Mediterranean Journal of Clinical Psychology

ISSN 2282-1619

MJCP

Volume 11, n 2, 2023

Clinical Psychology

Clinimetric properties of the 21-item Depression, Anxiety and Stress Scales (DASS-21)

*Giovanni Mansueto [1, 2], Danilo Carrozzino [3], Kaj Sparle Christensen [4], Chiara Patierno [3], Fiammetta Cosci [1, 5 *]*

Abstract

Background: The Depression, Anxiety and Stress Scale-21 (DASS-21) is a widely used patient-reported outcome measure. While psychometric properties of the DASS-21 have been studied, insufficient attention has been devoted to the assessment of its clinimetric properties. This study verified the clinimetric properties of the Italian version of the DASS-21 according to Clinimetric Patient-Reported Outcome Measures (CLIPROM) criteria.

Methods: This is a cross-sectional study involving 951 university students from April to September 2020. Participants were asked to fill in the DASS-21 via an online survey. Participation was voluntary. Item Response Theory (IRT) models were used to test dimensionality, scalability, and sensitivity of DASS-21.

Results: IRT analyses showed that the DASS-21 total score was a multidimensional measure of psychological distress. Fit to the Rasch model was achieved after excluding five misfitting items and adjusting the sample size, resulting in a 16-item version of the DASS-21. The 16-item version entailed the clinimetric property of sensitivity but included inter-correlated items. Brief versions of the DASS-21 subscales of depression, anxiety, and stress, which did not include locally dependent items, fitted the Rasch model expectations, and had an acceptable unidimensionality and scalability, were identified.

Conclusion: The 16-item version of the DASS-21 may be used as an overall indicator of dysthymia and should be supplemented with the brief versions of the depression, anxiety, and stress subscales, which were found to be valid clinimetric indices.

[1] Department of Health Sciences, University of Florence, Florence, Italy

[2] Department of Psychology, Sigmund Freud University, Milan, Italy

[3] Department of Psychology "Renzo Canestrari", Alma Mater Studiorum University of Bologna, Bologna, Italy

[4] Research Unit for General Practice and Section for General Medical Practice, Department of Public Health, Aarhus University, Aarhus, Denmark

[5] Department of Psychiatry & Neuropsychology, Maastricht University, Maastricht, the Netherlands

E-mail corresponding author: fiammetta.cosci@unifi.it

Keywords:
Anxiety; Assessment; Clinimetrics; Depression; Stress; Validity; Clinical psychology.

Received: 23 May 2023
Accepted: 25 July 2023
Published: 31 August 2023

Citation: Mansueto, G., Carrozzino, D., Christensen, K.S., Patierno, C., Cosci, F. (2023). Clinimetric properties of the 21-item Depression, Anxiety and Stress Scales (DASS-21). *Mediterranean Journal of Clinical Psychology 11*(2). https://doi.org/10.13129/2282-1619/mjcp-3795

1. Introduction

The Depression, Anxiety and Stress Scales (DASS) is a patient-reported outcome measure developed by Lovibond and Lovibond (1993) to assess the full range of core symptoms of depression and anxiety (Lovibond & Lovibond, 1995a). Its original version had 42 items organized in three subscales (Lovibond & Lovibond, 1993). A shorter 21-item version (DASS-21) was proposed with three 7-item subscales assessing depression, anxiety, and subjective tension or stress over the previous week (Lovibond & Lovibond, 1995b). The DASS, and its various versions, has been widely used in psychopathological assessment of clinical (e.g., Sakakibara et al., 2009; Smith et al., 2021) and non-clinical populations (e.g., Sekhar et al., 2021), with a strong increase of use during the Covid-19 era (e.g., Mboua et al., 2021).

The psychometric properties of the DASS-21 have been largely studied (Lee et al., 2019; Yeung et al., 2020). Good internal consistency (Antony et al., 1998; Bottesi et al., 2015; Brown et al., 1997; Clara et al., 2001; Lee et al., 2019; Lovibond & Lovibond, 1995a; Page et al., 2007), temporal stability (Bottesi et al., 2015), convergent validity (Antony et al., 1998; Bottesi et al., 2015; Brown et al., 1997; Crawford & Henry, 2001; Lovibond & Lovibond, 1995a), divergent validity (Bottesi et al., 2015), and construct validity (Mahmoud et al., 2010) were found. Criterion validity showed to be good only for the depression subscale (Lee et al., 2019; Moya et al., 2022). A two-factor model was the one better fitting (Lee et al., 2019; Yeung et al., 2020).

However, classical psychometric principles have a limited utility in clinical practice (Charlson et al., 2022; Cosci, 2021; Fava et al., 2004; Fava, 2022). For psychometric scales, a number of homogeneous items for assessing a single condition may be important, but for measuring a phenomenon like for instance over time change of clinical manifestations, the index cannot be homogeneous and redundant (Wright & Feinstein, 1992). In order to overcome the major limitations of classical psychometrics, Alvan R. Feinstein (Feinstein, 1982; Feinstein, 1983; Feinstein, 1987) proposed in 1980s clinimetrics, the science of innovative methods for clinical assessment. This scientific domain is aimed at evaluating a number of measurement properties and clinical issues not usually captured by the traditional psychometric model (Fava et al., 2012). Such an approach is nowadays the referral for clinical measurements (Cosci, 2021; Fava, 2022). Clinimetric Patient-Reported Outcome Measures (CLIPROM) criteria (Carrozzino et al., 2021a) were proposed to guide the development and validation process of self-reported tools.

According to CLIPROM criteria (Carrozzino et al., 2021a), relevant clinimetric properties of a self-rating scale are sensitivity, unidimensionality, and scalability. Sensitivity (Carrozzino et al., 2021a; Fava et al., 2018) allows to discriminate between subjects belonging to different categories, for instance patients and healthy controls, or to discriminate between different

degrees of severity of a symptom (e.g., mild, moderate, severe depression) or different clinical phenomena (e.g., melancholic vs non melancholic depression). Unidimensionality and scalability allow to measure construct validity, that is whether each item of a rating scale gives unique clinical information, belongs to an underlying construct, and whether the total score is a valid measure of the severity of the assessed dimension (e.g., depression) (Bech, 2004; Bech, 2012; Carrozzino et al., 2021a; Fava et al., 2018).

According to CLIPROM criteria (Carrozzino et al., 2021a), clinimetric properties should be tested via Item Response Theory (IRT) models, that is Rasch and Mokken analyses (Bech, 2012; Carrozzino et al., 2021a). In the Rasch analysis, the clinimetric properties are verified testing the overall fit to the Rasch model, the dimensionality of the rating scale, the differential item functioning, and the local independence of items (Bech, 2012; Carrozzino et al., 2020; Rasch, 1980). In the Mokken analysis, scalability is tested via the Loevinger's coefficient of homogeneity (Loevinger, 1948), measuring the extent to which each item assesses a specific degree of severity of the underlying clinical dimension (Bech, 2012; Fava et al., 2018).

Up to now, the clinimetric properties of the DASS-21 have been poorly examined even though the scale has been largely used in clinical and non-clinical realms with the aim of assessing depression, anxiety, stress, three rather common clinical dimensions. Three studies tested unidimensionality using the Rasch analysis in non-clinical (Medvedev et al., 2020; Shea et al., 2009) and clinical (Parkitnya et al., 2012) samples and provided support for the construct validity of the three DASS-21 subscales. However, DASS-21 total scale showed to be multidimensional and had a significant misfit to the Rasch model (Medvedev et al., 2020; Shea et al., 2009; Parkitny et al., 2012). This means that DASS-21 showed relatively poor clinimetric properties even though, being a tool addressed to measure clinical dimensions such as depression, anxiety, and stress, it is recommendable that it entails good clinimetric properties to be useful in the hands of clinicians.

1.1 Study Hypotheses

In order to make available deeper knowledge on the clinimetric properties of a scale so largely used in clinical psychology, the Italian version of the DASS-21 (Bottesi et al., 2015) was analysed using both Rasch and Mokken analyses with the aim of testing dimensionality (of the whole scale and subscales), scalability (i.e., whether the DASS-21 is a statistically sufficient measure of the degree of the underlying dimensions of depression, anxiety, and stress), and sensitivity (i.e., whether the DASS-21 discriminates between individuals displaying different levels of depression, anxiety, or stress) according to CLIPROM criteria (Carrozzino et al., 2021a). We

hypothesize that clinimetric properties of the DASS-21 might be improved via shortened versions, which can be thus implemented in research and clinics.

2. Methods

2.1. Procedure and participants

This is a cross-sectional study involving university students (n = 951; 77.5% women; mean age of 24.86 ± 5.62 years). Data were collected between April and September 2020 via an online survey. The link to the survey was sent to all institutional email addresses of students enrolled at the University of Florence in the academic year 2019-2020 (n = 51,715) together with an invitation to disseminate the link via email or social media among friends and acquaintances. There were no specific requirements for participation, which was voluntary and not compensated. All respondents were provided with information about the study and gave their digital informed consent for study participation. The optimal number of participants to recruit was determined using methodological recommendations (Hagell & Westergren, 2016), which suggest a sample size ranging from 250 to 500 individuals to perform Rasch analyses.

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethical Committee of the University of Florence (#98, protocol n. 0092804).

2.2. Measure

The DASS-21 has three subscales assessing depression, anxiety, and subjective tension or stress over the previous week. Each subscale (Lovibond & Lovibond, 1995) consists of 7 items rated on a 4-point Likert scale with responses ranging from 0 (i.e., "did not apply to me at all") to 3 (i.e., "applied to me very much or most of the time"). For the present research, the Italian version was used. It displayed good measurement properties (Bottesi et al., 2015).

2.3. Statistical analyses

Item Response Theory (IRT) models (i.e., Rasch and Mokken analyses) were conducted according to CLIPROM criteria (Carrozzino et al., 2021a) to evaluate the following clinimetric properties:

-overall fit to the model, which was tested using the chi-square item-trait interaction statistics providing a summary measure of how the DASS-21 conforms to the Rasch model expectations (Pallant & Tennant, 2007; Tennant & Conaghan, 2007). A non-significant chi-square probability value indicates a good level of overall fit to the Rasch model (Pallant & Tennant, 2007; Tennant & Conaghan, 2007).

- Standardized fit residual values for items and individuals were examined for any indication of misfit (Pallant & Tennant, 2007; Tennant & Conaghan, 2007).

- Scalability, which was evaluated using the Mokken analysis to test whether the DASS-21 was a statistically sufficient measure of the degree of the underlying dimensions of depression, anxiety, and stress (Bech, 2012). A Loevinger's coefficient of homogeneity (Loevinger, 1947) ranging from 0.30 to 0.39 indicates a just acceptable level of scalability, while a coefficient ≥ 0.40 is a clear indication of the scalability of the rating scale under evaluation (Bech, 2012).

- Dimensionality, which was examined to determine the construct validity of the DASS-21 testing whether it was a valid clinimetric index of the underlying dimensions of depression, anxiety, and stress. Principal Component Analysis (PCA) of residuals was conducted to identify the two most different subsets of items (i.e., the most positively and negatively factor-loading items on the first component). Paired *t*-tests were then performed to compare scores on the two subsets of items (Christensen et al., 2019; Nielsen et al., 2017). If more than 5% of *t*-tests were significant, the DASS-21 was not considered unidimensional (Christensen et al., 2019; Nielsen et al., 2017).

- Local dependency, which was investigated evaluating whether the response to one item was dependent on the response to another item after controlling for the underlying construct under examination (Marais & Andrich, 2008). A residual correlation value > 0.20 indicates the presence of local dependency between items (Marais & Andrich, 2008).

- Differential Item Functioning (DIF), which was tested to assess whether a certain form of item bias can occur when different groups of individuals (e.g., males and females) respond differently to an item despite equal levels of the dimension under evaluation (Pallant & Tennant, 2007; Tennant & Conaghan, 2007).

- Person Separation Reliability Index (PSI), which was evaluated to estimate the clinimetric sensitivity of the DASS-21, testing its ability to discriminate between individuals with different levels of the dimensions under assessment (Carrozzino et al., 2021a; Pallant & Tennant, 2007).

Consistently with Shea et al. (2009) who identified an additional DASS-21 subscale, we tested the clinimetric properties also of the DASS-21 anxiety/stress subscale.

3. Results

3.1. Fit to the Rasch model

A significant item-trait interaction statistic was found analyzing the DASS-21 total score (χ^2 = 513.54, degrees of freedom [df] = 189, $p < 0.001$), which means misfit to the Rasch model (Table 1, Analysis 1). Standardized fit residuals for items (SD = 3.53) were not within acceptable

limits while standardized fit residuals for persons (SD = 1.46) were within acceptable limits. The overall fit to the Rasch model was not achieved after the exclusion of misfitting items (Table 1, Analysis 2-6) but the model fit was achieved (χ^2 = 107.05, df = 144, p = 0.99) adjusting the sample size (Table 1, Analysis 7).

Table 1. Model fit statistics for DASS-21 total score (n=940)

Sample	Analysis	Model fit (overall)	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant t-tests (%)	Local dependency (residual correlations >0.20)
All items	1	$\chi2$(189)=513.54, p<0.001	-0.40 (3.53)	-0.81 (1.46)	0.93	16.06	Item 3&16, 4&7, 4&15, 4&19, 6&18, 10&13, 10&15, 10&16, 10&21, 11&12, 16&17, 16&21, 17&21
-delete item 2	2	$\chi2$(180)=369.49, p<0.001	-0.42 (2.86)	-0.82 (1.52)	0.93	15.97	Item 3&16, 4&7, 4&15, 4&19, 6&18, 10&13, 10&15, 10&16, 10&21, 11&12, 16&17, 16&21, 17&21
-delete item 13	3	$\chi2$(171)=300.46, p<0.001	-0.38 (2.32)	-0.32 (1.34)	0.92	14.80	Item 3&16, 4&15, 4&19, 6&18, 10&15, 10&16, 10&21, 11&12, 16&17, 16&21, 17&21
-delete item 5	4	$\chi2$(162)=276.10, p<0.001	-0.35 (2.04)	-033 (1.34)	0.92	13.35	Item 3&16, 4&15, 4&19, 6&18, 10&15, 10&16, 10&21, 11&12, 16&17, 16&21, 17&21
-delete item 3	5	$\chi2$(153)=213.60, p<0.001	-0.33 (1.78)	-0.33 (1.30)	0.91	13.03	Item 4&15, 4&19, 6&18, 10&15, 10&16, 10&21, 11&12, 16&17, 16&21, 17&21
-delete item 11	6	$\chi2$(144)=198.67, p=0.002	-0.34 (1.64)	-0.34 (1.30)	0.90	13.47	Item 4&15, 4&19, 6&18, 10&15, 10&16, 10&21, 16&17, 16&21, 17&21
-adjusted sample size (n=500)	7	$\chi2$(144)=107.05, p=0.99	-0.34 (1.64)	-0.34 (1.30)	0.90	13.47	Item 4&15, 4&19, 6&18, 10&15, 10&16, 10&21, 16&17, 16&21, 17&21

$\chi2$: chi-square; p: probability; SD: standard deviation; PSI: person separation index (with extremes)

Model fit statistics of the DASS-21 subscales are reported in Table 2.

As to the depression subscale, a significant item-trait interaction statistic was found (χ^2 = 192.59, df = 63, p < 0.001), indicating misfit to the Rasch model (Table 2, Analysis 1). Standardized fit residuals for items (SD = 4.13) were not within acceptable limits while standardized fit residuals for persons (SD = 1.04) were within acceptable limits. The overall fit to the Rasch model was not achieved excluding the misfitting item (Table 2, Analysis 2). On the contrary, once again,

the model fit was achieved (χ^2 = 46.06, df = 54, p = 0.77) adjusting the sample size (Table 2, Analysis 3).

As to the anxiety subscale, a significant item-trait interaction statistic was found (χ^2 = 261.88, df = 63, p < 0.001), indicating misfit to the Rasch model (Table 2, Analysis 4). Standardized fit residuals for items (SD = 3.85) were not within acceptable limits while standardized fit residuals for persons (SD = 1.10) were within acceptable limits. The overall fit to the Rasch model was not achieved excluding the misfitting items (Table 2, Analysis 5-6) but adjusting the sample size (χ^2 = 41.15, df = 30, p = 0.08) (Table 2, Analysis 7).

As to the stress subscale, a significant item-trait interaction statistic was observed (χ^2 = 97.19, df = 63, p = 0.004), indicating misfit to the Rasch model (Table 2, Analysis 8). Standardized fit residuals for items (SD = 2.41) and persons (SD = 1.21) were within acceptable limits. Model fit was achieved (χ^2 = 60.03, df = 54, p = 0.27) excluding the misfitting item (Table 2, Analysis 9).

Table 2. Model fit statistics for original and revised DASS-21 scales

Action	Analysis	Overall model fit	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant *t*-tests (%)	Local dependency (residual correlations >0.20)
DASS-21-Depression (n=885)	1	χ2(63)=192.59, p<0.001	-0.63 (4.13)	-0.34 (1.04)	0.86	3.28	None
-delete item 5	2	χ2(54)=75.73, p=0.03	-0.38 (1.40)	-0.36 (1.01)	0.85	2.55	None
-adjusted sample size (n=500)	3	χ2(54)=46.06, p=0.77	-0.38 (1.40)	-0.36 (1.01)	0.85	2.55	None
DASS-21-Anxiety (n=818)	4	χ2(63)=261.88, p<0.001	-0.58 (3.85)	-0.34 (1.10)	0.71	2.08	None
-delete item 2	5	χ2(24)=69.43, p<0.001	-0.39 (1.96)	-0.33 (1.07)	0.70	2.11	None
-delete item 15	6	χ2(30)=61.97, p<0.001	-0.08 (0.84)	-0.31 (1.02)	0.64	1.20	None
-adjusted sample size (n=500)	7	χ2(30)=41.15, p=0.08	-0.08 (0.84)	-0.31 (1.02)	0.64	1.20	None
DASS-21-Stress (n=919)	8	χ2(63)=97.19, p=0.004	-0.05 (2.41)	-0.39 (1.21)	0.87	5.77	None
-delete item 11	9	χ2(54)=60.03, p=0.27	0.17 (1.50)	-0.42 (1.29)	0.84	4.73	None

χ2: chi-square; p: probability; SD: standard deviation; PSI: person separation index (with extremes)

As to the anxiety-stress dimension (Table 3), the analysis showed a significant item-trait interaction statistic (χ^2 = 332.84, df = 126, p < 0.001), indicating misfit to the Rasch model (Table 3, Analysis 1). Standardized fit residuals for items (SD = 3.34) were not within acceptable limits while standardized fit residuals for persons (SD = 1.24) were within acceptable limits. The overall fit to the Rasch model was not achieved excluding misfitting items (Table 3, Analysis 2-5) but model fit was achieved (χ^2 = 61.77, df = 90, p = 0.99) adjusting the sample size (Table 3, Analysis 7).

Table 3. Model fit statistics for DASS-21-Anxiety-Stress scale

Action	Analysis	Overall model fit	Item fit residual, mean (SD)	Person fit residual, mean (SD)	PSI	Dimensionality, significant *t*-tests (%)	Local dependency (residual correlations >0.20)
DASS-21-Anxiety-Stress (n=936)	1	χ2(126)=332.84, p<0.001	-0.18 (3.34)	-0.32 (1.24)	0.90	11.22	Item 4&15, 11&12
-delete item 2	2	χ2(117)=185.21, p<0.001	-0.22 (2.2.)	-0.32 (1.18)	0.90	11.24	Item 4&15, 11&12
-delete item 11	3	χ2(108)=156.50, p=0.001	-0.18 (1.90)	-0.33 (1.17)	0.88	9.72	Item 4&15
-delete item 15	4	χ2(99)=125.50, p=0.04	-0.09 (1.66)	-0.32 (1.17)	0.87	11.34	None
-delete item 12	5	χ2(90)=114.03, p=0.04	-0.39 (1.96)	-0.33 (1.07)	0.86	8.99	None
-adjusted sample size (n=500)	7	χ2(90)=61.77, p=0.99	-0.39 (1.96)	-0.33 (1.07)	0.86	8.99	None

χ2: chi-square; p: probability; SD: standard deviation; PSI: person separation index (with extremes)

3.2. Scalability and dimensionality

Based on Mokken analysis, DASS-21 total score had an acceptable scalability (Loevinger's coefficient of homogeneity = 0.50). Significant *t*-tests outside the critical value of 5% suggested multidimensionality for the DASS-21 (Table 1, Analysis 1-7).

The depression subscale showed acceptable scalability (Loevinger's coefficient of homogeneity = 0.67) and unidimensionality (< 5% of *t*-tests were significant) (Table 2, Analysis 1). Dimensionality (Table 2, Analysis 2) and scalability (Loevinger's coefficient of homogeneity of 0.72) improved excluding the misfitting item.

The anxiety subscale had acceptable scalability (Loevinger's coefficient of homogeneity = 0.50) and unidimensionality (< 5% of *t*-tests were significant) (Table 2, Analysis 4). Dimensionality improved excluding misfitting items (Table 2, Analysis 5-6).

The stress subscale showed acceptable scalability (Loevinger's coefficient of homogeneity = 0.59) but multidimensionality (significant *t*-tests were outside the critical value of 5%) (Table 2, Analysis 8). The DASS-21 stress subscale became unidimensional excluding the misfitting item (< 5% of *t*-tests were significant) (Table 2, Analysis 9).

The DASS-21 anxiety-stress dimension showed to be multidimensional (>5% of *t*-tests were significant) (Table 3, Analysis 1-7).

3.3. Local dependency, DIF, PSI

Residual correlations more than 0.20, indicating local dependency, were found between the following DASS-21 item-pairs: items 3 (i.e., " I couldn't seem to experience any positive feeling at all") and 16 (i.e., "I was unable to become enthusiastic about anything"); items 4 (i.e., "I experienced breathing difficulty" [e.g., excessively rapid breathing, breathlessness in the absence of physical exertion]) and 7 (i.e., "I experienced trembling" [e.g., in the hands]); items 4 and 15 (i.e., "I felt I was close to panic"); items 4 and 19 (i.e., "I was aware of the action of my heart in the absence of physical exertion"); items 6 (i.e., "I tended to over-react to situations") and 18 (i.e., "I felt that I was rather touchy"); items 10 (i.e., "I felt that I had nothing to look forward to") and 13 (i.e., "I felt down-hearted and blue"); items 10 and 15 (i.e., "I felt I was close to panic"); items 10 and 16 (i.e., "I was unable to become enthusiastic about anything"); items 10 and 21 (i.e., "I felt that life was meaningless"); items 11 (i.e., "I found myself getting agitated") and 12 (i.e., "I found it difficult to relax"); items 16 (i.e., "I was unable to become enthusiastic about anything") and 17 (i.e., "I felt I wasn't worth much as a person"); items 16 and 21 (i.e., "I felt that life was meaningless"); and items 17 (i.e., "I felt I wasn't worth much as a person") and 21.

No indication of local dependency between items of the DASS-21 subscales was found.

As to the DASS-21 anxiety-stress dimension, local dependency was detected between items 4 and 15, and between items 11 and 12.

Items 11, 12, 16, and 21 (i.e., "I felt that life was meaningless") of the DASS-21 showed uniform DIF for sex. There were no other indications of uniform or non-uniform DIF.

PSI was 0.93, indicating that the DASS-21 total score could reliably discriminate between individuals with different levels of the underlying trait under examination (Table 1, Analysis 1). As to the depression subscale, the PSI was 0.86, suggesting that this measure could reliably

discriminate between respondents with different levels of the underlying construct under evaluation (Table 2, Analysis 1). The DASS-21 anxiety subscale had a PSI of 0.71, indicating that this measure could reliably distinguish between different groups but not between different subjects (Table 2, Analysis 4). As to the stress factor, the PSI was 0.87, suggesting that the DASS-21 stress subscale could reliably discriminate between individuals with different levels of the underlying dimension under evaluation (Table 2, Analysis 8). The DASS-21 anxiety-stress dimension had a PSI of 0.90, indicating that this subscale could reliably discriminate between subjects with different levels of the underlying trait under examination (Table 3, Analysis 1).

4. Discussion

In their DASS-21 validation analysis, which was mainly conducted using classical psychometrics (e.g., confirmatory factor analysis), Henry and Crawford (2005) concluded that this is a valid measure of an underlying factor of general psychological distress. The findings of our clinimetric analysis run counter to this conclusion since the DASS-21 was found to be a multidimensional index that cannot be used to assess general psychological distress. This is consistent with previous Rasch analyses (Medvedev et al., 2020; Parkitny et al., 2012; Shea et al., 2009). Caution should be therefore paid when using the DASS-21 as a measure of psychological distress, at least in university students.

Fit to the Rasch measurement model was achieved after excluding five misfitting items (i.e., items 2, 3, 5, 11, 13) and adjusting the sample size. The result is a 16-item version of the DASS-21 which assesses depressive (e.g., "I felt that I had nothing to look forward to") and anxious symptoms (e.g., "I felt scared without any good reason"), interpersonal sensitivity (e.g., "I felt that I was rather touchy"), and neuroticism (e.g., "I tended to over-react to situations"). From a conceptual perspective, the total score of the 16-item DASS-21 may reflect the clinical construct of dysthymia as outlined by Farmer et al. (2002) and by Bech et al. (Bech et al., 2016a; Bech et al., 2016b). This is a dimension of vulnerability characterized by a tendency to feel anxious, worried, and pessimistic about the future (Bech et al. 2016a; Bech et al., 2016b; Farmer et al. 2002). The 16-item version of the DASS-21 may also reflect the construct of dysthymia recently refined by Guidi and Fava (2022). That is a clinical condition at the other end of a continuum with euthymia and characterized by demoralization, chronic worrying, mental pain, subjective incompetence, rigidity, and abnormal reactivity to environmental stimuli (Guidi & Fava, 2022). The 16-item DASS-21 might thus be a useful complementary tool in the clinical realm of diagnostic instrument for dysthymia (e.g., the Clinical Interview for Dysthymia – Guidi & Fava, 2022) since it allows to assess specific aspects of this clinical manifestation as soon as it is properly diagnosed. Our results indeed indicate that the total score of the 16-item version

of the DASS-21 entailed the clinimetric property of sensitivity. However, PCA of residuals indicated multidimensionality and this index also included inter-correlated items, which provide clinically redundant information. This implies that the total score of the 16-item version of the DASS-21 should be supplemented with the use of the DASS-21 subscales of depression, anxiety, and stress, which did not include locally dependent items, fitted the Rasch model expectations, were unidimensional, and had an acceptable scalability. In brief, clinicians should use both 16-item version of the DASS-21 and DASS-21 subscales of depression, anxiety, stress if they want to have a more accurate assessment.

Regarding DASS-21 subscales, after removing the misfitting item and adjusting the sample size, the scalability of the resulting 6-item version of the subscale of depression improved. This short form of the depression subscale was found to fit the Rasch model and to be unidimensional. It can therefore be used by clinicians as a dimensional measure of depression, preferably together with other clinimetric indices, such as the Major Depression Inventory (Carrozzino et al., 2021b; Olsen et al., 2003) and the Hamilton Rating Scales for Depression (Carrozzino et al., 2020; Timmerby et al., 2017), to take advantage of incremental validity (Hunsley & Meyer, 2003; Sechrest, 1963).

The DASS-21 anxiety subscale showed an unsatisfactory initial fit to the Rasch model and the exclusion of misfitting items did not improve the overall fit (Medvedev et al., 2020). Model fit for the 5-item version was achieved only after adjusting the sample size. This subscale was found to entail the clinimetric property of dimensionality or construct validity, implying that it can be used by clinicians and researchers as a dimensional measure when they need to assess anxiety severity. It should be remarked that most of the items of this subscale (e.g., "I experienced breathing difficulty", "I experienced trembling [e.g., in the hands]", "I was worried about situations in which I might panic and make a fool of myself", "I was aware of the action of my heart in the absence of physical exertion") cover somatic symptoms of panic rather than the wide spectrum of anxiety. This makes this brief index suitable for clinicians to easily and briefly assess somatic symptoms of panic while, based on the clinical content of items, it should be used in combination with other clinimetric instruments, such as the 8-item version of the Anxiety Symptom Scale (ASS_8; Bech, 2012; Bech et al., 2014) and the 6-item version of the Hamilton Rating Scale for Anxiety (HAM-A_6; Bech, 2007; Loldrup et al., 1989), to provide a comprehensive assessment of anxiety.

As to the DASS-21 subscale of stress, fit to the Rasch model was achieved for the 6-item version, which had acceptable unidimensionality. This subscale can be therefore used in clinical psychology as a dimensional measure of psychological distress, particularly of negative

affectivity, as originally introduced by Watson and Clark (1984). However, given the limited number of items, this subscale should be administered in combination with comprehensive measures of psychological distress such as the revised version of the Hopkins Symptom Checklist (SCL-90-R; Carrozzino et al., 2022), the Psychosocial Index (Sonino & Fava, 1998), and the Kellner's Symptom Questionnaire (SQ; Benasi et al., 2020).

As to the DASS-21 anxiety/stress subscale, evidence of multidimensionality was found and the fit to the Rasch model was achieved only after removing misfitting items (i.e., items 2, 11, 12, and 15) and adjusting the sample size. The resulting 10-item version seems to be informative only as an overall indicator of stress related to anxious symptoms. Future studies are needed to improve its construct validity.

The present study has some limitations. The results are based on a convenience sampling that limits the generalizability of the findings. Future studies using better-targeted (i.e., clinical) populations with higher levels of depression, anxiety, and psychological distress (e.g., dysthymic, demoralized, depressed patients) are needed to provide support to our findings. Second, a cross-sectional design was used, thus precluding the assessment of predictive and incremental validity of the DASS-21. Future research using a longitudinal design is encouraged. A comparison with other assessment tools was not performed, thus precluding the evaluation of concurrent validity. Future studies comparing the DASS-21 with related and previously validated measures of depression, anxiety, and psychological distress (including both clinician-rated scales and self-reported questionnaires) are recommended.

5. Conclusion

This is the first study applying CLIPROM criteria (Carrozzino et al., 2021a) to the validation process of the DASS-21. Our findings indicate that there is no valid justification for using in clinics and research the total score of the DASS-21 as a global measure of general psychological distress because of misfit to the Rasch model, multidimensionality, and local dependency of items. Fit to the Rasch measurement expectations was obtained after removal of misfitting symptoms and adjustment of sample size, which resulted in the 16-item version of the DASS-21 that has a valid justification of being used. However, its total score should be supplemented with the DASS-21 subscales of depression, anxiety, and stress. The brief versions of these subscales, which did not include misfitting and clinically redundant items, can be used by clinicians and researchers as unidimensional measures together with other clinimetric indices in order to have a detailed description of the clinical phenomenon and take advantage of incremental validity.

Depression and anxiety are highly comorbid conditions (Cosci & Fava, 2021) that involve the experience of psychological distress and require an adequate assessment based on clinimetric criteria, rather than on classical psychometric principles, that do not fit in with the complexity of clinical reality (Bech, 2012; Carrozzino et al., 2021a; Fava, 2022). Progress can be achieved in assessment research and practice creating clinimetric versions of existing psychometric tools (as it is the case for DASS-21), and promoting an accurate, entailed, although comprehensive, assessment of subjects and patients.

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethical Committee of the University of Florence (#98, protocol n. 0092804).

Informed Consent Statement

An online consent form was used to obtain consent from participants, and all participants consented that their data should be used for this research. Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The present research's data will be available on request.

Conflict of interest statement

The authors declare no conflict of interest.

Authors' Contribution

G.M: investigation, writing – original draft; D.C.: writing – original draft, writing - review & editing; K.S.C.: formal analysis, writing - review & editing; C.P.: writing – original draft, writing - review & editing; F.C.: conceptualization, supervision, writing - review & editing.

References

1. Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10*(2), 176-181. https://doi.org/10.1037/1040-3590.10.2.176
2. Bech, P. (2004). Modern psychometrics in clinimetrics. *Psychotherapy and Psychosomatics. 73*(3), 134-138. https://doi.org/10.1159/000076448
3. Bech, P. (2007). Dose-response relationship of pregabalin in patients with generalized anxiety disorder: A pooled analysis of four placebo-controlled trials. *Pharmacopsychiatry*, *40*(4), 163-168. https://doi.org/10.1055/s-2007-984400
4. Bech, P. (2012). *Clinical Psychometrics*. Wiley-Blackwell.
5. Bech, P., Bille, J., Møller, S. B., Hellström, L. C., & Østergaard, S. D. (2014). Psychometric validation of the Hopkins Symptom Checklist (SCL-90) subscales for depression, anxiety, and interpersonal sensitivity. *Journal of Affective Disorders, 160*, 98-103. https://doi.org/10.1016/j.jad.2013.12.005
6. Bech, P., Carrozzino, D., Austin, S. F., Møller, S. B., & Vassend, O. (2016a). Measuring euthymia within the Neuroticism Scale from the NEO Personality Inventory: a Mokken analysis of the Norwegian general population study for scalability. *Journal of Affective Disorders, 193*, 99-102. https://doi.org/10.1016/j.jad.2015.12.039
7. Bech, P., Kessing, L. V., & Bukh, J. D. (2016b). The validity of dysthymia to predict clinical depressive symptoms as measured by the Hamilton Depression Scale at the 5-year follow-up of patients with first episode depression. *Nordic Journal of Psychiatry, 70*(8), 563-566. https://doi.org/10.1080/08039488.2016.1180712
8. Benasi, G., Fava, G. A., & Rafanelli, C. (2020). Kellner's symptom questionnaire, a highly sensitive patient-reported outcome measure: systematic review of clinimetric properties. *Psychotherapy and Psychosomatics*, *89*(2), 74-89. https://doi.org/10.1159/000506110
9. Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry, 60*, 170-181. https://doi.org/10.1016/j.comppsych.2015.04.005
10. Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D.H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35*(1), 79-89. https://doi.org/10.1016/s0005-7967(96)00068-x
11. Carrozzino, D., Patierno, C., Guidi, J., Berrocal Montiel, C., Cao, J., Charlson, M. E., Christensen, K. S., Concato, J., De las Cuevas, C., de Leon, J., Eöry, A., Fleck, M. P., Furukawa, T. A., Horwitz, R. I., Nierenberg, A. A., Rafanelli, C., Wang, H., Sonino, N., & Fava, G. A. (2021a). Clinimetric criteria for patient-reported outcome measures. *Psychotherapy and Psychosomatics, 90*(4), 222-232. https://doi.org/10.1159/000516599
12. Carrozzino, D., Christensen, K. S., & Cosci, F. (2021b). Construct and criterion validity of patient-reported outcomes (PROs) for depression: a clinimetric comparison. *Journal of Affective Disorders, 283*, 30-35. https://doi.org/10.1016/j.jad.2021.01.043

13. Carrozzino, D., Christensen, K. S., Patierno, C., Siri, C., Zecchinelli, A., Pezzoli, G., & Cosci, F. (2022). The Hopkins Symptom Checklist (SCL-90-R): A Patient-Reported Outcome Measure in Parkinson's Disease. *Journal of Geriatric Psychiatry and Neurology, 35*(5), 689-697. https://doi.org/10.1177/08919887211060020
14. Carrozzino, D., Patierno, C., Fava, G. A., & Guidi, J. (2020). The Hamilton rating scales for depression: a critical review of clinimetric properties of different versions. *Psychotherapy and Psychosomatics, 89*(3), 133-150. https://doi.org/10.1159/000506879
15. Charlson, M. E., Carrozzino, D., Guidi, J., & Patierno, C. (2022). Charlson Comorbidity Index: A Critical Review of Clinimetric Properties. *Psychotherapy and Psychosomatics, 91*(1), 8-35. https://doi.org/10.1159/000521288
16. Christensen, K. S., Oernboel, E., Nielsen, M. G., & Bech, P. (2019). Diagnosing depression in primary care: A Rasch analysis of the major depression inventory. *Scandinavian Journal of Primary Health Care, 37*(1), 105-112. https://doi.org/10.1080/02813432.2019.1568703
17. Clara, I., Cox, B., & Enns, M. (2001). Confirmatory factor analysis of the Depression-Anxiety-Stress Scales in depressed and anxious patients. *Journal of Psychopathology and Behavioral Assessment, 23*(1), 61-67. https://doi.org/10.1023/A:1011095624717
18. Crawford, J. R., & Henry, J. D. (2003). The Depression Anxiety Stress Scales (DASS): normative data and latent structure in a large non-clinical sample. *British Journal of Clinical Psychology, 42*(2), 111-131. https://doi.org/10.1348/014466503321903544
19. Cosci, F. (2021). Clinimetric perspectives in clinical psychology and psychiatry. *Psychotherapy and Psychosomatics, 90*(4), 217-221. https://doi.org/10.1159/000517028
20. Cosci, F., & Fava, G. A. (2021). When anxiety and depression coexist: the role of differential diagnosis using clinimetric criteria. *Psychotherapy and Psychosomatics*, *90*(5), 308-317. https://doi.org/10.1159/000517518
21. Farmer, A., McGuffin, P., & Williams, J. (2002). *Measuring Psychopathology*. Oxford University Press.
22. Fava, G. A. (2022). Forty years of clinimetrics. *Psychotherapy and Psychosomatics*, *91*(1), 1-7. https://doi.org/10.1159/000520251
23. Fava G. A., Carrozzino D., Lindberg L., & Tomba, E. (2018). The clinimetric approach to psychological assessment: A tribute to Per Bech, MD (1942- 2018). *Psychotherapy and Psychosomatics*, *87*(6), 321–326. https://doi.org/10.1159/000493746
24. Fava, G. A., Ruini, C., & Rafanelli, C. (2004). Psychometric theory is an obstacle to the progress of clinical research. *Psychotherapy and Psychosomatics*, *73*(3), 145-148. https://doi.org/10.1159/000076451
25. Fava, G. A., Tomba, E., & Sonino, N. (2012). Clinimetrics: the science of clinical measurements. *International Journal of Clinical Practice, 66*(1), 11-15. https://doi.org/10.1111/j.1742-1241.2011.02825.x
26. Feinstein, A. R. (1982). T. Duckett Jones memorial lecture. The Jones criteria and the challenges of clinimetrics. *Circulation, 66*(1), 1-5. https://doi.org/10.1161/01.CIR.66.1.1
27. Feinstein, A.R. (1983). An additional basic science for clinical medicine: IV. The development of clinimetrics. *Annals of Internals Medicine, 99*(6), 843-848. https://doi.org/10.7326/0003-4819-99-6-843
28. Feinstein, A. R. (1987). *Clinimetrics*. Yale University Press.

29. Guidi, J., & Fava, G. A. (2022). The Clinical Science of Euthymia: A Conceptual Map. *Psychotherapy and Psychosomatics, 91*(3), 156-167. https://doi.org/10.1159/000524279
30. Hagell, P., & Westergren, A. (2016). Sample Size and Statistical Conclusions from Tests of Fit to the Rasch Model According to the Rasch Unidimensional Measurement Model (RUMM) Program in Health Outcome Measurement. *Journal of Applied Measurement, 17*(4), 416-431.
31. Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 44*(2), 227-239. https://doi.org/10.1348/014466505X29657
32. Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychological Assessment, 15*(4), 446-455. https://doi.org/10.1037/1040-3590.15.4.446
33. Lee, J., Lee, E. H., & Moon, S. H. (2019). Systematic review of the measurement properties of the Depression Anxiety Stress Scales–21 by applying updated COSMIN methodology. *Quality of Life Research, 28*(9), 2325-2339. https://doi.org/10.1007/s11136-019-02177-x
34. Loevinger, J. (1947). A systematic approach to the construction and evaluation of tests of ability. *Psychological Monographs*, *61*(4), i–49. https://doi.org/10.1037/h0093565
35. Loevinger, J. (1948). The technic of homogeneous tests compared with some aspects of scale analysis and factor analysis. *Psychological Bulletin, 45*(6), 507–529. https://doi.org/10.1037/h0055827
36. Loldrup, D., Langemark, M., Hansen, H. J., Olesen, J., & Bech, P. (1989). Clomipramine and mianserin in chronic idiopathic pain syndrome. *Psychopharmacology*, *99*(1), 1-7. https://doi.org/10.1007/BF00634443
37. Lovibond, P. F., & Lovibond, S. H. (1995a). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335-343. https://doi.org/10.1016/0005-7967(94)00075-U
38. Lovibond, S. H., & Lovibond, P. F. (1995b). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Psychology Foundation Monograph.
39. Lovibond, S. H., & Lovibond, P. F. (1993). *Manual for the Depression Anxiety Stress Scales (DASS)* (1st ed.). Psychology Foundation Monograph.
40. Mahmoud, J., Hall, L. A., & Staten, R. S. (2010). The psychometric properties of the 21-Item Depression Anxiety and Stress Scale (DASS-21) among a sample of young adults. *Southern Online Journal of Nursing Research, 10*(4), 21-34.
41. Marais, I., & Andrich, D. (2008). Effects of Varying Magnitude and Patterns of Response Dependence. *Journal of Applied Measurement, 9*(2), 105-124.
42. Mboua, P. C., Siakam, C., & Mabo, N. L. (2021). Impact of the resumption of classes on the mental health of students of the Faculty of Letters and Social Sciences of the University of Dschang, in the context of Covid 19. *Journal of Affective Disorders Reports, 5*:100147. https://doi.org/10.1016/j.jadr.2021.100147
43. Medvedev, O. N., Krägeloh, C. U., Titkova, E. A., & Siegert, R. J. (2020). Rasch analysis and ordinal-to-interval conversion tables for the Depression, Anxiety and Stress Scale. *Journal of Health Psychology, 25*(10-11), 1374-1383. https://doi.org/10.1177/1359105318755261

44. Moya, E., Larson, L. M., Stewart, R. C., Fisher, J., Mwangi, M. N., & Phiri, K. S. (2022). Reliability and validity of depression anxiety stress scale (DASS)-21 in screening for common mental disorders among postpartum women in Malawi. *BMC Psychiatry, 22*(1), 352. https://doi.org/10.1186/s12888-022-03994-0
45. Nielsen, M. G., Ørnbøl, E., Vestergaard, M., Bech, P., & Christensen, K. S. (2017). The construct validity of the major depression inventory: a Rasch analysis of a self-rating scale in primary care. *Journal of Psychosomatic Research, 97*, 70-81. https://doi.org/10.1016/j.jpsychores.2017.04.001
46. Olsen, L. R., Jensen, D. V., Noerholm, V., Martiny, K., & Bech, P. (2003). The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychological Medicine, 33*(2), 351-356. https://doi.org/10.1017/s0033291702006724
47. Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology, 46*(3), 283-297. https://doi.org/10.1348/014466506X158996
48. Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology, 46*(1), 1-18. https://doi.org/10.1348/014466506X96931
49. Parkitny, L., McAuley, J. H., Walton, D., Costa, L. O. P., Refshauge, K .M., Wand, B. M., Di Pietro, F., & Moseley, G. L. (2012). Rasch analysis supports the use of the depression, anxiety, and stress scales to measure mood in groups but not in individuals with chronic low back pain. *Journal of Clinical Epidemiology, 65*(2), 189-198. https://doi.org/10.1016/j.jclinepi.2011.05.010
50. Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests.* The University of Chicago Press.
51. Sakakibara, B. M., Miller, W. C., Orenczuk, S. G., Wolfe, D. L.; & SCIRE Research Team (2009). A systematic review of depression and anxiety measures used with individuals with spinal cord injury. *Spinal Cord, 47*(12), 841-51. https://doi.org/10.1038/sc.2009.93
52. Sechrest, L. (1963). Incremental validity: A recommendation. *Educational and Psychological Measurement, 23*(1), 153-158. https://doi.org/10.1177/001316446302300113
53. Sekhar, P., Tee, Q. X., Ashraf, G., Trinh, D., Shachar, J., Jiang, A., Hewitt, J., Green, S., & Turner, T. (2021). Mindfulness-based psychological interventions for improving mental well-being in medical students and junior doctors. *Cochrane Database Systematic Reviews, 12*(12), CD013740. https://doi.org/10.1002/14651858.CD013740.pub2
54. Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry. 9*(1), 1-10. https://doi.org/10.1186/1471-244X-9-21
55. Smith, E. L., Garety, P. A., Harding, H., & Hardy, A. (2021). Are there reliable and valid measures of anxiety for people with psychosis? A systematic review of psychometric properties. *Psychol Psychother. 94*(1), 173-198. https://doi.org/10.1111/papt.12265
56. Sonino, N., & Fava, G. A. (1998). A simple instrument for assessing stress in clinical practice. *Postgradraduate Medical Journal, 74*(873), 408-410. https://doi.org/10.1136/pgmj.74.873.408

57. Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research, 57*(8), 1358-1362. https://doi.org/10.1002/art.23108

58. Timmerby, N., Andersen, J. H., Søndergaard, S., Østergaard, S. D., & Bech, P. (2017). A systematic review of the clinimetric properties of the 6-item version of the Hamilton Depression Rating Scale (HAM-D6). *Psychotherapy and Psychosomatics, 86*(3), 141-149. https://doi.org/10.1159/000457131

59. Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96*(3), 465–490. https://doi.org/10.1037/0033-2909.96.3.465

60. Wright, J. G., & Feinstein, A. R. (1992). A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *Journal of Clinical Epidemiology, 45*(11), 1201-1218. https://doi.org/10.1016/0895-4356(92)90161-f

61. Yeung, A. Y., Yuliawati, L., & Cheung, S. H. (2020). A systematic review and meta-analytic factor analysis of the Depression Anxiety Stress Scales. *Clinical Psychology: Science and Practice, 27*(4), e12362. https://doi.org/10.1111/cpsp.12362

DOI: 10.13129/2282-1619/mjcp-3795