.73/.02
11. I find it hard to describe how I feel about people.	DDF	.46/−.23	.54/−.22	.48/−.26	.36/−.03
12. People tell me to describe my feelings more.	DDF	.50/−.40	.52/−.42	.44/−.44	.53/−.37
13. I do not know what's going on inside me.	DIF	.72/−.28	.66/−.22	.69/−.20	.74/−.14
14. I often do not know why I am angry.	DIF	.65/−.16	.68/−.28	.66/−.17	.56/−.15
Cronbach's α		.84	.86	.85	.83
Mean (S.D.) of scale		24.7 (6.8)	24.2 (7.4)	30.5 (9.4)	30.1 (8.6)
EOT					
5. I prefer to analyze problems rather than just describe them.[a]	EOT	−.02/.58	−.07/.56	.06/.51	−.07/.75
8. I prefer to just let things happen rather than to understand why they turned out that way.	EOT	.20/−.55	.11/−.49	.29/−.53	.21/−.61
10. Being in touch with emotions is essential.[a]	EOT	.06/.50	−.15/.47	−.21/.48	.36/.38
15. I prefer talking to people about their daily activities rather than their feelings.	EOT	.17/−.58	.37/−.65	.21/−.51	.39/−.57
19. I find examination of my feelings useful in solving personal problems.[a]	EOT	.01/.71	.09/.62	.12/.52	.42/.57
Cronbach's α		.62	.58	.44	.65
Mean (S.D.) of scale		10.7 (2.7)	11.2 (2.9)	11.5 (3.1)	12.2 (3.7)
TAS total (16 items)					
Cronbach's α		.81	.83	.82	.78
Mean (S.D.) of scale		35.4 (7.7)	35.4 (8.6)	42.0 (10.7)	42.3 (9.7)
Items with insufficient loading on scales					
16. I prefer to watch 'light' entertainment shows rather than psychological dramas.	EOT	.05/−.48	.03/−.39	.07/−.38	−.03/−.59
17. It is difficult for me to reveal my innermost feelings, even to close friends.	DDF	.42/−.48	.46/−.45	.42/−.35	.53/−.52
18. I can feel close to someone, even in moments of silence.[a]	EOT	−.06/.43	.02/.35	−.27/.44	−.12/.30
20. Looking for hidden meanings in movies or plays distracts from their enjoyment.	EOT	.01/−.31	.06/−.32	.12/−.23	.07/−.26

[a] Items are reverse scored.

ings; DIDF) and five items of the original TAS-20 subscale EOT were retained in the second subscale (EOT). In order to investigate the reliability of the two scales, Cronbach's α was calculated for each subsample. The reliability of the first scale is good, whereas the internal consistency of the second scale is unsatisfactory (Table 5). No items were deleted because of decreasing the internal consistency of one of the scales. In Table 6, the means and the Cronbach's α of the original TAS-20 scales are represented. Again, the internal consistency of the EOT scale is unsatisfactory.

The test–retest reliability of the TAS-20

Seventy patients (44 women and 26 men) completed the TAS-20 3 months after the first time. The average scores on the TAS-20 and its subscales for the patients who completed the TAS-20 a second time did not differ from those of the patients who did not participate in the retest. The test–retest reliability of the DIF, DDF, EOT, and the TAS-20 total score in this study are represented in Table 7.

Correlations between subscales

The Pearson's product moment correlations between DIF and DDF were substantial. The correlations for female and male students were respectively $r = .45$ ($P < .01$) and $r = .58$ ($P < .01$). For female and male patients, they were $r = .52$ ($P < .01$) and $r = .59$ ($P < .01$). The Pearson's correlations between DIF and EOT were low, being $r = .07$ (ns) and $r = 0{,}16$ ($P < .05$) for, respectively, female and male students, and $r = .20$ ($P = .05$) and $r = .05$ (ns) for, respectively, female and male patients. The correlations of DDF and EOT for female and male students were, respectively, $r = .30$ ($P < .01$) and $r = .26$ ($P < .01$), and for female and male patients, $r = .42$ ($P < .01$) and $r = .24$ ($P < .05$), respectively.

Table 6
Internal consistency and mean scores of the original TAS-20 scales in the four samples

	Female students ($n = 347$)	Male students ($n = 172$)	Female patients ($n = 93$)	Male patients ($n = 66$)
DIF				
Cronbach's α	.82	.83	.85	.79
Mean (S.D.) of scale	15.31 (4.80)	14.50 (4.97)	19.38 (7.11)	19.44 (6.20)
DDF				
Cronbach's α	.78	.77	.67	.72
Mean (S.D.) of scale	11.78 (3.76)	12.27 (4.04)	14.47 (4.24)	13.88 (4.26)
EOT				
Cronbach's α	.66	.60	.52	.65
Mean (S.D.) of scale	17.43 (3.82)	18.15 (4.06)	19.13 (4.41)	20.36 (5.08)
TAS-20 total				
Cronbach's α	.81	.82	.82	.79
Mean (S.D.) of scale	44.52 (8.87)	44.92 (9.78)	52.98 (12.13)	53.68 (11.29)

Table 7
Three-month test–retest reliability of the TAS-20 and its subscales in psychiatric outpatients

	Mean (S.D.) at test 1	Mean (S.D.) at test 2	r*
DIF	19.2 (6.4)	18.3 (6.3)	.71
DDF	14.0 (4.3)	13.3 (4.3)	.68
EOT	19.5 (4.9)	19.1 (4.5)	.66
TAS-20	52.7 (12.5)	50.6 (12.4)	.74

* Pearson's Product–Moment Correlation Coefficient, all correlations $P < .001$.

The discriminative validity of the TAS-20

Using the original TAS scales DIF [$F(3,677) = 27.71$, $P < .001$], DDF [$F(3,677) = 14.45$, $P < .001$], TAS-20 total score [$F(3,677) = 31.14$, $P < .001$], as well as EOT [$F(3,677) = 11.75$, $P < .001$] discriminate between the student participants and the psychiatric outpatients. Post hoc analysis shows that within both groups male and female respondents do not differ from each other with respect to their scores on one of these scales (Table 6).

Discussion

From the literature, it appears that the TAS-20 contains some, but not all, aspects of alexithymia. The fantasy aspect of the alexithymia construct is not represented in the TAS-20 and in most studies, the dimension EOT cannot be measured sufficiently reliable.

As previous research showed that the TAS-20 factor structure might be dependent on the sample studied, we studied both a patient and a nonpatient sample, applying confirmatory MGM analysis and exploratory SCA in order to find a stable factorial structure across samples. MGM did not yield the original three-factor structure of the TAS-20. SCA showed two factors to be present: a subscale (DIDF) comprising 11 items originating from the original subscales DIF and DDF and a subscale (EOT) with five items originating from the original subscale EOT.

In our study, the original DIF and DDF subscales merged into one single factor. This is in accordance with the results of several other studies. One may understand that the items of these two scales cluster together, because theoretically verbalisation and differentiation of feelings are interconnected [32,48].

The original DIF and DDF subscales have reasonable good internal consistencies, as has the DIDF subscale in our study. However, the internal consistency of the EOT subscale is inadequate. This is in agreement with the rather low

internal consistency of the EOT subscale generally reported in literature. We presume that the rather low internal consistency of the EOT subscale is due to the content of some items thought to represent externally oriented thinking.

Until now, the test–retest reliability of the TAS-20 has been primarily investigated in nonpatient samples, usually with a short time interval. This is the first study in which the test–retest reliability has been investigated in a clinical sample over a longer period. The test–retest reliability in psychiatric outpatients was satisfactory for the DIF subscale and the TAS-20 total score. The test–retest reliability of DDF and EOT subscales, however, just failed to reach the conventional level of adequacy.

In our study, as was the case in various other studies, the TAS discriminates well between psychiatric patients and adult nonpatients. Apart from the discriminative validity of the TAS-20, much effort has been put in studying its construct validity [3]. However, research into the criterion validity of the TAS-20 is scarce. TAS-20 threshold values indicating the presence of alexithymia at a clinically relevant level were reported in one article [49], but without references to empirical research. Taylor et al. [3] reported these threshold values to be based on a study [15] amongst 39 students in which the TAS scores are related to those of the BIQ, but the data on which the threshold values are based were not published. This is an important issue because as long as its criterion validity has not been firmly established, one may question whether the TAS-20 as a self-report questionnaire is an adequate instrument to assess alexithymia because alexithymic patients usually are not very self-reflective [3].

Alexithymia is clinically a relevant concept. Much effort has been put in studying the association between alexithymia and psychiatric pathology such as unexplained physical symptoms, eating disorders, and addictions. The TAS-20 appears to be the most commonly used instrument in this research. However, the present study shows some important shortcomings in terms of validity and reliability of the TAS-20. Therefore, we agree with the authors [3] of the instrument in recommending the TAS-20 to be used in empirical research in combination with other instruments such as the BIQ [3] or the subscales 'openness to feelings' and 'openness to fantasy' of the NEO Personality Inventory [50]. Furthermore, we do not encourage the use of the subscales of the TAS-20, because DIF and DDF probably represent one and the same dimension of alexithymia and the reliability of the EOT subscale repeatedly proves to be unreliable. For clinical practice, we do not recommend the use of the TAS-20 in favour of an assessment scale as the BIQ [3] or the alexithymia scale of the KAPP [6].

Acknowledgments

The authors wish to thank B. Bermond, PhD, Department of Psychology, Faculty of Behavioural and Social Sciences at the University of Amsterdam for providing the student data.

References

[1] Alexander F, French TM, Pollock GH. Psychosomatic specificity. Experimental study and results. Chicago: University Chicago Press, 1968.
[2] Sifneos PE. The prevalence of 'alexithymic' characteristics in psychosomatic patients. Psychother Psychosom 1973;22:255–62.
[3] Taylor GJ, Bagby RM, Parker JDA. Disorders of affect regulation: alexithymia in medical and psychiatric illness. Cambridge: Cambridge University Press, 1997.
[4] Taylor GJ. Alexithymia: concept, measurement, and implications for treatment. Am J Psychiatry 1984;141:725–32.
[5] Krystal JH, Giller EL, Cicchetti DV. Assessment of alexithymia in posttraumatic stress disorder and somatic illness: introduction of a reliable measure. Psychosom Med 1986;48:84–94.
[6] Weinryb RM, Roessel RJ. Karolinska Psychodynamic Profile; KAPP. Acta Psychiatr Scand, Suppl 1991;83:1–23.
[7] Cohen K, Auld F, Demers L, Catchlove R. Alexithymia. The development of a valid and reliable projective measure (the objectively scored Archetypal-9 Test). J Nerv Ment Dis 1985;173:621–7.
[8] Acklin MW, Alexander G. Alexithymia and somatization. A Rorschach study of four psychosomatic groups. J Nerv Ment Dis 1988; 176:343–50.
[9] Lane RD, Quinlan DM, Schwartz GE, Walker PA, Zeitlin SB. The Levels of Emotional Awareness Scale: a cognitive–developmental measure of emotion. J Pers Assess 1990;55:124–34.
[10] Haviland MG, Reise SP. A California Q-set alexithymia prototype and its relationship to ego-control and ego-resiliency. J Psychosom Res 1996;41:597–608.
[11] Apfel RJ, Sifneos PE. Alexithymia: concept and measurement. Psychother Psychosom 1979;32:180–90.
[12] Kleiger JH, Kinsman RA. The development of a MMPI alexithymia scale. Psychother Psychosom 1980;34:17–24.
[13] Bermond B, Vorst HCM, Vingerhoets AJ, Gerritsen W. The Amsterdam Alexithymia Scale: its psychometric values and correlations with other personality traits. Psychother Psychosom 1999;68:241–51.
[14] Bagby RM, Parker JD, Taylor GJ. The twenty-item Toronto Alexithymia Scale: I. Item selection and cross-validation of the factor structure. J Psychosom Res 1994;38:23–32.
[15] Bagby RM, Taylor GJ, Parker JD. The twenty-item Toronto Alexithymia Scale: II. Convergent, discriminant, and concurrent validity. J Psychosom Res 1994;38:33–40.
[16] Haviland MG, Warren WL, Riggs ML. An observer scale to measure alexithymia. Psychosomatics 2000;41:385–92.
[17] Linden W, Wen F, Paulhus DL. Measuring alexithymia: reliability, validity, and prevalence. Adv Pers Assess 1995;10:51–95.
[18] Taylor GJ. The alexithymia construct: conceptualization, validation, and relationship with basic dimensions of personality. New Trends Exp Clin Psychiatry 1994;10:61–74.
[19] Bach M, Bach D, de Zwaan M, Serim M, Bohmer F. Validierung der deutschen Version der 20-Item Toronto-Alexithymie-Skala bei Normalpersonen und psychiatrischen Patienten. Psychother, Psychosom, Med Psychol 1996;46:23–8.
[20] Bressi C, Taylor G, Parker J, Bressi S, Brambilla V, Aguglia E, Allegranti I, Bongiorno A, Giberti F, Bucca M, Todarello O, Callegari C, Vender S, Gala C, Invernizzi G. Cross validation of the factor structure of the 20-item Toronto Alexithymia Scale: an Italian multicenter study. J Psychosom Res 1996;41:551–9.
[21] Fukunishi I, Nakagawa T, Nakamura H, Kikuchi M, Takubo M. 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. Psychol Rep 1997;80:787–99.

[22] Pandey R, Mandal MK, Taylor GJ, Parker JD. Cross-cultural alexithymia: development and validation of a Hindi translation of the 20-Item Toronto Alexithymia Scale. J Clin Psychol 1996;52:173–6.
[23] Parker JDA, Bagby RM, Taylor GJ, Endler NS, Schmitz P. Factorial validity of the 20-item Toronto Alexithymia Scale. Eur J Pers 1993; 7:221–32.
[24] Rastam M, Gillberg C, Gillberg IC, Johansson M. Alexithymia in anorexia nervosa: a controlled study using the 20-item Toronto Alexithymia Scale. Acta Psychiatr Scand 1997;95:385–8.
[25] Salminen JK, Saarijarvi S, Aarela E, Toikka T, Kauhanen J. Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. J Psychosom Res 1999;46: 75–82.
[26] Simonsson-Sarnecki M, Lundh L-G, Törestad B, Bagby RM, Taylor GJ, Parker JDA. A Swedish translation of the 20-item Toronto Alexithymia Scale: cross-validation of the factor structure. Scand J Psychol 2000;41:25–30.
[27] Erni Th, Lötscher K, Modestin J. Two-factor solution of the 20-item Toronto Alexithymia Scale confirmed. Psychopathology 1997;30: 335–40.
[28] Taylor GJ, Ryan D, Bagby RM. Toward the development of a new self-report alexithymia scale. Psychother Psychosom 1985;44:191–9.
[29] Garner DM, Olmsted MP, Polivy J. Development and validation of a multi-dimensional eating disorder inventory for anorexia and bulimia. Int J Eating Disord 1983;2:15–34.
[30] Cacioppo JT, Petty RE. The need for cognition. J Pers Soc Psychol 1960;42:116–31.
[31] Haviland MG, Hendryx MS, Cummings MA, Shaw DG, MacMurray JP. Multidimensionality and state dependency of alexithymia in recently sober alcoholics. J Nerv Ment Dis 1991;179:284–90.
[32] Taylor GJ, Bagby RM, Parker JDA. The revised Toronto Alexithymia Scale: some reliability, validity, and normative data. Psychother Psychosom 1992;57:34–41.
[33] Crowne DP, Marlowe D. A new scale of social desirability independent of psychopathology. J Consult Clin Psychol 1960;24:349–54.
[34] Loas G, Corcos M, Stephan Ph, Pellet J, Bizouard P, Venisse JL, Perez-Diaz F, Guelfi JD, Flament M, Jeammet P. Factorial structure of the 20-item Toronto Alexithymia Scale; Confirmatory factorial analyses in nonclinical and clinical samples. J Psychosom Res 2001;50:255–61.
[35] Bach M, Bach D, de Zwaan M. Independency of alexithymia and somatization. A factor analytic study. Psychosomatics 1996;37: 451–8.
[36] Haviland MG, Hendryx MS, Shaw DG, Henry JP. Alexithymia in women and men hospitalized for psychoactive substance dependence. Compr Psychiatry 1994;35:124–8.
[37] Lumley MA, Mader C, Gramzow J, Papineau K. Family factors related to alexithymia characteristics. Psychosom Med 1996;58:211–6.
[38] Lumley MA, Norman S. Alexithymia and health care utilization. Psychosom Med 1996;58:197–202.
[39] Bagby RM, Nyklicek I, Vingerhoets A, Parker JDA, Rector NA. Validation of the 20-item Toronto Alexithymia Scale in a Dutch community sample. Eur J Pers Assess 2000 (in press).
[40] Haviland MG, Reise SP. Structure of the twenty-item Toronto Alexithymia Scale. J Pers Assess 1996;66:116–25.
[41] American Psychiatric Association.. Diagnostic and statistical manual of mental disorders. 4th ed. Washington (DC): American Psychiatric Association, 1994.
[42] UNESCO. International Standard Classification of Education (ISCED). Paris: UNESCO, 1976.
[43] Jonsson JO. Class origin, cultural origin and educational attainment: the case of Sweden. Eur Sociol Rev 1987;3:229–42.
[44] Cattell RB. The scree test for the number of factors. Multivar Behav Res 1966;1:245–76.
[45] Nunnally J. Psychometric theory. New York: McGraw-Hill, 1978.
[46] Kiers HAL. User's manual. SCA a program for simultaneous components analysis of variables measured in two or more populations. Groningen: ProGamma, 1990.
[47] Kiers HAL, Berge JMF. Alternating least squares algorithms for simultaneous components analysis with equal component weight matrices in two or more populations. Psychometrica 1989;54:467–73.
[48] Lane RD, Schwartz GE. Levels of emotional awareness: a cognitive–developmental theory and its application to psychopathology. Am J Psychiatry 1987;144:133–43.
[49] Parker J, Taylor G, Bagby M. Alexithymia and the processing of emotional stimuli: an experimental study. New Trends Exp Clin Psychiatry 1993;9:9–14.
[50] Costa PT, McCrae RR. Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa (FL): Psychological Assessment Resources, 1992.