psych

MDPI

Article

Further Refinement of the Center for Epidemiological Studies Depression Scale-10: Complementary Evidence from Item Response Theory and Classical Test Theory

Anita Padmanabhanunni and Tyrone B. Pretorius *

Department of Psychology, University of the Western Cape, Cape Town 7530, South Africa; apadmana@uwc.ac.za
* Correspondence: tpretorius@uwc.ac.za

Abstract: The assessment of mental health, particularly depression, in university student populations is crucial for effective intervention and support. This study investigates the psychometric properties of the Center for Epidemiological Studies Depression Scale-10 (CES-D10) among 322 university students in the Western Cape province of South Africa, employing both classical test theory and item response theory. Participants were also assessed using the Perceived Stress Scale, the Satisfaction with Life Scale, and a short form of the Beck Hopelessness Scale. The results reveal satisfactory reliability indices for the CES-D10 based on Cronbach's alpha and McDonald's omega. However, Item 8 was identified as problematic across multiple metrics, including exploratory and confirmatory factor analysis (CFA) and Rasch analysis; therefore, the exclusion of this item is recommended for improved scale performance. The 9-item version displayed superior fit in the CFA and better construct validity than the 10-item scale. Scores on the CES-D10 were positively correlated with perceived stress and hopelessness and negatively correlated with life satisfaction, supporting the criterion-related validity of the scale. The study extends the psychometric validation literature of the CES-D10 by incorporating Rasch analysis, underscoring the benefits of using multiple statistical frameworks to achieve robust findings. These results have relevance for mental health assessment among university students in developing contexts, providing an evidence-based tool for early intervention.

Keywords: center for epidemiological studies depression scale-10; depression; rasch analysis; classical test theory; item response theory

Citation: Padmanabhanunni, A.; Pretorius, T.B. Further Refinement of the Center for Epidemiological Studies Depression Scale-10: Complementary Evidence from Item Response Theory and Classical Test Theory. *Psych* **2024**, *6*, 21–33. https://doi.org/10.3390/psych6010002

Academic Editor: Alexander Robitzsch

Received: 16 November 2023
Revised: 15 December 2023
Accepted: 19 December 2023
Published: 20 December 2023

1. Introduction

The mental health of university students remains a significant public health concern. Studies undertaken in both high- and low-to-middle-income countries over the past decade have confirmed that the prevalence of anxiety, stress, and depression is higher among university students than among other population groups [1,2]. The transition to university life can be challenging for many students as they encounter a new academic environment and expand their social network. For many students, attending university also necessitates living away from home for the first time. These adjustments—coupled with academic pressure, financial constraints, and increased independence—can contribute to heightened stress and susceptibility to mental health issues, particularly depression [3].

In a cross-sectional comparative study conducted during the COVID-19 pandemic, Naser and colleagues [1] reported that depression and anxiety were markedly higher among Jordanian university students than among healthcare workers and the general population. This discrepancy was attributed to the transition to remote learning in the context of significant resource constraints and students' worries about the impact on their future academic trajectory [1]. Vahedian-Azimi and colleagues [2] compared the severity of psychological distress among different Iranian population groups and found that depression and anxiety were higher among medical students than among the general population and medical staff. This discrepancy was attributed to students' exposure to

COVID-19 patients, limited access to adequate medical equipment, fears of infection, and concerns about errors in treatment [2]. A Spanish study [4] reported that university students had higher levels of depression, anxiety, and stress than staff, and this finding was attributed to the distinct implications of the pandemic for students. A range of systematic reviews and meta-analytic studies [3,5,6] have confirmed that depression is a significant mental health concern among students.

The consequences of untreated depression for university students can be profound. Depression has been linked to poor academic performance, diminished quality of life, social isolation, substance use, and suicidal ideation [7]. For this reason, systematic screening has been advocated to facilitate early detection and intervention. Such screenings require assessment measures that can discriminate between those who have depression and those who do not, thereby accurately identifying at-risk students [8]. The Center for Epidemiological Studies Depression Scale (CES-D) is one of the most widely used brief instruments for assessing depression. The scale was originally developed for research in general population epidemiological studies and primary care settings, but it has been extensively utilized as a screening measure for depressive symptoms among various populations [8,9] and groups with specific medical conditions [10]. The CES-D is a 20-item self-report scale that measures the presence and severity of symptoms of depression occurring over the past week [11].

The scale is often represented by the four-factor structure proposed by Radloff [11], which relies on principal component analysis and consists of depressed affect (e.g., feeling sad), positive affect (e.g., feeling hopeful), somatic or vegetative activity (e.g., sleep disturbances and appetite loss) and interpersonal problems (e.g., feeling lonely). These components were not developed based on the existing contemporary diagnostic criteria for depression. For this reason, some researchers have argued [12] that the scale may not be aligned with the contemporary diagnostic criteria for depression, which could potentially affect its reliability or validity.

Previous validation studies [13–15] have reported one-, two-, and three-factor structures of the scale, raising the question of whether the CES-D captures the same phenomena across different languages, cultures, and populations. For example, studies investigating the factor structure of the scale among adolescents in diverse cultural contexts, including China [15], Saudi Arabia [16], Malaysia [13], Korea [14], and the United States [17] have reported a two-factor structure [16], a three-factor structure [13–15], and a four-factor structure [17,18]. Kleim and colleagues [19] investigated the factor structure of the CES-D using non-parametric models in a representative German sample of adults and reported support for the two-factor and four-factor models. These findings suggest that the scale may not be psychometrically equivalent across different cultural settings. Apart from culture, certain dimensions of the scale, such as interpersonal problems, have not been found in studies [20–22] using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) in Asian populations. This finding could mean that the experience or expression of depression in these populations may not strongly involve interpersonal problems or may manifest in a way that is not captured by the existing scale. However, these differences in structures may also reflect choices made in terms of the method of factor extraction and rotation of factors.

There has also been extensive criticism of the psychometric properties of various items on the CES-D. For example, items targeting physical symptoms, such as "I felt that everything I did was an effort," have been reported [23] to potentially inflate CES-D results for chronic pain sufferers. In addition, robust gendered differences in the item "I had crying spells" have been found [24] to lead to an inflation of women's scores due to cultural norms pertaining to emotional expression rather than differences in depressive symptoms. The majority of factor analysis studies of the CES-D have relied on CFA and EFA [25].

Several abbreviated versions of the CES-D have been created for situations in which using the complete scale might be too cumbersome. For example, Levine [26] evaluated a seven-item version of the scale based on a sample of youth from the United States and reported that the short-form scale had moderately acceptable unidimensional validity

based on CFA and internal consistency reliability and that it was superior to the original scale. He and colleagues [27] developed a nine-item Chinese version of the scale and found it to be reliable and valid. Various ten-item versions of the CES-D have been developed as well [28,29]. The current study evaluates Andresen's [30] 10-item version, which has demonstrated comparable accuracy to the original CES-D scale in classifying those with depressive symptoms [31]. Using Andersen's [30] CES-D10, Zhang and colleagues [31] confirmed that the factors most associated with predicting outcomes in the CES-D20 were paralleled in the CES-D10. These researchers also confirmed that the short-form scale consisted of the same underlying factors of positive and negative affect found in the original development of the CES-D20. The scale's brevity and ease of administration make it a valuable instrument for large-scale epidemiological studies, longitudinal research, and clinical settings. The CES-D10 has been widely validated and has demonstrated good psychometric properties, including high internal consistency and test-retest reliability (see Yang and colleagues for an overview [32]).

Most short-form versions of the CES-D have been developed utilizing traditional test theory metrics, including item-total correlations and factor loadings [33–35]. However, one limitation of classical test theory (CTT) is its dependency on the specific sample being studied, as the performance of items may vary when applied to different populations [36]. In contrast, item response theory (IRT) methods, including Rasch analysis, offer more stable and generalizable estimates of item and test performance across diverse samples [36]. Rasch analysis, a specific form of IRT, allows for a nuanced understanding of individual item characteristics and their relationships to the underlying trait being measured (in this case, depressive symptoms). It provides additional information about item difficulty, discrimination, and how well each item fits within the overall construct of interest. Moreover, Rasch analysis is designed to evaluate the unidimensionality of a scale, which is crucial to ensure that the measure is capturing a single underlying construct [37].

It is, nevertheless, important to acknowledge that while IRT, with its emphasis on item parameter invariance, can offer advantages in terms of comparability across different samples or over time, particularly in diverse or longitudinal contexts, this does not imply that CTT is less effective in providing meaningful and generalizable results within the scope of its application. CTT, focusing on total test scores and reliability estimates like Cronbach's alpha, remains a robust and reliable tool for assessing test reliability and internal consistency in specific samples. Therefore, both IRT and CTT have unique strengths and are suited to different research needs, with the choice of method depending on the specific objectives and conditions of the study.

The current study was undertaken in South Africa, where a recent national survey of university students [38] revealed that the 30-day prevalence rate for major depressive disorder was 15.4%, a figure that surpasses the prevalence rate observed in the general population. This disparity underscores the necessity for assessment instruments that are tailored to effectively identify depressive symptoms in this group, who may experience unique stressors and challenges distinct from the general population. Several studies undertaken in sub-Saharan Africa have investigated the psychometric properties of the CES-D. For example, Baron and colleagues [33] conducted a validation study of the CES-D-10 among Zulu, Xhosa, and Afrikaans-speaking populations in South Africa. Using EFA, they found that a two-factor solution accounted for 42.7% and 46.7% of the variance in the CES-D-10 scores among Zulu and Xhosa-speaking samples, respectively. In contrast, a single-factor model was more suitable for the Afrikaans sample, explaining 51.6% of the variance. Furthermore, in the Zulu, Afrikaans, and Xhosa-speaking samples, the item-rest correlations for the CES-D-10 were consistently above 0.37, 0.40, and 0.60, respectively. However, items 5 ("I felt hopeful about the future") and 8 ("I was happy") were exceptions, as they consistently demonstrated the lowest item-rest correlations across these groups. Kilburn and colleagues [39] examined the performance of the CES-D10 among rural youth in Kenya, Malawi, Tanzania, Zambia, and Zimbabwe using CFA and EFA and found that a two-factor solution had the best model fit. Adams and colleagues [40] validated the 8-item

CES-D among an aging Shangaan-speaking South African sample using EFA and CFA and reported that a two-factor solution accounted for 50.6% of the variance in the overall scale. The reliability of the scale differed by gender in that men were less likely to endorse sleep problems and depressed mood than women. These studies suggest that while the CES-D is a valuable tool, further assessment of its psychometry properties is warranted, particularly in diverse cultural contexts.

This study aims to augment the existing body of evidence supporting the reliability, validity, and dimensionality of the CES-D10 by employing Rasch analysis in tandem with traditional CTT methods as well as factor analysis (EFA and CFA). The study aims to offer a comprehensive evaluation of the scale, considering both sample-specific characteristics and generalizable item properties. This dual-method approach aims to strengthen the empirical foundations of the CES-D, making it more robust and applicable across varied settings and populations.

2. Materials and Methods

2.1. Participants and Procedure

The study was conducted during the period 15 March 2022 to 19 August 2022. Participants were students (n = 322) at a university in the Western Cape province of South Africa. We used Google Forms to construct an electronic version of the instruments described in the Instruments section. This link was sent to a random sample of 1500 students selected using a random number generator, along with an invitation to participate in a study related to the mental health consequences of COVID-19. The 322 participants who accepted the invitation constituted a response rate of 21.5%. A post-hoc analysis indicated that the sample size of 322 leads to a 5.31% margin of error (95% confidence interval). The majority of the sample lived in an urban area (87.3%) and were women (77%). The mean age of the sample was 26.01 years (SD = 10.19). The university from which the participant sample was drawn caters to mature students who range in age from 17–63 years.

2.2. Instruments

Participants completed the following questionnaires: the CES-D10 [30], the Perceived Stress Scale (PSS) [41], the Satisfaction with Life Scale (SWLS) [42], and a short form of the Beck Hopelessness Scale (BHS-9) [43]. In addition, participants completed a brief demographic questionnaire.

The CES-D10 consists of 10 items to which participants respond on a 4-point scale, ranging from "rarely or none of the time" (0) to "most or all of the time" (3). An example item of the CES-D10 is: "I felt that everything I did was an effort." Zhang and colleagues [31] reported an alpha coefficient of 0.88 for the CES-D10 and found that the short-form version was as accurate as the 20-item version in screening for depression. In South Africa, Baron and colleagues [33] validated the CES-D10 in different language groups and reported reliability coefficients ranging from 0.69 to 0.89.

The PSS is a 10-item measure of the extent to which individuals appraise events in their lives as stressful. Responses to the 10 items are made on a 5-point scale ranging from "never" (0) to "very often" (4). An example item of the PSS is: "How often have you felt difficulties were piling up so high that you could not overcome them?" The author of the scale reported reliability coefficients ranging from 0.84 to 0.86. The relationship between life event scores and scores on the PSS served as evidence for the tool's validity [41]. In South Africa, Engelbrecht [44] confirmed the construct validity of the scale and reported a reliability coefficient of 0.93 for the PSS.

The SWLS is a 5-item measure of the cognitive component of subjective well-being, and it assesses participants' cognitive appraisals of how satisfied they are with their lives. Responses to the five items are made on a 7-point scale ranging from "strongly disagree" (1) to "strongly agree" (7). An example item of the SWLS is: "So far, I have gotten the important things I want in life." Diener and colleagues [42] reported a Cronbach's alpha of 0.87 in the first application of the scale, and the relationship between scores on the SWLS and other

measures of subjective well-being serves as evidence of the scale's validity. Pretorius and Padmanabhanunni [45] used CTT and IRT to demonstrate that the SWLS is a unidimensional, reliable, and valid measure of life satisfaction when used with schoolteachers in South Africa.

The BHS-9 is a short form of the original 20-item Beck Hopelessness Scale [46]. Responses to the nine items of the BHS-9 are made on a dichotomous true/false scale. An example item of the BHS-9 is: "All I can see ahead of me is unpleasantness rather than pleasantness." Balsamo and colleagues reported a Mokken scale reliability of 0.87 and an alpha coefficient of 0.86 for the 9-item version of the BHS in a sample of psychiatric in-patients [43].

2.3. Ethical Considerations

Ethical approval to undertake the study was granted by the Humanities and Social Sciences Ethics Committee of the University of the Western Cape (ethics reference number: HS22/2/9, February 2022). The study was conducted in accordance with the guidelines of the Declaration of Helsinki. Participants were assured of anonymity, and no identifying information was collected. Participation was entirely voluntary, and participants could withdraw from the survey at any time. Participants provided informed consent on the landing page of the electronic link.

2.4. Data Analyses

We used IBM SPSS for Windows version 28 and IBM SPSS Amos version 28 (IBM Corp., Armonk, NY, USA) for all CTT analyses. We used Pearson's r to examine the criterion-related validity of the CES-D10 by establishing the relationship between scores on the CES-D10 and measures of perceived stress, hopelessness, and life satisfaction. Both alpha and omega were used to examine the reliability of the CES-D10. In addition, the contribution of each item to the overall reliability of the scale was assessed by determining whether the exclusion of an item would lead to an increase in the scale's reliability. To examine the factor structure of the CES-D10, we used EFA (maximum likelihood with oblique promax rotation) and CFA. Prior to the EFA, we used Bartlett's test of sphericity and the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy to determine whether the data were suitable for factor analyses. A significant Bartlett's test and a KMO greater than 0.50 would indicate that there is sufficient common variance, and thus, factor analyses would be appropriate [47]. We also used parallel analysis to confirm the number of factors that accounted for the variability in the data. In parallel analysis, a random number of alternative data sets are generated, and eigenvalues obtained in these random data sets are compared to the actual eigenvalues obtained in the data set under analysis. Meaningful factors are those where the eigenvalue obtained for the real data set exceeds those obtained in the parallel data sets. Parallel analysis was conducted with SPSS syntax, which is freely available online [48]. With respect to CFA, we used the following fit indices [49] to examine the extent to which the one-factor model fits the data: chi-squared (χ^2: should be non-significant, which would indicate a perfect fit), goodness-of-fit index (GFI: should be greater than 0.90), confirmatory fit index (CFI: should be greater than 0.90), root-mean-square error of approximation (RMSEA: should be less than 0.08), and the Tucker–Lewis index (TLI: should be greater than 0.90) [49]. Akaike's information criterion (AIC), a model comparison index, was also used (the model with the lowest AIC value is considered the best fit). The factor loadings in CFA should be significant and greater than 0.40 [50].

Construct validity was assessed in terms of inter-item correlations and item-total correlations. Inter-item correlations should ideally be between 0.15 and 0.85, and average inter-item correlations should be between 0.15 and 0.50 [51]. Inter-item correlations that are too high (greater than 0.85) would indicate item redundancy, and those that are too low (less than 0.15) would indicate that the items do not assess the same content domain. Item-total correlations provide an indication of the extent to which each item contributes to

the measurement of the latent variable; ideally, these correlations should be greater than 0.50 [52].

A Rasch analysis was conducted using Winsteps version 5.1.4 [53]. In Rasch analysis, item and person separation indices provide further evidence of construct validity if the person separation index is greater than or equal to 2 together with a person reliability greater than or equal to 0.80 and an item separation index greater than or equal to 3, along with an item separation reliability greater than or equal to 0.80. Person separation and reliability assess whether an item can discriminate between high and low scorers on the latent variable, while item separation and reliability confirm whether an item difficulty hierarchy exists in the scale. The fit of items in the Rasch model is assessed using infit and outfit mean squares (MnSq), which should range from 0.5 to 1.5. Mean square values outside of this range are indicative of misfitting items. Finally, the dimensionality of the scale was assessed through a principal component analysis of the residuals after the presumed latent trait was excluded. If the eigenvalue of a possible second dimension (referred to as the first contrast) is greater than 2, it would indicate the existence of a possible second dimension.

3. Results

The intercorrelations, descriptive statistics, and reliability coefficients for the scales used in the study are reported in Table 1.

Table 1. Intercorrelations, descriptive statistics, and reliabilities of study variables.

Scale	1	2	3	4
1. CES-D10	—			
2. PSS	0.66 **	—		
3. SWLS	−0.53 **	−0.53 **	—	
4. BHS-9	0.50 **	0.47 **	−0.52 **	—
Mean	14.15	23.89	19.35	2.29
SD	6.77	6.28	7.06	2.45
Alpha	0.84	0.85	0.86	0.84
Omega	0.85	0.86	0.86	0.84

Note. CES-D10, Center for Epidemiological Studies Depression Scale-10; PSS, Perceived Stress Scale; SWLS, Satisfaction with Life Scale; BHS-9, Beck Hopelessness Scale-9. ** $p < 0.001$.

The findings in Table 1 indicate that all the scales demonstrated satisfactory reliability (α and ω: 0.84–0.86). Table 1 also shows a significant positive relationship between depression and perceived stress (r = 0.66, $p < 0.001$, large effect), as well as hopelessness (r = 0.50, $p < 0.001$, large effect). Conversely, there was a significant negative relationship between depression and life satisfaction (r = −0.53, $p < 0.001$, large effect).

The results of the KMO (0.86) and Bartlett's test ($p < 0.001$) indicated substantial correlations in the data; therefore, an EFA (maximum likelihood with oblique promax rotation) was used to examine the factor structure of the CES-D10. EFA extracted two factors which accounted for 42.13% and 10.82% of the variance, respectively. The results of the EFA are reported in Table 2.

Table 2 indicates that all the negative items load on factor 1. The two positive items load on factor 2. However, Items 12 ("I was happy") and 8 ("I felt hopeful about the future") cross-loaded on both factors. Parallel analysis, however, indicated that one factor was sufficient to account for the variability in the data as the eigenvalue of factor 1 in the real data set (4.21) was greater than the eigenvalue of the parallel data sets (1.37), while the eigenvalue of factor 2 (1.08) was lower than that obtained in the parallel data sets (1.24).

Consequently, we examined a one-factor model of the CES-D using CFA. However, CFA indicated that item 8 had an unacceptably low factor loading of less than 0.40. We thus examined two one-factor models of the CES-D10: one model included all 10 items, and the other excluded Item 8. The CFA models are presented in Figure 1.

Table 2. Results of the EFA for the CES-D10.

Item *	Factor Loadings	
	1	**2**
Factor 1		
1. I was bothered by things that usually don't bother me.	**0.68**	0.37
5. I had trouble keeping my mind on what I was doing.	**0.66**	0.37
6. I felt depressed.	**0.76**	0.53
7. I felt that everything I did was an effort.	**0.55**	0.21
10. I felt fearful.	**0.60**	0.42
11. My sleep was restless.	**0.58**	0.39
14. I felt lonely.	**0.66**	0.10
20. I could not get "going".	**0.72**	0.03
Factor 2		
8. I felt hopeful about the future.	0.31	**0.70**
12. I was happy.	0.45	**0.54**

* Item numbers are drawn from the original 20-item scale. Factor loadings in bold indicate primary loading.

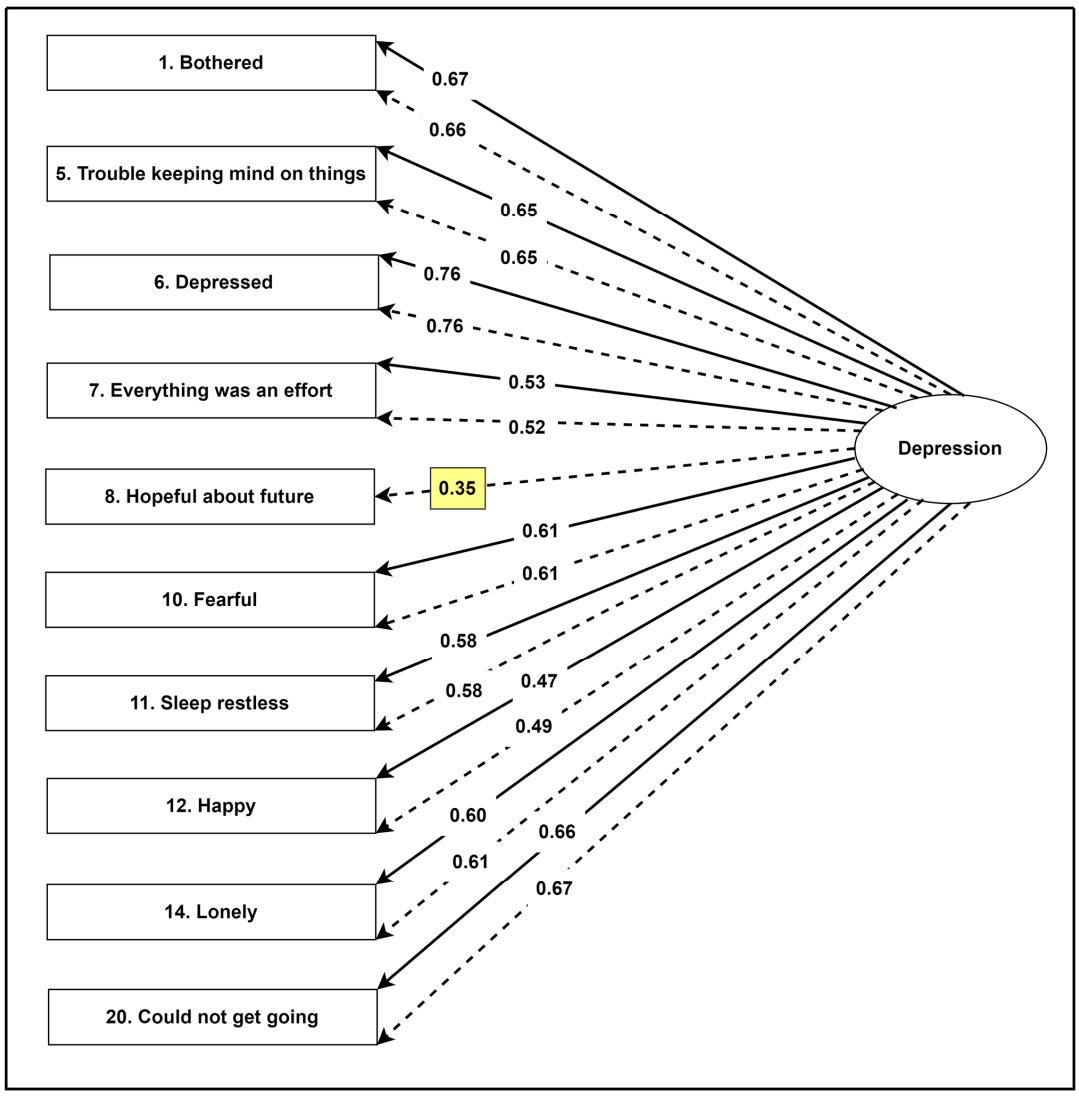

Figure 1. One-factor model of the CES-D10. Note. Solid line = parameter estimates for the 9-item model; dotted line = parameter estimates for the 10-item model; rectangles are observed measures; ellipse is a latent variable. All parameter estimates are standardized. The yellow rectangle reflects a factor loading < 0.40.

The fit indices demonstrate that both models fit the data to an acceptable degree (GFI, TLI, CFI > 0.90; RMSEA < 0.08). However, the fit indices for the two models indicate that the 9-item model (χ^2 = 65.56; p < 0.001; GFI = 0.96; TLI = 0.94; CFI = 0.95; RMSEA = 0.07) was marginally better fit to the data than the 10-item model (χ^2 = 96.86; p < 0.001; GFI = 0.94; TLI = 0.91; CFI = 0.93; RMSEA = 0.07). In addition, the model comparison index (AIC) indicated that the 9-item model (AIC = 101.56) had lower values than the 10-item model (136.86) and thus was the better fitting model. Figure 1 also reflects that all items had parameter estimates greater than 0.40 except Item 8, which had an unacceptably low factor loading of less than 0.40 (β = 0.35). As the EFA suggested the presence of a method factor (a distinction between positively- and negatively-worded items), we also examined a two-factor model of the CES-D10. The fit indices indicated that the two-factor model was an acceptable fit for the data (GFI, TLI, CFI > 0.90; RMSEA < 0.08). However, the AIC for the two-factor model (123.09) was higher than that of the 9- and 10-item unidimensional model, indicating that it was not the best-fitting model.

The item-level indices for the 10 items of the CES-D10 are reported in Table 3. These indices include inter-item correlation, item-total correlation, mean, standard deviation, reliability if item removed (CTT), and infit and outfit MnSq (Rasch).

Table 3. Item-level indices for the CES-D10.

Items and Indices	1	5	6	7	8	10	11	12	14	20
1. Bothered	—									
5. Keeping mind on things	0.47 **	—								
6. Depressed	0.52 **	0.49 **	—							
7. Everything was an effort	0.34 **	0.37 **	0.45 **	—						
8. Hopeful about future	0.17 *	0.19 **	0.31 **	0.07	—					
10. Fearful	0.43 **	0.36 **	0.46 **	0.28 **	0.23 **	—				
11. Sleep restless	0.39 **	0.34 **	0.42 **	0.32 **	0.23 **	0.48 **	—			
12. Happy	0.31 **	0.27 **	0.38 **	0.17 *	0.36 **	0.33 **	0.22 **	—		
14. Lonely	0.41 **	0.35 **	0.49 **	0.30 **	0.27 **	0.29 **	0.31 **	0.33 **	—	
20. Could not get going	0.36 **	0.50 **	0.45 **	0.37 **	0.31 **	0.40 **	0.41 **	0.33 **	0.47 **	—
Mean	1.26	1.55	1.38	1.71	1.30	1.10	1.51	1.43	1.49	1.43
SD	0.98	1.00	1.14	1.09	1.06	1.03	1.10	0.94	1.14	1.04
α if item removed	0.82	0.83	0.81	0.84	0.85	0.83	0.83	0.84	0.83	0.82
Item-total correlation	0.68 **	0.67 **	0.78 **	0.58 **	0.49 **	0.66 **	0.65 **	0.56 **	0.66 **	0.72 **
Infit MnSq	0.80	0.84	0.85	1.23	1.36	0.97	1.08	0.96	1.12	0.81
Outfit MnSq	0.84	0.90	0.80	1.34	1.60	0.94	1.09	1.01	1.12	0.78

** p < 0.001, * p < 0.01.

The findings in Table 3 indicate that the inter-item correlations were all significant except for the correlation between Items 8 and 7. Generally, the findings reflect low to medium correlations between Item 8 and the rest of the items of the CES-D10. The range of inter-item correlations from 0.22 to 0.52 reflects low to moderate correlations, which indicate that there is no item redundancy. Item 8 was the only item that would lead to a marginal increase in the total scale reliability (α = 0.85) if excluded. The average inter-item correlation was 0.21. All the item-total correlations were significant and greater than 0.50. However, Item 8 had the lowest correlation with the scale total (r = 0.49) and a medium effect size, whereas the correlation coefficients for the rest of the items represented a large effect size (r > 0.50). In terms of the Rasch infit and outfit MnSq, all items except Item 8 had optimal MnSq values (0.50–1.50). The outfit MnSq of Item 8 exceeded 1.50 (MnSq = 1.60), which is less than optimal.

The scale-level indices for the CES-D10 are reported in Table 4. These findings include the CTT and Rasch indices for the 9- and 10-item versions of the scale.

Table 4. Scale-level indices for two versions of the short CES-D.

Index	CESD-10	CES-D9	Acceptable Value
Alpha	0.84	0.85	>0.70
Omega	0.85	0.85	>0.70
Item separation reliability *	0.92	0.93	≥0.80
Item separation index *	3.32	3.54	≥3.00
Person separation reliability *	0.80	0.80	≥0.80
Person separation index *	1.98	2.00	≥2.00
First contrast *,[1]	1.57	1.40	<2.00

* Rasch indices; [1] eigenvalue of the unexplained variance in the first contrast.

The findings in Table 4 reflect that the two scale versions had similar and acceptable indices at the scale level, except for the person separation index of the 10-item version, which was lower than the acceptable level of 2.00.

4. Discussion

The development of the original 10-item version of the CES-D10 [30] and subsequent validity studies thereof (for example, [33–35]) have largely been carried out from the perspective of CTT. However, this method is sample-dependent; therefore, the results may vary from sample to sample. IRT is theoretically a less sample-dependent approach and thus produces indices that remain relatively stable across samples. The current study sought to provide further evidence for the reliability, validity, and dimensionality of the CES-D10 from the perspectives of CTT and IRT.

First, the reliability of the CES-D10 scale in the current study was satisfactory, as determined using Cronbach's alpha and McDonald's omega. Second, Item 8 ("I felt hopeful about the future") proved problematic across several indices. It cross-loaded on two factors in EFA, with the most salient loading on a second factor rather than the primary factor. The loading of item 8 on a second factor does not suggest that the CES-D10 should be considered multidimensional since a single-item factor would indicate that the single item and the latent factor that it defines share 100% of the variance, making them empirically indistinguishable. In this regard, Costello and Osborne [54] recommend that a factor should only be retained if it has at least three items with factor loadings greater than 0.32. In CFA, the factor loading for Item 8 was less than the recommended 0.40. CFA fit indices showed that, while both a 9-item (excluding Item 8) and a 10-item version of the scale were good fits for the data, the 9-item version demonstrated a superior fit. Item 8 had low correlations and one non-significant correlation with the other nine items, which indicates that the items do not share the same content domain [55]. In addition, the correlation of Item 8 with the total scale was less than the recommended 0.50, which indicates that the contribution of this item to the measurement of the latent variable was less than that of the other nine items. Item 8 was also the only item that, if excluded, would marginally increase the reliability of the total scale. Rasch analysis also indicated that this item is a misfit in the scale, as the outfit MnSq value for Item 8 was greater than 1.50.

In terms of criterion-related validity, we found that depression was positively associated with perceived stress and hopelessness and negatively associated with life satisfaction. These findings are consistent with the theoretical expectations, as higher levels of stress and hopelessness and lower levels of life satisfaction have previously been associated with higher levels of depression [56–58].

The 9- and 10-item versions of the CES-D had roughly similar indices at the scale level (i.e., reliability, item separation indices, and dimensionality), suggesting that either version would be acceptable for use. However, the 10-item version had a low person separation index. This finding indicates that the construct validity of the 10-item version

is compromised by Item 8 because it renders the scale unable to discriminate between individuals with high and low measures of depression.

The study findings have several practical and theoretical implications for university settings within developing contexts. University counseling centers or mental health services are often under-resourced. This situation, coupled with the high stress levels and increased vulnerability to mental health issues observed among university students, necessitates the availability of a brief, reliable, and valid tool for the early detection of depressive symptomology. Employing a reliable diagnostic tool can facilitate early interventions, which could impact students' mental health and well-being. Depression has a notable impact on academic performance and overall well-being. For universities in developing contexts—where students may already grapple with challenges such as financial constraints, limited access to resources, and societal pressures—a reliable tool like the 9-item CES-D can assist mental health professionals in identifying students in distress. Mental health professionals would then be empowered to offer necessary support to these students, which could, in turn, improve academic outcomes and reduce dropout rates.

Theoretically, the study emphasizes the importance of employing both CTT and IRT for a comprehensive and generalizable assessment of the reliability and validity of the scale. The identified problematic nature of Item 8, as highlighted in the study, underscores the importance of ensuring the cultural relevance of psychometric tools. The low person separation index for the 10-item version has implications for the construct validity of the scale. In the context of developing countries, where the CES-D10 may be translated into different languages or modified to account for local cultural norms, understanding the limitations of particular items could guide adaptations to ensure the scale remains reliable and valid.

In summary, the refined 9-item version of the CES-D could serve as an effective tool for mental health assessment among university students in developing contexts, promoting mental health interventions and academic support services. By using a reliable and valid tool, mental health professionals can be more confident in diagnoses and subsequent treatment plans for university students.

In sum, the current study makes important contributions on both theoretical and practical fronts. Theoretical contributions are exemplified by our comprehensive approach to evaluating the psychometric properties of the CES-D10 within a specific cultural context. We highlight the importance of considering cultural relevance and adaptation in the assessment of depressive symptomology, particularly in developing regions. The identification of item 8 as problematic underscores the need for a nuanced approach when utilizing widely used scales across diverse populations. From a practical perspective, our findings provide mental health professionals, educators, and policymakers with a valuable tool for early detection and intervention of depressive symptoms among university students. The recommendation to exclude item 8 enhances the practical utility of the CES-D10, while our criterion-related validity results emphasize its real-world applicability. Overall, this study bridges theoretical insights with practical implications, contributing to a more comprehensive understanding of mental health assessment and support in university settings within developing contexts.

The study has certain limitations. The study is cross-sectional and, therefore, captures data at a single point in time. This approach does not allow for assessment of the instrument's sensitivity to change over time, which is particularly important for tracking the effectiveness of interventions. The study relies on a student sample from one institution located in a specific geographic area, which may impact the generalizability of the findings. Although this study provides valuable insights into the psychometric properties of the CES-D10, it does not explicitly account for cultural variations that might influence the interpretation of specific items, including Item 8. However, the exclusion of item 8 cannot be solely grounded in the findings of the current study. The problematic nature of item 8 should be replicated in other samples and validated with different statistical methods, such as non-parametric IRT, to further substantiate its potential limitations. The study

employed both CTT and IRT; however, using additional methods such as longitudinal designs or multi-group comparisons could offer more comprehensive insights regarding the reliability and validity of the instrument. While the study did establish criterion-related validity with respect to stress, hopelessness, and life satisfaction, it did not examine how the CES-D10 correlates with clinical diagnosis of depression or other mental health conditions. An examination of such correlations might offer a more direct validation of the tool's utility in clinical settings.

5. Conclusions

This study offers a nuanced examination of the psychometric properties of the CES-D10 scale, contributing valuable insights through the lenses of both CTT and IRT. The findings suggest that the 9-item version of the CES-D may be more reliable and valid for assessing depression than the 10-item version, particularly among university student populations in developing contexts. This enhanced version has the potential to serve as a more precise tool for mental health practitioners and researchers, thereby facilitating improved mental health outcomes for students. The identification of limitations related to individual items of the scale and the construct validity of the scale sets the foundation for future research aimed at refining the measurement of depression among diverse populations. Given the urgency of addressing mental health challenges in academic settings, particularly in resource-constrained environments, the implications of this study are timely.

Author Contributions: Conceptualization, A.P. and T.B.P.; Formal analysis, T.B.P.; Investigation, A.P. and T.B.P.; Methodology, T.B.P.; Project administration, A.P.; Supervision, A.P.; Visualization, T.B.P.; Writing – original draft, A.P. and T.B.P.; Writing – review & editing, A.P. and T.B.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the South African Medical Research Council (SIR grant).

Institutional Review Board Statement: Ethical approval to undertake the study was granted by the Humanities and Social Sciences Ethics Committee of the University of the Western Cape (ethics reference number: HS22/2/9, February 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Naser, A.Y.; Dahmash, E.Z.; Al-Rousan, R.; Alwafi, H.; Alrawashdeh, H.M.; Ghoul, I.; Abidine, A.; Bokhary, M.A.; Al-Hadithi, H.T.; Ali, D.; et al. Mental health status of the general population, healthcare professionals, and university students during 2019 coronavirus disease outbreak in Jordan: A cross-sectional study. *Brain Behav.* **2020**, *10*, e01730. [CrossRef] [PubMed]
2. Vahedian-Azimi, A.; Moayed, M.S.; Rahimibashar, F.; Shojaei, S.; Ashtari, S.; Pourhoseingholi, M.A. Comparison of the severity of psychological distress among four groups of an Iranian population regarding COVID-19 pandemic. *BMC Psychiatry* **2020**, *20*, 402. [CrossRef] [PubMed]
3. Satinsky, E.N.; Kimura, T.; Kiang, M.V.; Abebe, R.; Cunningham, S.; Lee, H.; Lin, X.; Liu, C.H.; Rudan, I.; Sen, S.; et al. Systematic review and meta-analysis of depression, anxiety, and suicidal ideation among Ph.D. students. *Sci. Rep.* **2021**, *11*, 14370. [CrossRef] [PubMed]
4. Odriozola-González, P.; Planchuelo-Gómez, Á.; Irurtia, M.J.; de Luis-García, R. Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a Spanish university. *Psychiatry Res.* **2020**, *290*, 113108. [CrossRef] [PubMed]
5. Sarokhani, D.; Delpisheh, A.; Veisani, Y.; Sarokhani, M.T.; Esmaeli Manesh, R.; Sayehmiri, K. Prevalence of Depression among University Students: A Systematic Review and Meta-Analysis Study. *Depress. Res. Treat.* **2013**, *2013*, 373857. [CrossRef]
6. Li, W.; Zhao, Z.; Chen, D.; Peng, Y.; Lu, Z. Prevalence and associated factors of depression and anxiety symptoms among college students: A systematic review and meta-analysis. *J. Child. Psychol. Psychiatry* **2022**, *63*, 1222–1230. [CrossRef] [PubMed]
7. De Luca, S.M.; Franklin, C.; Yueqi, Y.; Johnson, S.; Brownson, C. The Relationship Between Suicide Ideation, Behavioral Health, and College Academic Performance. *Community Ment. Health J.* **2016**, *52*, 534–540. [CrossRef]
8. Vilagut, G.; Forero, C.G.; Barbaglia, G.; Alonso, J. Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. *PLoS ONE* **2016**, *11*, e0155431. [CrossRef]

9. Dozeman, E.; van Schaik, D.J.F.; van Marwijk, H.W.J.; Stek, M.L.; van der Horst, H.E.; Beekman, A.T.F. The center for epidemiological studies depression scale (CES-D) is an adequate screening instrument for depressive and anxiety disorders in a very old population living in residential homes. *Int. J. Geriatr. Psychiatry* **2011**, *26*, 239–246. [CrossRef]
10. Kellahan, S.R.; Huang, X.; Lew, D.; Xian, H.; Eisen, S.; Kim, A.H.J. Depressed Symptomatology in Systemic Lupus Erythematosus Patients. *Arthritis Care Res.* **2023**, *75*, 749–757. [CrossRef]
11. Radloff, L.S. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1977**, *1*, 385–401. [CrossRef]
12. Cosco, T.D.; Lachance, C.C.; Blodgett, J.M.; Stubbs, B.; Co, M.; Veronese, N.; Wu, Y.-T.; Prina, A.M. Latent structure of the Centre for Epidemiologic Studies Depression Scale (CES-D) in older adult populations: A systematic review. *Aging Ment. Health* **2020**, *24*, 700–704. [CrossRef] [PubMed]
13. Ghazali, S.R.; Elklit, A.; Balang, R.V.; Chen, Y.Y. Concurrent validity and exploratory factor analysis of the Malay version of Center for Epidemiologic Studies-Depression Scale (Cesd) among Malaysian adolescents. *Psychiatry* **2016**, *17*, 71–78.
14. Heo, E.-H.; Choi, K.-S.; Yu, J.-C.; Nam, J.-A. Validation of the Center for Epidemiological Studies Depression Scale among Korean Adolescents. *Psychiatry Investig.* **2018**, *15*, 124–132. [CrossRef] [PubMed]
15. Dou, D.; Shek, D.T.L.; Zhu, X.; Zhao, L. Dimensionality of the Chinese CES-D: Is It Stable across Gender, Time, and Samples? *Int. J. Environ. Res. Public Health* **2021**, *18*, 11818. [CrossRef]
16. Dardas, L.A.; Xu, H.; Shawashreh, A.; Franklin, M.S.; Wainwright, K.; Pan, W. Psychometric Properties and Factor Structure of the Center for Epidemiologic Studies Depression Scale in a National Sample of Arab Adolescents. *Issues Ment. Health Nurs.* **2019**, *40*, 437–443. [CrossRef]
17. Roberts, R.E.; Andrews, J.A.; Lewinsohn, P.M.; Hops, H. Assessment of Depression in Adolescents Using the Center for Epidemiologic Studies Depression Scale. *Psychol. Assess.* **1990**, *2*, 122–128. [CrossRef]
18. Skriner, L.C.; Chu, B.C. Cross-Ethnic Measurement Invariance of the SCARED and CES-D in a Youth Sample. *Psychol. Assess.* **2014**, *26*, 332–337. [CrossRef]
19. Kliem, S.; Beller, J.; Tibubos, A.N.; Zenger, M.; Schmalbach, B.; Brähler, E. A reanalysis of the center for epidemiological studies depression scale (CES-D) using non-parametric item response theory. *Psychiatry Res.* **2020**, *290*, 113132. [CrossRef]
20. Jiang, L.; Wang, Y.; Zhang, Y.; Li, R.; Wu, H.; Li, C.; Wu, Y.; Tao, Q. The Reliability and Validity of the Center for Epidemiologic Studies Depression Scale (CES-D) for Chinese University Students. *Front. Psychiatry* **2019**, *10*, 315. [CrossRef]
21. Wang, M.; Armour, C.; Wu, Y.; Ren, F.; Zhu, X.; Yao, S. Factor Structure of the CES-D and Measurement Invariance Across Gender in Mainland Chinese Adolescents. *J. Clin. Psychol.* **2013**, *69*, 966–979. [CrossRef] [PubMed]
22. Zhu, X.; Shek, D.T.L.; Dou, D. Factor structure of the Chinese CES-D and invariance analyses across gender and over time among Chinese adolescents. *J. Affect. Disord.* **2021**, *295*, 639–646. [CrossRef] [PubMed]
23. Ohayon, M.M.; Schatzberg, A.F. Using chronic pain to predict depressive morbidity in the general population. *Arch. Gen. Psychiatry* **2003**, *60*, 39–47. [CrossRef] [PubMed]
24. Rivera-Medina, C.L.; Caraballo, J.N.; Rodríguez-Cordero, E.R.; Bernal, G.; Dávila-Marrero, E. Factor Structure of the CES-D and Measurement Invariance Across Gender for Low-Income Puerto Ricans in a Probability Sample. *J. Consult. Clin. Psychol.* **2010**, *78*, 398–408. [CrossRef] [PubMed]
25. Carleton, R.N.; Thibodeau, M.A.; Teale, M.J.N.; Welch, P.G.; Abrams, M.P.; Robinson, T.; Asmundson, G.J.G. The center for epidemiologic studies depression scale: A review with a theoretical and empirical examination of item content and factor structure. *PLoS ONE* **2013**, *8*, e58067. [CrossRef] [PubMed]
26. Levine, S.Z. Evaluating the seven-item Center for Epidemiologic Studies Depression Scale short-form: A longitudinal US community study. *Soc. Psychiatry Psychiatr. Epidemiol.* **2013**, *48*, 1519–1526. [CrossRef] [PubMed]
27. He, J.; Chen, Z.; Guo, F.; Zhang, J.; Yang, Y.; Wang, Q. A short Chinese version of center for epidemiologic studies depression scale. *Chin. J. Behav. Med. Brain Sci.* **2013**, *12*, 1133–1136.
28. Cole, J.C.; Rabin, A.S.; Smith, T.L.; Kaufman, A.S. Development and Validation of a Rasch-Derived CES-D Short Form. *Psychol. Assess.* **2004**, *16*, 360–372. [CrossRef]
29. Lee, S.W.; Stewart, S.M.; Byrne, B.M.; Wong, J.P.S.; Ho, S.Y.; Lee, P.W.H.; Lam, T.H. Factor Structure of the Center for Epidemiological Studies Depression Scale in Hong Kong Adolescents. *J. Personal. Assess.* **2008**, *90*, 175–184. [CrossRef]
30. Andresen, E.M.; Malmgren, J.A.; Carter, W.B.; Patrick, D.L. Screening for Depression in Well Older Adults: Evaluation of a Short Form of the CES-D. *Am. J. Prev. Med.* **1994**, *10*, 77–84. [CrossRef]
31. Zhang, W.; O'Brien, N.; Forrest, J.I.; Salters, K.A.; Patterson, T.L.; Montaner, J.S.G.; Hogg, R.S.; Lima, V.D. Validating a shortened depression scale (10 item CES-D) among HIV-positive people in British Columbia, Canada. *PLoS ONE* **2012**, *7*, e40793. [CrossRef] [PubMed]
32. Yang, W.; Xiong, G.; Garrido, L.E.; Zhang, J.X.; Wang, M.-C.; Wang, C. Factor Structure and Criterion Validity Across the Full Scale and Ten Short Forms of the CES-D Among Chinese Adolescents. *Psychol. Assess.* **2018**, *30*, 1186–1198. [CrossRef] [PubMed]
33. Baron, E.C.; Davies, T.; Lund, C. Validation of the 10-item centre for epidemiological studies depression scale (CES-D-10) in Zulu, Xhosa and Afrikaans populations in South Africa. *BMC Psychiatry* **2017**, *17*, 6. [CrossRef] [PubMed]
34. Lee, A.E.Y.; Chokkanathan, S. Factor structure of the 10-item CES-D scale among community dwelling older adults in Singapore. *Int. J. Geriatr. Psychiatry* **2008**, *23*, 592–597. [CrossRef] [PubMed]

35. Mohebbi, M.; Nguyen, V.; McNeil, J.J.; Woods, R.L.; Nelson, M.R.; Shah, R.C.; Storey, E.; Murray, A.M.; Reid, C.M.; Kirpach, B.; et al. Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. *Gen. Hosp. Psychiatry* **2018**, *51*, 118–125. [CrossRef] [PubMed]
36. Meijer, R.R.; Sijtsma, K.; Smid, N.G. Theoretical and empirical comparison of the Mokken and the Rasch approach to IRT. *Appl. Psychol. Meas.* **1990**, *14*, 283–298. [CrossRef]
37. Sijtsma, K.; Meijer, R.R.; Andries van der Ark, L. Mokken scale analysis as time goes by: An update for scaling practitioners. *Pers. Individ. Dif.* **2011**, *50*, 31–37. [CrossRef]
38. Bantjes, J.; Kessler, M.; Lochner, C.; Breet, E.; Bawa, A.; Roos, J.; Davids, C.; Muturiki, M.; Kessler, R.C.; Stein, D.J. The mental health of university students in South Africa: Results of the national student survey. *J. Affect. Disord.* **2023**, *321*, 217–226. [CrossRef]
39. Kilburn, K.; Prencipe, L.; Hjelm, L.; Peterman, A.; Handa, S.; Palermo, T. Examination of performance of the Center for Epidemiologic Studies Depression Scale Short Form 10 among African youth in poor, rural households. *BMC Psychiatry* **2018**, *18*, 201. [CrossRef]
40. Adams, L.B.; Farrell, M.; Mall, S.; Mahlalela, N.; Berkman, L. Dimensionality and differential item endorsement of depressive symptoms among aging Black populations in South Africa: Findings from the HAALSI study. *J. Affect. Disord.* **2020**, *277*, 850–856. [CrossRef]
41. Cohen, S. Perceived stress in a probability sample of the United States. In *The Social Psychology of Health*; Spacapan, S., Oskamp, S., Eds.; Sage Publications: Thousand Oaks, CA, USA, 1988; pp. 31–67.
42. Diener, E.D.; Emmons, R.A.; Larsen, R.J.; Griffin, S. The satisfaction with life scale. *J. Personal. Assess.* **1985**, *49*, 71–75. [CrossRef] [PubMed]
43. Balsamo, M.; Carlucci, L.; Innamorati, M.; Lester, D.; Pompili, M. Further Insights Into the Beck Hopelessness Scale (BHS): Unidimensionality Among Psychiatric Inpatients. *Front. Psychiatry* **2020**, *11*, 727. [CrossRef] [PubMed]
44. Engelbrecht, M.C. Construct validity and reliability of the perceived stress scale for nursing students in South Africa. *Curationis* **2022**, *45*, 2276. [CrossRef] [PubMed]
45. Pretorius, T.B.; Padmanabhanunni, A. Assessing the cognitive component of subjective well-being: Revisiting the satisfaction with life scale with classical test theory and item response theory. *Afr. J. Psychol. Assess.* **2022**, *4*, 9. [CrossRef]
46. Beck, A.T.; Weissman, A.; Lester, D.; Trexler, L. The measurement of pessimism: The hopelessness scale. *J. Consult. Clin. Psychol.* **1974**, *42*, 861–865. [CrossRef] [PubMed]
47. Shkeer, A.S.; Awang, Z. Exploring the items for measuring the marketing information system construct: An exploratory factor analysis. *Int. Rev. Manag. Mark.* **2019**, *9*, 87–97. [CrossRef]
48. O'connor, B.P. 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. [CrossRef]
49. Kline, R.B. *Principles and Practice of Structural Equation Modeling*; Guilford Publications: New York, NY, USA, 2015.
50. Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. *Multivariate Data Analysis*; Prentice Hall: Englewood Cliffs, NJ, USA, 2010.
51. Paulsen, J.; BrckaLorenz, A. *Internal Consistency Statistics*; FSSE Psychometric Portfolio; Indiana University: Bloomington, IN, USA, 2017.
52. DeVon, H.A.; Block, M.E.; Moyle-Wright, P.; Ernst, D.M.; Hayden, S.J.; Lazzara, D.J.; Savoy, S.M.; Kostas-Polston, E. A Psychometric Toolbox for Testing Validity and Reliability. *J. Nurs. Scholarsh.* **2007**, *39*, 155–164. [CrossRef]
53. Linacre, J.M. Winsteps®Rasch Measurement Computer Program (Version 5.6.0); Computer Software. Available online: https://www.winsteps.com/ (accessed on 11 August 2023).
54. Costello, A.B.; Osborne, J. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* **2005**, *10*, 7. [CrossRef]
55. Piedmont, R.L. Inter-item Correlations. In *Encyclopedia of Quality of Life and Well-Being Research*; Michalos, A.C., Ed.; Springer: Dordrecht, The Netherlands, 2014; pp. 3303–3304.
56. Charles, N.E.; Strong, S.J.; Burns, L.C.; Bullerjahn, M.R.; Serafine, K.M. Increased mood disorder symptoms, perceived stress, and alcohol use among college students during the COVID-19 pandemic. *Psychiatry Res.* **2021**, *296*, 113706. [CrossRef]
57. Demirtas, A.S.; Yildiz, B. Hopelessness and perceived stress: The mediating role of cognitive flexibility and intolerance of uncertainty. *Dusunen Adam J. Psychiatry Neurol. Sci.* **2019**, *32*, 259. [CrossRef]
58. Haikalis, M.; Doucette, H.; Meisel, M.K.; Birch, K.; Barnett, N.P. Changes in College Student Anxiety and Depression from Pre- to During-COVID-19: Perceived Stress, Academic Challenges, Loneliness, and Positive Perceptions. *Emerg. Adulthood* **2022**, *10*, 534–545. [CrossRef] [PubMed]