Research Article

Exploring the Unidimensionality of the GAD-7 for South African First Responders: Evidence from Multiple Psychometric Approaches

Anita Padmanabhanunni [†], Tyrone Pretorius [†, *]

Department of Psychology, University of the Western Cape, Robert Sobukwe Road, Cape Town, South Africa; E-Mails: apadamana@uwc.ac.za; tpretorius@uwc.ac.za

† These authors contributed equally to this work.

* **Correspondence:** Tyrone Pretorius; E-Mail: tpretorius@uwc.ac.za

Abstract

Generalized Anxiety Disorder (GAD) is a prevalent mental health condition and the GAD-7 is widely used as a screening tool and measure of GAD severity. However, research has yielded conflicting findings regarding the factor structure of the instrument, with some studies supporting a unidimensional interpretation while others suggest a two-factor solution. These inconsistencies highlight the importance of validating the GAD-7 across diverse populations. This study aimed to assess the psychometric properties and factor structure of the GAD-7 among South African first responders (n = 429). Participants completed a demographic questionnaire along with the GAD-7, the Patient Health Questionnaire, the Post-Traumatic Stress Disorder Checklist, and subscales from the Maslach Burnout Inventory, namely Emotional Exhaustion and Depersonalization. The dimensionality of the GAD-7 was examined using Mokken Scale Analysis (MSA), parallel analysis, and Confirmatory Factor Analysis (CFA) with a Schmid-Leiman transformation. Results supported the unidimensionality of the GAD-7. MSA indicated that all items were loaded onto a single scale, with strong inter-item correlations and a high H-coefficient. The parallel analysis confirmed that only one factor was necessary, and the Schmid-Leiman transformation extracted a general factor accounting for 84% of the item variance. Furthermore, the reliability of the GAD-7 was excellent, with alpha

and omega coefficients exceeding 0.90, indicating high internal consistency. The strong reliability and validity of the GAD-7 in this study reinforce its utility as a reliable tool for assessing generalized anxiety in a non-clinical setting.

Keywords
Generalized anxiety disorder; GAD-7; dimensionality; Mokken analysis

1. Introduction

Generalized Anxiety Disorder (GAD) is one of the most common and widespread mental health conditions. It is characterized by persistent, excessive, and uncontrollable anxiety and worry about a wide range of everyday activities or events, such as work, health, social interactions, and routine tasks [1]. People with GAD often find it difficult to manage or control their worry, even when there is little or no reason for concern. This ongoing anxiety can interfere significantly with daily functioning, relationships, and overall well-being.

The worry experienced in GAD tends to be disproportionate to the actual situation and may shift from one focus to another, creating a constant state of unease. Unlike occasional anxiety, which is a natural stress response, the anxiety in GAD is chronic and can persist for months or years. This condition is often accompanied by physical symptoms such as restlessness, fatigue, difficulty concentrating, irritability, muscle tension, and sleep disturbances. Severe GAD has been associated with deficits in self-care and healthcare utilization.

The GAD-7 is the most commonly used screening tool and severity indicator for GAD [2, 3]. The current study examines the factor structure of the instrument among South African first responders. This population group is at increased risk of adverse mental health outcomes owing to their repeated exposure to traumatic events in the course of their work [4, 5]. Existing studies have underscored that first responders typically present for treatment with co-morbid symptoms of post-traumatic stress disorder (PTSD), depression, and GAD [6-8]. The limited available research in South Africa has confirmed that this population is at increased risk of developing a range of co-occurring mental health conditions. For instance, a study of ambulance personnel in the Western Cape province of the country reported a 30% prevalence rate for PTSD, which was significantly higher than the lifetime prevalence rate (2.3%) observed in the general population [9]. This elevated rate was attributed to the increased exposure to traumatic events experienced by first responders in this region compared to other provinces and the general population. First responders with pre-existing or comorbid mental health conditions were also found to experience heightened levels of distress [9]. A retrospective study further highlighted the heightened exposure to trauma among first responders in the Western Cape Province, revealing that 34% of cases managed (25.5 per 1,000 population) were trauma-related [10]. The majority of these cases (50.2%) were assault-related injuries, and most victims were women [10]. Research involving paramedic trainees in the Western Cape Province reported that 94% had experienced a traumatic event and 16% met criteria for PTSD [11]. Those with PTSD also had higher rates of depression [11]. The overlap of symptoms among co-occurring mental health disorders can lead to challenges differentiating between conditions, which can complicate the implementation of targeted interventions [6-8]. Hence, there is a critical need

for valid and reliable assessment measures that can accurately capture and differentiate between mental health disorders among first responder populations. This study aims to contribute to research in this area by assessing the psychometric properties and factor structure of the GAD-7 among first responders.

The original validation study of the GAD-7 was in a primary care setting and the scale demonstrated satisfactory internal consistency and factorial validity [12]. The psychometric properties of the GAD-7 have been evaluated in other primary care settings and produced inconsistent results. Kertz and colleagues, for instance, examined the factor structure of the GAD-7 in a sample of hospital patients and found that while the instrument demonstrated strong internal consistency and convergent validity, Confirmatory Factor Analysis (CFA) did not support its proposed unidimensional structure [13]. Using Exploratory Factor Analysis (EFA) and CFA, Johnson and colleagues investigated the instrument's psychometric properties among a heterogeneous sample of patients [14]. While the EFA supported a one-factor solution, the CFA provided a poorer fit to the data and the factor loading for item 6 (irritability) was particularly low. Using EFA, Beard and Björgvinsson assessed the dimensionality of the GAD-7 among psychiatric patients and proposed a two-factor solution. One factor reflected bodily symptoms while the other evaluated the cognitive and emotional experience of anxiety [15].

Research utilizing nationally representative surveys has offered support for the unidimensional structure of the scale. For example, Lowe and colleagues, using CFA, reported a one-factor solution for the GAD-7 in a nationally representative study in Germany [16]. Similarly, Shevlin and colleagues, using data from nationally representative samples in the United Kingdom, Ireland, Spain, and Italy, reported support for the unidimensional structure of the GAD-7 [17]. Among student populations, the results have been less consistent. For instance, Lee and Kim assessed the factor structure of the GAD-7 with Korean university students and found that the one-factor model provided only a marginal fit to the data [18]. They identified three items (4, 5, and 6) that shared unique residual variance beyond that explained by the GAD factor. They suggested that these items may reflect somatic symptoms relative to the other items, which were more cognitive. Byrd-Bredbenner and colleagues examined the psychometric properties of the GAD-7 among university students in the United States (US) and found support for its unidimensional structure [2]. In contrast, a separate study also conducted with university students in the US reported that CFA supported both a one-factor and a two-factor solution for the instrument [3]. The two-factor solution divided items into somatic tension symptoms and cognitive-emotional symptoms. Dhira and colleagues, in a study of university students in Bangladesh, found that the unidimensional model had a marginal fit and that a modified one-factor model was a better fit to the data [19].

In South Africa, the GAD-7 has been extensively used among various population groups and has demonstrated sound internal consistency reliability [20, 21]. However, similar to the findings in international research, inconsistencies have been noted in the instrument's factor structure. Van Wijk and colleagues, for instance, explored the clinical validity of the GAD-7 among South African workers and found evidence supporting a unidimensional structure [22]. Similarly, a study of university students reported that the one-factor model had the best fit [23]. In contrast, Hart and colleagues reported that one-factor and two-factor models demonstrated a reasonable fit among a sample of young women residing in an urban setting [20]. Kigozi examined the construct validity and reliability of the GAD-7 in a sample of tuberculosis patients, finding that a modified two-factor model provided the best fit [21]. The conflicting findings on the factor structure of the GAD-7

highlight the need for further examination of the instrument's psychometric properties across different cultural and demographic groups. These inconsistencies suggest that the GAD-7 may perform differently depending on the studied population, underscoring the importance of validating its use in various settings.

The current study aims to contribute to the existing literature by examining the factor structure and psychometric properties of the GAD-7 among a population group that is at high risk of adverse mental health outcomes. While the instrument has demonstrated reliability across various populations [13, 16, 21], inconsistencies in its factor structure—both internationally and within South Africa—highlight the need for a focused investigation into its validity and applicability for first responders who face distinct psychological stressors and co-occurring mental health conditions.

2. Materials and Methods

2.1 Participants and Procedures

The study used a cross-sectional research design and convenience sampling. The participants were first responders (n = 429) from the Western Cape Province of South Africa, consisting of police officers (n = 309) and paramedics (n = 120). An online version of the questionnaires was created using Google Forms. Police officers and paramedics could participate if employed in the Western Cape Province. No specific exclusion criteria were applied. To ensure diversity, recruitment occurred at multiple police stations and hospitals across both urban and peri-urban areas. Demographic information (e.g., age, gender, years of service, and work location) was collected to assess the representativeness of the sample post hoc. Multiple recruitment strategies were used to reach a broad and diverse sample of first responders. First, the link to the survey was distributed via email to police stations and local hospitals. To ensure that email invitations were received, a "request read receipt" feature was included in the email distribution. This allowed us to confirm that the intended recipients opened the emails. Follow-up emails were sent to administrative contacts two weeks after the initial invitation.

Second, research assistants visited police stations and hospitals to recruit participants in person. This face-to-face interaction promoted participation among individuals less familiar with online surveys. Third, permission was obtained from administrators of Facebook groups dedicated to first responders to post the survey link and an invitation to participate. These groups are active communication platforms for first responders to share resources and discussions relevant to their work. This method facilitated access to individuals who might not be reached through email or in-person recruitment, particularly those in remote areas or on different shifts. While convenience sampling inherently has limitations, our recruitment methods aimed to mitigate these by engaging participants from varied settings and roles.

The sample was predominantly male (55%) and worked in an urban area (92.3%). The majority of the sample was married and had on average 2.4 dependents (SD = 1.86). The mean age of the sample was 39 years (SD = 9.93) and the mean length of service as a first responder was 13.23 years (SD = 9.65).

2.2 Instruments

Participants completed a short demographic questionnaire along with a survey that included various instruments measuring mental health indicators including the Generalized Anxiety Disorder Scale (GAD-7) [12], the Patient Health Questionnaire (PHQ-9) [24], the Post-Traumatic Stress Disorder Checklist (PCL) [25] and two subscales of the Maslach Burnout Inventory (MBI) namely Emotional Exhaustion and Depersonalization [26]. The GAD-7 is a self-report questionnaire used to assess the severity of generalized anxiety disorder symptoms. It consists of seven items that evaluate common anxiety symptoms over the past two weeks. Responses are rated on a 4-point scale ranging from 0 (not at all) to 3 (nearly every day), with higher scores indicating greater anxiety severity. An example of an item from the GAD-7 is: "*Feeling nervous, anxious, or on edge*". In South Africa, the GAD-7 has demonstrated good reliability, with a Cronbach's alpha of 0.86 in a sample of tuberculosis patients [21] and 0.89 in a sample of university students [23].

The PHQ-9 is a 9-item instrument used to measure depression. Responses to the nine items are rated on a 4-point scale, ranging from 0 (not at all) to 3 (nearly every day). An example of a PHQ-9 item is: "*Feeling bad about yourself—or that you are a failure or have let yourself or your family down?*". The initial validation study of the instrument reported reliability coefficients of 0.89 and 0.86 for two different samples. The PHQ-9 has been used in patients in a primary health care patient population and among university students in South Africa and demonstrated sound reliability (α = 0.76; α = 0.83) [27, 28]. The inclusion of the PHQ-9 allows for the assessment of the discriminant validity of the GAD-7 by examining the extent to which it measures anxiety independently of overlapping depressive symptoms, which are typically comorbid with anxiety disorders.

The PCL is a self-report questionnaire designed to assess symptoms of PTSD. Post-traumatic stress symptomology frequently co-occurs with GAD [7, 8], especially in populations exposed to trauma. Including the PCL supports the evaluation of GAD-7 by comparing its performance with a related but distinct construct. The PCL consists of 20 items, aligned with the criteria outlined in the DSM-5 for PTSD, and evaluates how much an individual has been bothered by specific symptoms over the past month. Responses are rated on a 5-point scale ranging from 0 (*not at all*) to 4 (*extremely*), with higher scores indicating greater PTSD symptom severity. An example of an item from the PCL is "*Repeated, disturbing memories, thoughts, or images of a stressful experience from the past?*" The PCL had been validated in the South African context among, for example, patients receiving care for HIV (α = 0.97) [29] and university students (α = 0.93) [30].

The MBI is a widely used self-report tool designed to assess burnout, particularly in occupational settings. It measures three key dimensions of burnout: emotional exhaustion, depersonalization (cynicism), and personal accomplishment. Anxiety symptoms frequently contribute to or overlap with burnout. Hence, the inclusion of the MBI allows for an evaluation of the predictive validity of the GAD-7 in relation to occupational stress and burnout among a high-risk population. The MBI consists of 22 items, and responses are rated on a 7-point Likert scale, ranging from 0 (*never*) to 6 (*every day*). Higher scores on emotional exhaustion and depersonalization indicate greater burnout, while lower scores on personal accomplishment suggest reduced feelings of competence and success at work. An example of an item from the MBI is: "*I feel emotionally drained from my work*". The MBI has shown good reliability in a sample of South African police officers, with Cronbach's alpha values of 0.88 for emotional exhaustion and 0.78 for depersonalization [31].

2.3 Ethics

The study adhered to the ethical principles outlined in the Declaration of Helsinki and received approval from several bodies. Ethical clearance was granted by the Humanities and Social Sciences Research Ethics Committee of the University of the Western Cape (ethics reference: HS23/2/4, 23 May 2023) and the South African Police Services (ethics reference: 3/34/2, 27 June 2023). Additionally, permission was obtained from the Western Cape Department of Health to access hospitals (reference: WC_202307_041, 15 September 2023) and from a private ambulance company to recruit their employees (reference: 12 December 2023). Participation in the study was voluntary, with informed consent obtained on the first page of the electronic survey link, and no personal identifying information was collected. The survey platform was configured to prevent the collection of email addresses that could potentially identify participants. Detailed information about the purpose of the study, procedures, and participants' rights was provided on the first page of the online survey. The information page also provided the contact details of non-governmental organizations providing free mental health care services, and participants were encouraged to access these resources if completing the survey evoked any distress. Participants had to explicitly agree to participate by selecting a consent option before proceeding to the study.

2.4 Data Analysis

We used Mokken scale analysis (MSA), parallel analysis, and confirmatory factor analysis (CFA: maximum likelihood) with a Schmid-Leiman transformation to examine the dimensionality of the GAD-7. MSA was conducted using the package "Mokken" [32] in R software [33]. Parallel analysis was conducted using SPSS syntax available on the Internet [34]. CFA was conducted with IBM SPSS Amos for Windows Version 28 (IBM Corp., Armonk, NY, USA), and the Schmid-Leiman transformation was obtained with package "psych" [35] in R software. All classical test theory indices (e.g., alpha, item-total correlations, etc.) were obtained with IBM SPSS for Windows Version 29 (IBM Corp., Armonk, NY, USA).

MSA is a non-parametric approach to scale development and evaluation that assesses the dimensionality and scalability of items in an instrument [36]. MSA uses an algorithm referred to as automated item selection procedure (AISP) to determine how many scales the items of an instrument load on. If an item is unscalable and does not load on any scale, it will be assigned a zero, and if, for example, the items load on two scales, items will be assigned the values one and two. Thus, on a unidimensional scale, all items will be assigned a value of one [37]. A scalability coefficient indicates the contribution of each item to the measurement of the latent variable H_i. In general, H_i-coefficients greater than 0.30 reflect items that largely satisfactorily contribute to the overall scale [36]. In addition to these individual H_i-coefficients, MSA also provides an overall H-coefficient that demonstrates the strength of a unidimensional scale. It is proposed that a H-coefficient greater than 0.50 indicates a substantial scale [38]. One of the important assumptions in MSA is that the likelihood of endorsing an item does not decrease as values of the latent variable increase [38]. This is referred to as the assumption of monotonicity and is assessed in MSA by a *Crit* value, with *Crit* values greater than 80 reflecting serious violations of monotonicity [39]. In addition to the *Crit* value, MSA indicates whether any item violated monotonicity (#vi) and whether these violations were significant (#zsig).

The flexibility of MSA and its assumption of monotonicity makes it particularly valuable for psychometric analyses, especially in contexts where parametric assumptions, such as normality, may not hold. This is crucial when working with ordinal data, as with the GAD-7. By systematically examining individual items and the overall scale, MSA offers a robust framework for assessing the unidimensionality and reliability of measurement instruments.

Parallel analysis is one of the most accurate methods to determine the number of factors represented by a set of items [40]. This analysis method is particularly relevant given the inconsistencies in prior studies regarding the factor structure of the GAD-7. This method provides an objective, data-driven approach to factor retention, addressing a key limitation of traditional methods, which may overestimate the number of factors [41]. In parallel analysis, eigenvalues obtained in the current dataset are compared to several simulated datasets (N = 1000). Any eigenvalue in the current dataset that is greater than the 95th percentile of eigenvalues in the simulated datasets is considered to be a viable factor. A unidimensional solution is obtained if only one eigenvalue is greater than the 95th percentile of eigenvalues in the simulated datasets [41].

CFA allows for a rigorous evaluation of competing models, such as unidimensional and multidimensional factor structures, which is critical given the inconsistent findings regarding the dimensionality of the GAD-7 in previous research. This approach provides insight into whether the GAD-7 operates as a single overarching construct (generalized anxiety) or if it comprises distinct but related subdimensions, such as somatic and cognitive symptoms. We used CFA to examine three models of the factor structure of the GAD-7: a one-factor model, a correlated two-factor model (with cognitive-emotional and somatic tension as the two factors), and a second-order factor model in which the two factors load on a higher order factor representing a total anxiety score. In contrast to exploratory methods, CFA tests predefined models, making it particularly suitable for assessing whether the structure of an instrument aligns with theoretical expectations or prior empirical findings.

The extent to which the model of the factor structure fit the data was examined using the following fit indices as recommended by Gaskin and colleagues [42]: Chi-squared (χ^2), relative χ^2 (χ^2/pdf), the comparative fit index (CFI), the standardized root mean square residual (SRMR), the root mean square error of approximation (RMSEA), and the p-value of close fit (PCLOSE). Acceptable fit is indicated by the following rule-of-thumb criteria [42]:

- χ^2 should be non-significant; however, that would indicate a perfect fit [43].
- Relative χ^2 should be between 1 and 3.
- CFI should be greater than 0.95.
- SRMR should be less than 0.08.
- RMSEA should be less than 0.06.
- PCLOSE is a statistical test that RMSEA equals zero. It should be nonsignificant (p > 0.05), indicating a good model fit.

In addition to the above fit indices, it is recommended that a model comparison index such as the Akaike information criterion (AIC) be included when comparing models, and models with lower AIC values are regarded as better-fitting models [44]. In addition to these fit indices, we also selected standardized estimates and direct, indirect, and total effects, which are optional outputs in Amos.

While the CFA indicates the extent to which a particular structure fits the data to an acceptable degree, it does not indicate whether the factors in such a structure explain a sufficient amount of item variance. We thus used the Schmid-Leiman procedure to transform the second-order factor

structure into a bifactor structure that enables the extraction of a general factor (total anxiety scale) and two specific factors (cognitive-emotional and somatic tension factors). This transformation provides the explained common variance (ECV) and omega hierarchical (Omega$_h$) for the general factor. ECV is the percentage of item variance explained by a factor. At the same time, Omega$_h$ is a reliability estimate of the general factor based on the correlation of lower-order factors. An Omega$_h$ greater than 0.80 for the general factor indicates high reliability and suggests that total scores are essentially unidimensional [45].

In sum, these methods were chosen for their complementary strengths in testing unidimensionality and exploring the factor structure of measuring instruments. MSA provided a robust, non-parametric approach to assess scalability and monotonicity. At the same time, parallel analysis offered a rigorous, data-driven method for determining the number of meaningful factors, addressing the limitations of more traditional approaches. CFA provided a framework for validating and comparing factor models, ensuring theoretical alignment and empirical validity. Together, these techniques enabled a comprehensive examination of the dimensionality and reliability of the GAD-7, supporting its application among South African first responders.

The internal consistency of the GAD-7 was examined using several reliability indices – alpha, omega, composite reliability, and Mokken scale reliability (MS$_{rho}$) – as well as inter-item correlations, factor loadings, and item-total correlations. Inter-item correlations indicate the homogeneity of items and should ideally be between 0.15 and 0.85 [46]; less than 0.15 would indicate that the items do not have much in common, and more than 0.85 would reflect the redundancy of items. Item-total correlations indicate the extent to which all items contribute to the measurement of the latent variable and should be greater than 0.50 [47]. To obtain the factor loadings, we conducted an exploratory factor analysis (EFA: principal components analysis) specifying a forced one-factor solution. Like item-total correlations, factor loadings also indicate the extent to which items contribute to the measurement of the latent variable and should be above 0.55 [48]. Item-total correlations, factor loadings, and H$_i$ coefficients provide evidence of construct validity.

We examined convergent and discriminant validity using average variance extracted (AVE), composite reliability (CR), and average shared variance (ASV). It is suggested that if AVE is greater than 0.50 and AVE is less CR it would provide support for convergent validity. Further, to demonstrate discriminant validity, it is proposed that a variable should have more in common with items that contribute to its measurement (AVE) than it has in common with other variables that it is related to (ASV). Lastly, criterion-related validity was examined through the correlations between anxiety and related variables depression, PTSD, emotional exhaustion, and depersonalization. Emotional exhaustion and depersonalization are components of burnout.

3. Results

The results of the MSA are reported in Table 1 and provide strong evidence for the unidimensionality of the GAD-7. AISP found that all seven items loaded on a single scale (denoted by the value 1). The individual H-coefficients (H$_i$) were all greater than 0.30 and ranged between 0.62 and 0.70, reflecting that all items contribute substantially to the measurement of anxiety. There were no significant violations of monotonicity as indicated by *Crit*, #vi, and #zsig.

Table 1 Results of the Mokken Analysis of the GAD-7.

GAD-7 item	AISP indicator	H_i	SE	Monotonicity		
				Crit	#vi	#zsig
1. Nervous tick	1	0.63	0.02	0	0	0
2. Unable to control worrying	1	0.62	0.03	0	0	0
3. Worrying too much	1	0.64	0.03	0	0	0
4. Trouble relaxing	1	0.65	0.02	0	0	0
5. Being restless	1	0.63	0.02	0	0	0
6. Easily irritable	1	0.63	0.03	0	0	0
7. Feeling afraid	1	0.70	0.03	0	0	0

Note. AISP = automated item selection procedure, #vi = violations of monotonicity, #zsig = significance of monotonicity violations.

The results of parallel analysis supported the findings from MSA that the GAD-7 is essentially unidimensional in that only the first eigenvalue in the current dataset (4.81) was greater than the 95th percentile of eigenvalues (1.25) in the simulated datasets. The second eigenvalue (0.59) was less than the 95th percentile of eigenvalues (1.15) in the simulated datasets, demonstrating that one factor is sufficient to account for item variance.

The inter-item correlations, descriptive statistics for the items, item-total correlations, and factor loadings are reported in Table 2.

Table 2 Item-level indices for the GAD-7.

GAD-7 item	1	2	3	4	5	6	7
1. Nervous	—						
2. Unable to control worrying	0.72**	—					
3. Worrying too much	0.64**	0.75**	—				
4. Trouble relaxing	0.64**	0.65**	0.70**	—			
5. Being restless	0.64**	0.62**	0.60**	0.75**	—		
6. Easily irritable	0.59**	0.56**	0.54**	0.61**	0.65**	—	
7. Feeling afraid	0.60**	0.58**	0.60**	0.62**	0.62**	0.65**	—
Mean	1.07	1.17	1.19	1.10	0.97	1.19	1.03
SD	0.99	1.02	0.97	1.02	1.02	1.02	1.03
ITC	0.77**	0.78**	0.77**	0.80**	0.79**	0.71**	0.73**
λ	0.84**	0.84**	0.84**	0.86**	0.84**	0.79**	0.80**

Note. ITC = item-total correlation, **p < 0.001.

Table 2 indicates that the inter-item correlations ranged between 0.54 and 0.75, thus indicating that all items come from the same content domain and there are no issues with regard to the redundancy of items. All the item-total correlations were above 0.50 and ranged between 0.71 and 0.80. These strong item-total correlations, together with the factor loadings above 0.55 and ranged between 0.79 and 0.86, demonstrated that all items contribute substantially to the measurement of anxiety.

The three models of the factor structure of the GAD-7 that were examined using CFA are shown in Figure 1 and the fit indices for the three models are reported in Table 3.
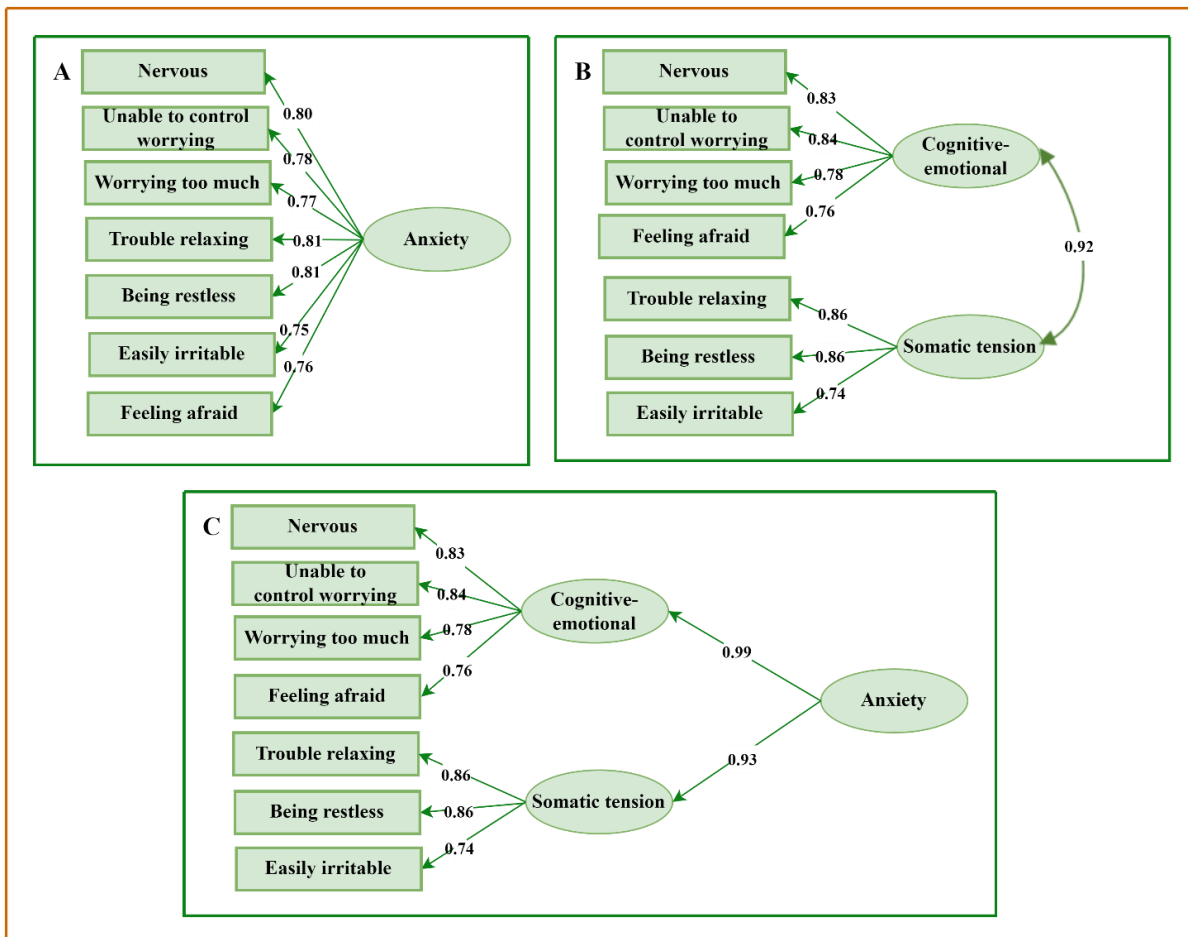

Figure 1 Three Models of the factor structure of the GAD-7 (Rectangles are observed variables, ellipses are latent variables. All regression coefficients are standardized and significant – $p < 0.001$).

Table 3 CFA fit indices for three models of the GAD-7.

Good fit index	Cutoff criteria	One-factor	Correlated two-factor	Second-order factor
$\chi^2(df)$		19.366(9)	20.048(9)	20.048(9)
p-value	nonsignificant	<0.001	<0.001	<0.001
Relative χ^2	Between 1 and 3	2.152	2.228	2.228
CFI	>0.95	0.995	0.995	0.995
SRMR	<0.08	0.018	0.018	0.018
RMSEA	<0.06	0.052	0.054	0.054
PCLOSE	>0.05	0.415	0.381	0.381
AIC	Lower values	71.366	72.048	72.048

Table 3 indicates that all three models fit the data to an acceptable degree (relative χ^2 <3, CFI >0.95, SRMR <0.08, RMSEA <0.06, PCLOSE >0.05). The correlated two-factor and the second-

order factor models had similar fit indices. The one-factor model had marginally a lower AIC. However, there was a substantial association ($r = 0.92$) between the two factors in the two-factor model.

As these results might suggest that the GAD-7 can be used as a single scale or two factors or two factors as well as a total scale, we used a Schmid-Leiman transformation to decompose sources of variance in order to determine the ECV for a general factor obtained as a result of the transformation. The results showed that the general factor accounted for 85% of item variance, while the two specific factors thus accounted for only 15%. In addition, Omega$_h$ was greater than 0.80 (Omega$_h$ = 0.84), which supports the interpretation of the GAD-7 as unidimensional.

The indices for the GAD-7 at the scale level are reported in Table 4. All the reliability indices were satisfactory and above 0.90. AVE was greater than ASV and less than CR. The MSA H-coefficient of 0.67 indicated a very strong scale.

Table 4 Scale level indices for the GAD-7.

Index	Value	Suggested cutoff
Cronbach's Alpha	0.92	≥0.75
McDonald's Omega	0.92	≥0.75
Composite reliability (CR)	0.96	≥0.75
Mokken scale reliability (MS$_{rho}$)	0.93	≥0.75
Average variance extracted (AVE)	0.68	≥0.50
Average shared variance	0.38	<AVE
Scale H (Mokken)	0.67	≥0.50

The correlations between anxiety and related constructs are reported in Table 5. These coefficients reflect substantial associations between anxiety and depression, PTSD, emotional exhaustion, and depersonalization.

Table 5 Intercorrelations between variables used to establish concurrent validity.

GAD-7 item	1	2	3	4	5
1. Anxiety	—				
2. Depression	0.73**	—			
3. PTSD	0.66**	0.59**	—		
4. Emotional exhaustion	0.55**	0.56**	0.63**	—	
6. Depersonalization	0.49**	0.46**	0.56**	0.80**	—

**p < 0.001.

4. Discussion

The current study assessed the psychometric properties and factor structure of the GAD-7 among a sample of South African first responders using different but complementary psychometric paradigms. GAD is conceptualized as a pervasive mental health condition characterized by excessive and uncontrollable worry and accompanied by a range of somatic and cognitive symptoms [1]. This theoretical understanding aligns with the high degree of interrelatedness among the symptoms of

GAD, as well as its overlap with symptoms of other disorders, including PTSD, depression, and burnout [6, 8]. Previous research on the factor structure of the GAD-7 has produced conflicting results, with some studies supporting a unidimensional structure. In contrast, others propose a two-factor model reflecting cognitive and somatic symptom dimensions [2, 13, 14]. These discrepancies can be attributed to variations in population characteristics, divergence in methodological approaches as well as contextual differences across studies [49].

The GAD-7 was initially assessed on a clinical sample. A common problem in research with such population groups is the significant deviations from normal distributions observed in variables measuring mental health constructs. This can amplify the shared variance across items or lead to variability in symptom clustering [49, 50]. Studies using non-clinical or heterogeneous populations may detect more variability in symptom clustering, supporting a multidimensional structure [2, 17]. For example, cognitive and somatic symptoms may manifest differently depending on factors such as trauma exposure or stress levels, particularly in high-risk populations. Methodological differences, such as exploratory versus CFA, may demonstrate multidimensionality due to statistical artifacts or a lack of stringent criteria for factor retention [50]. By leveraging the strengths of a combination of robust analytical methods, including MSA, parallel analysis, and CFA with a Schmid-Leiman transformation, the current study offers a comprehensive evaluation of the dimensionality of the GAD-7. The use of multiple methods contributes to addressing the limitations of prior studies that may have relied on narrower methodological approaches.

In the current study, the different paradigms converged in demonstrating that the GAD-7 is a unidimensional scale. This aligns with the conceptualization of GAD as a single, overarching construct [15, 16]. Regarding unidimensionality, AISP in MSA indicated that all items were loaded on a single scale. Parallel analysis confirmed that a single scale was sufficient as only one eigenvalue in the raw data set exceeded the 95th percentile of eigenvalues in the simulated datasets. This pattern of eigenvalues aligns with the conceptualization of GAD as a single construct, as it suggests that the majority of variance in item responses is explained by a single dominant factor rather than multiple distinct dimensions. High inter-item correlations, ranging between 0.54 and 0.75, further reflect this shared variance, demonstrating that all items consistently measure a common underlying construct of generalized anxiety.

CFA results provided further evidence supporting unidimensionality. While the one-factor, correlated two-factor, and second-order factor models all fit the data to an acceptable degree, the Schmid-Leiman transformation extracted a general factor accounting for 84% of the item variance. The two subscales—cognitive-emotional and somatic tension factors—explained only 15% of the variance. This overwhelmingly dominant general factor indicates that the items of the GAD-7 primarily reflect a single overarching construct of anxiety rather than distinct subcomponents. The high correlation ($r = 0.92$) between the two factors in the correlated two-factor model further suggests redundancy, supporting the interpretation of GAD as a unified phenomenon.

The MSA results reinforced this conclusion by demonstrating the strength of the unidimensional scale, with a high H-coefficient of 0.67. This high H-coefficient, alongside strong Hi-coefficients exceeding 0.60, provides further evidence that the GAD-7 reliably measures a single latent construct. The combination of findings from AISP, Parallel Analysis, CFA, and the Schmid-Leiman transformation provides compelling support for interpreting the GAD-7 as a unidimensional measure. The minimal variance explained by the subscales and their high intercorrelation reinforce the conclusion that the GAD-7 is best viewed as capturing a singular construct of anxiety, rather

than distinct subcomponents. The confirmation of the scale's unidimensionality underscores its utility for both research and clinical practice, as it simplifies scoring and interpretation, allowing for a more straightforward assessment of generalized anxiety.

The GAD-7 evidenced a high level of internal consistency with alpha, omega, composite reliability, and Mokken scale reliability, all exceeding 0.90. The alpha value obtained in the current study suggests that all items consistently measure the same underlying construct. This minimizes measurement error and supports the reliability of the total score as a representation of generalized anxiety. The omega value provides a nuanced estimate of reliability by considering the individual contributions of each item to the overall construct. The high omega score observed in this study confirms that the GAD-7 is both internally consistent and a robust measure of generalized anxiety. Taken together, these reliability indices demonstrate the strength and stability of the GAD-7 as an assessment instrument. The findings underscore its homogeneity as a measure of generalized anxiety and indicate that the instrument reliably captures the construct across all items and is a stable assessment tool.

The theoretical foundation of GAD as a pervasive and overarching construct characterized by excessive worry and tension is consistent with these findings. The homogeneity of the item set, evidenced by high inter-item correlations, item-total correlations greater than 0.50, and factor loadings above 0.55, suggests that the scale captures a unified construct rather than disparate symptom clusters. This supports the notion that cognitive and somatic symptoms are not distinct dimensions but interrelated manifestations of the same underlying anxiety disorder.

AVE was greater than 0.50 and less than CR, supporting convergent validity. AVE was greater than ASV which supports the discriminant validity as it indicates that the latent variable anxiety had more in common with the items that contributed to its measurement than related variables. This suggests that the latent variable—anxiety—shares more variance with its items than with items from other related constructs. In other words, the GAD-7 can distinguish anxiety from other constructs, such as depression or emotional exhaustion, ensuring that it measures anxiety specifically rather than overlapping significantly with other psychological conditions. Criterion-related validity was also demonstrated, as the GAD-7 correlated appropriately with related constructs, such as depression, PTSD, emotional exhaustion, and depersonalization. This means that the scale behaves as expected in relation to well-established measures of similar psychological conditions, based on clinical and theoretical models of comorbidity. For instance, anxiety is often comorbid with depression and PTSD [51, 52], and the instrument's relationships with these variables provide empirical support that it measures anxiety in a manner consistent with theoretical and clinical expectations.

In sum, the high inter-item correlations, eigenvalue patterns, and variance decomposition collectively support the interpretation of the GAD-7 as a unidimensional measure of generalized anxiety. This unidimensionality aligns with the theoretical understanding of GAD as a single, overarching construct and underscores the utility of the instrument in both research and clinical settings, where a reliable measure of generalized anxiety is essential.

This study has important implications for clinical research and practice. The fact that the GAD-7 primarily measures a single overarching construct of anxiety simplifies its scoring and interpretation. With a unidimensional structure, all item responses contribute to a single total score, thus precluding the need to calculate or interpret separate subscale scores. This scoring approach reduces complexity and makes the tool more accessible in diverse clinical and research settings. For

clinicians, the unidimensionality of the GAD-7 means that it can be quickly administered and interpreted in routine practice settings. A single total score provides a clear and concise reflection of the severity of anxiety symptoms, enabling expedient decision-making regarding diagnosis and treatment planning. This simplicity is particularly beneficial in highly demanding work environments, such as primary care or emergency service settings where time constraints limit the feasibility of using more complex assessment measures. The ease of administration and interpretation of the GAD-7 also makes it useful for non-specialist practitioners, such as community health workers, who may not have training in psychometric assessment but still need a reliable measure of anxiety for screening and monitoring purposes.

For researchers, the unidimensional structure supports the use of the GAD-7 in large-scale studies or surveys where ease of administration and consistency across diverse populations is essential. Furthermore, the strong psychometric properties of the instrument, such as high internal consistency and robust construct validity, enhance confidence in its use as a measure of generalized anxiety. The unidimensionality of the GAD-7 also minimizes the risk of misinterpretation, as researchers are not required to separate overlapping dimensions or reconcile scores across multiple subscales.

Compared to other widely used measures of anxiety, notably the State-Trait Anxiety Inventory (STAI) [53] and the Beck Anxiety Inventory (BAI) [54], the unidimensional structure of the GAD-7 offers distinct advantages. The STAI, for instance, differentiates between state anxiety, which is temporary and situational, and trait or dispositional anxiety, thereby capturing a broader range of anxiety-related constructs [53]. Similarly, the BAI includes a range of neurophysiological, autonomic, and cognitive symptoms, leading to a multidimensional factor structure [54]. While these tools can provide more granulated insights into the potential subdomains of anxiety, their complexity can make them less practical in settings where quick screening or interpretation is needed.

While the current study provides evidence for the unidimensionality and reliability of the GAD-7 among South African first responders, it has specific limitations that may influence the generalizability of the results. First responders represent a high-risk group with distinctive occupational stressors, including frequent exposure to potentially traumatic events and demanding working environments. This may influence the manifestation of anxiety symptoms and the performance of the GAD-7, potentially impacting the applicability of these findings to other populations. The current sample was drawn from a single province in South Africa, where previous research has documented higher levels of trauma exposure among first responders compared to those in other provinces [9]. This may affect the generalizability of the results to first responders in different areas or to the general population. The study utilized a non-clinical sample of first responders, and the results may not fully extend to clinical populations with diagnosed GAD or to other high-risk occupational groups such as healthcare workers or military personnel. The sample also lacked representation from broader demographic groups which could influence the performance of the GAD-7.

To validate the conclusions of the current study and determine its broader applicability, future research should examine the psychometric properties of the GAD-7 in diverse demographic and clinical groups in various settings. Studies involving the general population, other high-risk occupational groups, and culturally diverse settings are recommended to determine whether the unidimensionality and reliability observed in this study are across different contexts. Such research would provide a more comprehensive understanding of the instrument's utility as a measure of

generalized anxiety. It is also recommended that longitudinal studies assess the stability and sensitivity of the GAD-7 to shifts in anxiety symptoms over time, particularly among individuals undergoing treatment. This could determine whether the unidimensional structure reliably captures reductions in generalized anxiety. Intervention studies assessing the instrument's responsiveness to therapeutic approaches (e.g., cognitive behavioral therapy for GAD) can determine whether the total score aligns with clinical improvements across cognitive and somatic symptom domains.

5. Conclusions

The current study provides evidence supporting the GAD-7 as a reliable and valid tool for assessing generalized anxiety among first responders. The findings confirm the unidimensional structure of the scale, with high internal consistency and strong construct validity, making it a valuable resource for both clinical practice and research. The demonstrated convergent and discriminant validity further enhances its applicability by confirming that it accurately captures anxiety without overlap with related constructs like depression or PTSD. These results underscore the utility of the GAD-7 for tracking anxiety.

Author Contributions

AP and TBP contributed equally to the conceptualization and data collection. TP was responsible for the data analysis. All authors contributed equally to the writing, review, and editing of this article.

Funding

This research received no specific grant from any funding

Competing Interests

The authors have declared that no competing interests exist.

Data Availability Statement

The data sets generated and/or analyzed during the current study are available from the corresponding author (TBP) upon reasonable request.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5-TR. 5th ed. Washington, D.C.: American Psychiatric Association Publishing; 2022.
2. Byrd-Bredbenner C, Eck K, Quick V. GAD-7, GAD-2, and GAD-mini: Psychometric properties and norms of university students in the United States. Gen Hosp Psychiatry. 2021; 69: 61-66.
3. White AE, Karr JE. Psychometric properties of the GAD-7 among college students: Reliability, validity, factor structure, and measurement invariance. Transl Issues Psychol Sci. 2023. doi: 10.1037/tps0000382.

4. Bonumwezi JL, Tramutola D, Lawrence J, Kobezak HM, Lowe SR. Posttraumatic stress disorder symptoms, work-related trauma exposure, and substance use in first responders. Drug Alcohol Depend. 2022; 237: 109439-109439.

5. Huang G, Chu H, Chen R, Liu D, Banda KJ, O'Brien AP, et al. Prevalence of depression, anxiety, and stress among first responders for medical emergencies during COVID-19 pandemic: A meta-analysis. J Glob Health. 2022; 12: 05028.

6. Baker LD, Ponder WN, Carbajal J, Galusha JM, Hidalgo JE, Price M. Mapping PTSD, depression, and anxiety: A network analysis of co-occurring symptoms in treatment-seeking first responders. J Psychiatr Res. 2023; 168: 176-183.

7. Ponder WN, Walters K, Simons JS, Simons RM, Jetelina KK, Carbajal J. Network analysis of distress, suicidality, and resilience in a treatment seeking sample of first responders. J Affect Disord. 2023; 320: 742-750.

8. Price M, Legrand AC, Brier ZM, Hébert-Dufresne L. The symptoms at the center: Examining the comorbidity of posttraumatic stress disorder, generalized anxiety disorder, and depression with network analysis. J Psychiatr Res. 2019; 109: 52-58.

9. Ntatamala I, Adams S. The correlates of post-traumatic stress disorder in ambulance personnel and barriers faced in accessing care for work-related stress. Int J Environ Res Public Health. 2022; 19: 2046.

10. Abdullah N, Saunders C, McCaul M, Nyasulu P. A retrospective study of the pre-hospital trauma burden managed by the western cape government emergency medical services. S Afr J Pre Hosp Emerg Care. 2021; 2: 18-26.

11. Fjeldheim CB, Nöthling J, Pretorius K, Basson M, Ganasen K, Heneke R, et al. Trauma exposure, posttraumatic stress disorder and the effect of explanatory variables in paramedic trainees. BMC Emerg Med. 2014; 14: 11.

12. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Intern Med. 2006; 166: 1092-1097.

13. Kertz S, Bigda-Peyton J, Bjorgvinsson T. Validity of the Generalized Anxiety Disorder-7 Scale in an acute psychiatric sample. Clin Psychol Psychother. 2013; 20: 456-464.

14. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the General Anxiety Disorder 7-Item (GAD-7) Scale in a heterogeneous psychiatric sample. Front Psychol. 2019; 10: 1713.

15. Beard C, Björgvinsson T. Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. J Anxiety Disord. 2014; 28: 547-552.

16. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care. 2008; 46: 266-274.

17. Shevlin M, Butter S, McBride O, Murphy J, Gibson-Miller J, Hartman TK, et al. Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. BMC Psychiatry. 2022; 22: 154.

18. Lee B, Kim YE. The psychometric properties of the Generalized Anxiety Disorder scale (GAD-7) among Korean university students. Psychiatry Clin Psychopharmacol. 2019; 29: 864-871.

19. Dhira TA, Rahman MA, Sarker AR, Mehareen J. Validity and reliability of the Generalized Anxiety Disorder-7 (GAD-7) among university students of Bangladesh. PLoS One. 2021; 16: e0261590.

20. Hart C, Draper CE, Soepnel LM, Godongwana M, Mabetha K, Nyati LH, et al. Examining the psychometric properties of the Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7 among young urban South African women. J Affect Disord. 2025; 369: 61-70.

21. Kigozi G. Construct validity and reliability of the generalised anxiety disorder-7 scale in a sample of tuberculosis patients in the Free State Province, South Africa. S Afr J Infect Dis. 2021; 36: 298.

22. van Wijk CH, Martin JH, Maree DJ. Clinical validation of brief mental health scales for use in South African occupational healthcare. SA J Ind Psychol. 2021; 47: a1895.

23. Tadi NF, Pillay K, Ejoke UP, Khumalo IP. Sex differences in depression and anxiety symptoms: Measurement invariance, prevalence, and symptom heterogeneity among university students in South Africa. Front Psychol. 2022; 13: 873292.

24. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med. 2001; 16: 606-613.

25. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. J Trauma Stress. 2015; 28: 489-498.

26. Maslach C, Jackson SE. The measurement of experienced burnout. J Occup Behav. 1981; 2: 99-113.

27. Bhana A, Rathod SD, Selohilwe O, Kathree T, Petersen I. The validity of the patient health questionnaire for screening depression in chronic care patients in primary health care in South Africa. BMC Psychiatry. 2015; 15: 118.

28. Makhubela M, Khumalo IP. Psychometric evaluation of the PHQ-9 in university students: Factorial validity and measurement equivalence across three African countries. Curr Psychol. 2023; 42: 18061-18069.

29. Kagee A, Bantjes J, Saal W, Sterley A. Predicting posttraumatic stress disorder caseness using the PTSD checklist for DSM-5 among patients receiving care for HIV. J Trauma Stress. 2022; 35: 13-21.

30. Padmanabhanunni A, Wiid C. From fear to fortitude: Differential vulnerability to PTSD among South African university students. Traumatology. 2022; 28: 167-174.

31. Storm K, Rothmann S. A psychometric analysis of the Maslach Burnout Inventory - general survey in the South African Police Service. S Afr J Psychol. 2003; 33: 219-226.

32. van der Ark LA. New developments in Mokken scale analysis in R. J Stat Softw. 2012; 48: 1-27.

33. R Development Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: http://www.R-project.org/.

34. O'Connor BP. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. Behav Res Methods Instrum Comput. 2000; 32: 396-402.

35. Revelle WR. PSYCH: Procedures for personality and psychological research [Internet]. Evanston, IL: Northwestern University; 2017. Available from: https://CRAN.r-project.org/package=psych.

36. Mokken RJ. A theory and procedure of scale analysis. The Hague, Netherlands: De Gruyter Mouton; 2011.

37. Koopman L, Zijlstra BJ, van der Ark LA. A two-step, test-guided Mokken scale analysis, for nonclustered and clustered data. Qual Life Res. 2022; 31: 25-36.

38. Wind SA. An Instructional module on mokken scale analysis. Educ Meas Issues Pract. 2017; 36: 50-66.

39. Sijtsma K, van der Ark LA. A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. Br J Math Stat Psychol. 2017; 70: 137-158.

40. Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. Psychol Bull. 1986; 99: 432-442.

41. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. Organ Res Methods. 2004; 7: 191-205.

42. Gaskin J. Statwiki CFA. Available from: http://statwiki.gaskination.com/index.php?title=CFA.

43. Jöreskog KG, Olsson UH, Wallentin FY. Confirmatory factor analysis (CFA). In: Multivariate analysis with LISREL. Springer; 2016. pp. 283-339.

44. Arbuckle JL. Amos 21.0 User's Guide. Mount Pleasant, SC: Amos Development Corporation; 2012.

45. Rodriguez A, Reise SP, Haviland MG. Applying bifactor statistical indices in the evaluation of psychological measures. J Pers Assess. 2016; 98: 223-237.

46. Paulsen J, BrckaLorenz A. Internal consistency statistics: FSSE Psychometric Portfolio [Internet]. Bloomington, IN: NSSE; 2017. Available from: https://nsse.indiana.edu/fsse/.

47. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh. 2007; 39: 155-164.

48. Tabachnick BG, Fidell LS, Ullman JB. Using multivariate statistics. Boston, MA: Pearson; 2013.

49. Floyd FJ, Widaman KF, Butcher JN. Factor analysis in the development and refinement of clinical assessment instruments. Psychol Assess. 1995; 7: 286-299.

50. Watts AL, Boness CL, Loeffelman JE, Steinley D, Sher KJ. Does crude measurement contribute to observed unidimensionality of psychological constructs? A demonstration with DSM–5 alcohol use disorder. J Abnorm Psychol. 2021; 130: 512.

51. Conklin LR, Cassiello-Robbins C, Brake CA, Sauer-Zavala S, Farchione TJ, Ciraulo DA, et al. Relationships among adaptive and maladaptive emotion regulation strategies and psychopathology during the treatment of comorbid anxiety and alcohol use disorders. Behav Res Ther. 2015; 73: 124-130.

52. Klein Breteler J, Ikani N, Becker ES, Spijker J, Hendriks G. Comorbid depression and treatment of anxiety disorders, OCD, and PTSD: Diagnosis versus severity. J Affect Disord. 2021; 295: 1005-1011.

53. Spielberger CD. Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.

54. Beck AT, Epstein N, Brown G, Steer R. An inventory for measuring clinical anxiety: Psychometric properties. J Consult Clin Psychol. 1998; 56; 893-897.