



Systematic Review of the Psychometric Evidence Supporting Use of the Depression Anxiety Stress Scales, Short Form (DASS-21) with Youth

Aberdine R. Dwight¹ · Amy M. Briesch¹ · Jessica A. Hoffman¹ · Christopher Rutt²

Accepted: 7 February 2024
© The Author(s) 2024

Abstract

Background Although the Depression Anxiety Stress Scales, Short Form (DASS-21) was developed for adults, its authors noted no compelling reasons to not use the measure with youth as young as 12 years. Despite increasingly widespread use with youth, psychometric evidence in support of its use with this population needs to be investigated to fully understand its utility.

Objective The goal of this paper was to conduct a systematic review to identify and appraise evidence for the psychometric properties of the Depression Anxiety Stress Scales, Short Form (DASS-21) when used with youth populations.

Methods Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, this systematic review identified 18 studies that reported relevant information on the psychometrics of the DASS-21 with youth samples. Each article was coded for evidence of three indicators of particular importance to progress monitoring: reliability, validity, and sensitivity to change.

Results Studies have found varied factor structures (i.e., 1–4 factors) for the DASS-21 and validity evidence based on relations to other variables has been limited. As such, further research is needed to understand the utility of the depression, anxiety, and stress subscales. Furthermore, no information is currently available about the sensitivity to change of the measure.

Conclusions Most studies reporting psychometric properties in this review utilized the DASS-21 with mixed-gender groups of secondary school students, thus supporting its use with similar populations. Pending additional inquiry, however, both researchers and practitioners are encouraged to utilize the total score, for which stronger evidence of reliability and validity exist.

Keywords DASS-21 · Psychometric · Reliability · Validity · Factor analysis

✉ Amy M. Briesch
a.briesch@northeastern.edu

¹ Department of Applied Psychology, Northeastern University, 360 Huntington Avenue, Boston 02115, MA, USA

² Boston Child Study Center, Boston, MA, USA

Introduction

Internalizing behaviors including depression, anxiety, social withdrawal, and somatic problems impact the behavioral, social, and emotional functioning of up to 20% of youth in the United States (Walker et al., 2000). Internalizing symptoms, such as worry, low self-esteem, and low mood usually occur when individuals try to control their emotions or impulses and tend to influence individuals' emotions in a maladaptive manner (Merrill & Gueldner, 2010). Given the range of negative outcomes associated with internalizing distress, including functional impairment (Kendall et al., 2010), social skills deficits (Nelson et al., 2004), and substance abuse (Lopez et al., 2005), there is a critical need to support youth with internalizing symptomatology. One means to support youth is providing evidence-based care, which includes monitoring responsiveness to treatment.

Progress monitoring, or continual routine tracking, is an important aspect of evidence-based practice as it facilitates more timely decision making as opposed to traditional pre-post assessment (APA, 2006). Reliance on infrequent monitoring can lead to waiting extended periods of time to determine intervention effectiveness and may result in youth receiving ineffective services and clinicians being uninformed about more frequent progress updates. However, to effectively progress monitor response to interventions that target internalizing difficulties, it is essential that reliable tools be validated for this specific assessment purpose. Specifically, for progress monitoring data to defensibly inform decision making, it is important for tools to demonstrate evidence of reliability, validity, and treatment sensitivity (i.e., the ability to detect changes in student behavior over time; Briesch & Volpe, 2007). Many self-report rating scales exist for assessing internalizing concerns in youth, such as the Children's Depression Inventory, Second Edition (CDI-2; Kovacs, 2011), the Multidimensional Anxiety Scale for Children, Second Edition (MASC-2; March, 2013), the Revised Children's Manifest Anxiety Scale, Second Edition (RCMAS-2; Reynolds & Richmond, 2008), and the Reynolds Adolescent Depression Scale, Second Edition (Reynolds, 2004). Although these traditional, norm-referenced behavior rating scales have impressive psychometric properties, these measures have been designed for diagnostic or summative purposes, comparing student reports to normative benchmarks. Additionally, youth are typically asked to reflect the extent to which they generally experience particular thoughts, feelings, and/or behaviors as opposed to focusing assessment on a shorter period of time (e.g., daily, weekly). As such, the utility of these tools within a progress monitoring context is unknown.

A systematic review was recently conducted by Dart et al. (2019) to better understand the landscape of available progress monitoring measures for internalizing symptoms that can be utilized with youth. In searching for tools appropriate for progress monitoring, the authors used the inclusion criterion that "an assessment is able to detect small changes in functioning and is designed for frequent (i.e., at least weekly) administration (i.e., resistant to practice effects, practical, and feasible) to provide information on treatment progress and inform treatment decisions" (Dart et al., 2019, p. 267). Although the review identified a total of 8 self-report measures, psychometric evidence was only available for three of these measures. Whereas the Separation Avoidance Anxiety Inventory (SAAI; In-Albon et al., 2013) focuses exclusively on the assessment of separation anxiety across 12 possible situations, the Brief Problem Checklist (BPC; Chorpita et al., 2010) provides a broad assessment of both internalizing and externalizing concerns across 15 items. In contrast, the Outcome Questionnaire (OQ; Vermeersch et al., 2000) measures three domains of adolescent functioning and potential impairment including subjective discomfort, interpersonal

relations, and social role performance. Consideration of the limited scope of these measures in combination with the available psychometric evidence (i.e., the SAAI being the only measure with evidence and reliability, validity, and sensitivity to change), suggests a continuing need to identify self-report measures designed for the regular progress monitoring of youth with internalizing concerns.

Depression Anxiety Stress Scales

One tool that has been used to progress monitor internalizing concerns in adults is the Depression Anxiety Stress Scales-Short Form (DASS-21; Lovibond & Lovibond, 1995). This 21-item self-report rating scale is an abbreviated version of the full 42-item DASS, consisting of three subscales. The depression subscale measures dysphoria, hopelessness, low self-esteem, self-depreciation, lack of interest, anhedonia, and low positive affect. The anxiety subscale assesses autonomic arousal, musculo-skeletal symptoms, situational anxiety, and subjective experiences of anxious arousal. Finally, the stress subscale measures difficulty relaxing, tension, agitation, irritability, and negative affect (Lovibond & Lovibond, 1995). Each of the three subscales consists of 7 items, which ask respondents to indicate how much a statement applied to them over the past week using a 4-point scale (i.e., 0 = *Did not apply to me at all*; 3 = *Applied to me very much, or most of the time*). Higher scores indicate more frequent symptomatology. Items on the DASS-21 were selected from the full-length DASS based on several criteria, including strong factor loadings and item means (Lovibond & Lovibond, 1995). The DASS-21 is available for free from the Psychology Foundation of Australia (n.d.).

Initially validated in a nonclinical sample of participants aged 17–69 years old, numerous psychometric studies support use of the DASS-21 with adult populations (e.g., Clara et al., 2001; Ng et al., 2007). Additionally, the English-language measure has been successfully translated into, and validated across, multiple languages including Chinese (Cao et al., 2023), Hindi (e.g., Kumar et al., 2019), and Spanish (e.g., Bados et al., 2005). Although the lower age limit of the development samples was 17 years, Lovibond and Lovibond (1995) noted no compelling reasons not to use the DASS-21 with children as young as 12. As such, the DASS-21 may be able to fill the gap that exists with regard to psychometrically-defensible progress monitoring tools for internalizing concerns. Because the scale was developed for, and tested primarily with, adult populations, however, one cannot assume that similar psychometric properties exist when used with a younger population. As a result, it is necessary to better understand the extant evidence supporting use of the DASS-21 specifically with youth populations.

Purpose of Study

Despite increasingly widespread application of the DASS-21 with adult and youth populations, psychometric evidence in support of its use specifically with a youth population needs to be investigated to fully understand its utility. Given this gap in the literature, this paper documents a systematic review of the literature pertaining to the psychometric properties of the DASS-21 with adolescent populations. The goal of this systematic review was to present information for both researchers and clinicians while also identifying limitations of the literature and highlighting both ideas for future research and clinical implications.

More specifically, we sought to evaluate the appropriateness of using the DASS-21 for progress monitoring with youth (i.e., under 18 years old) populations by answering the following questions:

1. What evidence exists regarding the reliability (i.e., test–retest, internal consistency) of data obtained from the DASS-21 with children and adolescents?
2. What evidence exists regarding the validity (i.e., structural, convergent, discriminant) of data obtained from the DASS-21 with children and adolescents?
3. What evidence exists regarding the sensitivity to change of data obtained from the DASS-21 with children and adolescents?

Method

Search Procedure

This systematic literature review adhered to the reporting guidelines of systematic reviews (Preferred Reporting Items for Systematic Reviews and Meta-Analyses [PRISMA]; Moher et al., 2009). During March 2023, the APA PsycInfo, eBook Collection (EBSCOhost), ERIC, MEDLINE, Psychology and Behavioral Sciences Collection, SPORTDiscus, and Dissertation and Theses databases were searched for papers citing exploration of the psychometric properties of the DASS-21 with youth populations. No date restrictions were used. All possible combinations of the following search terms were used: (a) either “DASS-21” or “depression, anxiety, stress scale” and (b) “psychometric,” “reliability,” “validity,” “sensitivity to change,” “factor structure,” “factor analysis,” or “reproducibility of results.”

The database search identified 1266 articles (see Fig. 1). After removing duplicates, 468 original titles were screened based on their titles and abstracts to remove irrelevant articles. Specifically, studies were selected for inclusion in the review if two criteria were met. First, the authors published the paper in English and in a peer-reviewed scientific journal or as a dissertation. Second, the authors examined any psychometric properties of the DASS-21

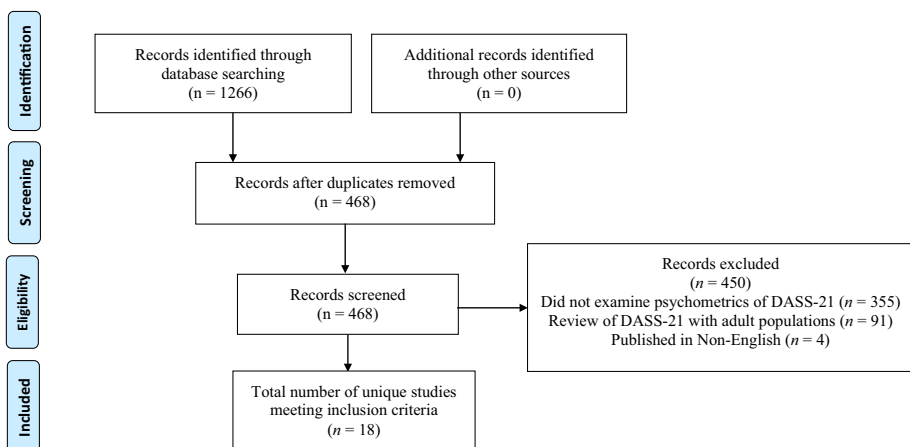


Fig. 1 Study inclusion flowchart

relevant to progress monitoring assessment with child or adolescent populations. Studies that primarily (i.e., > 50% of the sample) included data on adults (aged 18 and above) were excluded. One reviewer screened all titles and abstracts and, if needed, the full article, to determine eligibility for inclusion. Any questionable articles were discussed with a second reviewer until a consensus was reached. Screening resulted in the exclusion of 450 papers, due to the following exclusion criteria: (a) no examination of the psychometrics of the DASS-21: 355 articles (79%), (b) reviewed DASS-21 with adult populations only: 91 articles (20%), and (c) articles published in a language other than English: 7 articles (2%). After screening, 18 articles were included.

Coding Procedure and Analysis

After confirming eligible studies, each paper was reviewed to extract any psychometric evidence relevant to progress monitoring purposes (Briesch & Volpe, 2007). First, we looked for validity evidence based on internal structure by reviewing factor analytic findings. Specifically, we sought to determine whether the original three-factor model proposed by Lovibond and Lovibond (1995) was replicated in youth samples. Although factor analytic studies often compare multiple structural models, for the purposes of this study, we sought to identify the model identified by the authors as demonstrating the best model fit. Related, internal consistency reliability was coded if the authors assessed the degree of interrelatedness among the items either for the overall measure or for particular subscales. In reviewing factor analytic findings, we also coded whether each study provided evidence of measurement invariance across two or more demographic groups (e.g., gender, race, geography).

Second, we looked for validity evidence based on relations to other variables in the form of convergent and discriminant evidence. Convergent evidence was coded if the authors assessed the degree to which the scores from the DASS-21 related to scores on another measure to which it should be related. In contrast, discriminant evidence was coded if the authors assessed the degree to which the scores from the DASS-21 did not relate to scores on another scale that measured dissimilar constructs. Regarding the interpretation of coefficients, validity evidence is generally classified as weak ($r < 0.30$), moderate ($0.30 \geq r < 0.70$), or strong (≥ 0.70).

Finally, a psychometric criterion unique to progress monitoring is the responsiveness of a scale (Centers for Disease Control & Prevention, 2000). Scale responsiveness includes both whether data remain stable in the absence of an effect (i.e., test–retest reliability) and whether the measure is able to detect an actual effect when it occurs (i.e., sensitivity to change). Test–retest reliability is most often reported using Cronbach’s coefficient alpha (α), with values closer to 1.0 indicating stronger reliability. For the purposes of this review, coefficients between 0.60 and 0.70 indicated an acceptable level of reliability, whereas 0.8 or greater indicated a very good level. In cases where coefficient omega was reported, omega coefficient values above 0.8 were interpreted as good internal reliability. Sensitivity to change could be assessed using either single-measure or comparative approaches. Within a single-measure approach, the target measure is administered over time to a sample or individual whose behavior is expected to change to determine whether meaningful change is captured in the data (e.g., Hustus et al., 2020). Within a comparative approach, both the target measure and an external criterion measure are administered over time to determine whether similar change is captured by the two measures (e.g., Stratford et al., 1996).

Results

The 18 studies identified through the systematic review were published between 2005 and 2023 (see Table 1). Total study sample sizes varied extensively from 216 (Duffy et al., 2005) to 4,202 (Naumova, 2022) and may be explained, in part, by the fact that studies were split between those that required youth to complete the measure independently (e.g., online) and those that administered the measure in a group setting (e.g., during class time). The majority of studies focused on general samples of secondary school students (78%, $n = 14$), whereas four studies focused on college-age students. The greatest number of studies were published in Australia (28%, $n = 5$) and the United States (28%, $n = 5$), with three studies published in Europe, three in Asia, one in South America, and one containing a multi-national sample. Of the 18 identified studies, 1 (6%) reported on temporal stability, 13 (72%) reported on internal consistency, 16 (89%) reported on the factor structure of the measure, and 5 (28%) reported on convergent evidence. No studies investigated discriminant validity evidence or sensitivity to change of the DASS-21 with youth populations.

Table 1 Characteristics of the DASS-21 Studies Included in the Systematic Review

Study	N	Age <i>M (SD)</i>	Gender (% Female)	Target Population	Country
Anghel (2020)	255	21.97 (2.59)	51.0	Students in precollege program	Israel
Chin et al. (2019)					
Sample 1 (EFA)	426	19.7 (–)	59.9	Undergraduate students	USA
Sample 2 (CFA)	411	19.7 (–)	59.9	Undergraduate students	USA
Duffy et al. (2005)	216	13.13 (0.65)	55.4	Secondary school students	Australia
Evans et al. (2022)	306	15.35 (1.26)	54.0	Secondary school students	USA
Evans et al. (2021)	306	15.35 (1.26)	54.0	Secondary school students	USA
Jovanović et al. (2021)	1,906	16.54 (0.95)	60.7	Secondary school students	Serbia
Le et al. (2017)	1,745	16.5 (1.00)	46.3	Secondary school students	Vietnam
Mellor et al. (2015)					
Sample 1	371	14.56 (1.88)	51.0	Secondary school students	Australia
Sample 2	448	15.81 (1.38)	62.1	Secondary school students	Chile
Sample 3	558	15.92 (2.13)	62.2	Secondary school students	China
Sample 4	388	14.92 (1.79)	52.1	Secondary school students	Malaysia
Moore et al. (2017)	2,454	–	47.6	Secondary school students	USA
Naumova (2022)	4,202	16.43 (1.04)	64.9	Secondary school students	Macedonia
Norton (2007)	895	20.67 (4.15)	70.3	Undergraduate students	USA
Patrick et al. (2010)	425	15.89 (–)	51.6	Secondary school students	Australia
Shaw et al. (2017)	2,873	13.95 (1.86)	50.9	Secondary school students	Australia
Silva et al. (2016)	310	14.16 (2.12)	42.3	Secondary school students	Brazil
Szabó (2010)	484	13.62 (1.01)	48.0	Secondary school students	Australia
Tully et al. (2009)	4,039	14.56 (–)	53.0	Secondary school students	Australia
Willemsen et al. (2011)	677	13.80 (1.59)	55.8	Secondary school students	Belgium
Yap & Lee (2023)	400	18.8 (1.50)	52.2	Undergraduate students with temporomandibular disorders	Singapore

Note = – not provided

Validity Evidence Based on Internal Structure

Factor analysis of the DASS-21 has been conducted in 16 studies. EFA was used with 5 samples, CFA was used with 10 adolescent samples, and one study utilized both methods (see Table 2). Studies investigating the structural validity of the DASS-21 in adolescent samples have produced inconsistent findings. Extant research supports various internal structures of the DASS-21, including unidimensional, two-factor, DASS-consistent three-factor, and bifactor (i.e., tripartite, quadripartite) models.

Unidimensional Model

Two studies provided support for a unidimensional structure of the DASS-21 among adolescents. Patrick et al. (2010) and Shaw et al.'s (2017) analyses revealed that a single, general factor explained most of the common variance in the DASS-21 scores for adolescents and young people (almost 90%; Shaw et al., 2017). Both authors suggested that one factor should be extracted, indicating that the measure does not discriminate between depression, anxiety and stress in children and adolescents, but rather measures a single distress dimension.

Two-Factor Model

Two studies provided support for a two-factor model. Both Silva et al. (2016) and Yap and Lee (2023) conducted EFAs and found the best model fit when combining the items in the anxiety and stress scales into a single factor and the depression items in a second factor.

DASS-Consistent Three-Factor Model

The original three-factor model, with three correlated factors of depression, anxiety, and stress, has been supported across three studies (Anghel, 2020; Mellor et al., 2015; Norton, 2007). Additionally, Mellor et al. (2015) found evidence of measurement invariance across Australian, Chilean, Chinese, and Malaysian samples, suggesting that this three-factor model is supported across contexts. Norton et al. (2007) similarly found evidence of measurement invariance across racial groups (i.e., African American, Asian, Latinx, White); however, they did find that the strength of the factor intercorrelations varied across racial groups.

Bifactor Models

A number of studies have specified bifactor measurement models for the DASS-21. Bifactor measurement models account for both (a) a general factor, onto which all items load (e.g., distress) and (b) specific factors that account for some proportion of variance above and beyond the general factor (e.g., depression, anxiety, stress).

Tripartite Model. Although the DASS-21 was developed to discriminate between depression, anxiety, and stress as distinct states of negative emotionality (Lovibond & Lovibond, 1995), some studies have tested the fit of the tripartite model proposed by Clark and Watson (1991). This model suggests that whereas there are some distinctive symptoms of anxiety (i.e., physiological hyperarousal) and depression (i.e., anhedonia, lack of

positive affect), there are also many symptoms of negative affect (e.g., irritability, difficulty sleeping) that are common to both. Two studies provided support for the tripartite model. The tripartite model identified by Tully et al. (2009) contained a general factor (Negative Affectivity) on which all items were allowed to load and then two specific factors of depression and anxiety (anhedonia and physiological hyperarousal, respectively). This model was found to be invariant across younger (i.e., 12–14) and older (i.e., 15–18) adolescents. Willemsen et al. (2011) replicated the tripartite model in a sample of Belgian secondary school students and also found evidence of measurement invariance across males and females. In contrast, Duffy et al. (2005) found that a two-factor model comprising (1) physiological arousal and (2) generalized negativity (e.g., depression + negative affectivity) provided the best fit.

Quadripartite Model. Henry and Crawford (2005) first developed a quadripartite model, involving a common negative affectivity (NA) factor, as well as three specific factors of depression, anxiety, and stress when testing the construct validity of the DASS-21 for the general adult population. They assessed whether stress was synonymous with negative affectivity (NA) or if it represented a related, but separate, construct as it had remained unclear if the DASS stress subscale represented general psychological distress (i.e., NA) or a distinct stress component (Henry & Crawford, 2005). Data supported a quadripartite structure, including one general factor of psychological stress and three orthogonal factors of depression, anxiety, and stress.

Five studies have replicated the results of the quadripartite model with younger samples. Jovanović et al. (2021), Le et al. (2017), Naumova (2022), and Szabó (2010) reported this to be the best fitting model within their samples of secondary school students, though all noted that only the DASS-D and DASS-A contributed unique information beyond the general factor. Although Moore et al. (2017) also found strong support for the bifactor model with three specific factors, they reported that the specific factors failed to explain variance beyond the general factor of negative affectivity. Additionally, Jovanović et al. (2021) conducted tests of measurement invariance and found that invariance was supported across both gender and age.

Although Chin et al. (2019) attempted to replicate the quadripartite model, their research involving undergraduate students in the U.S. ultimately resulted in a unique bifactor model. Specifically, these authors found weak results for the anxiety and stress subscales, thus supporting a bifactor model that consisted of a general distress factor (onto which all 21 depression, anxiety, and stress items loaded) as well as a distinct 6-item depression factor.

Internal Consistency Reliability

Internal consistency coefficients (i.e., Cronbach's alpha, omega) were reported in nine studies for the depression, anxiety, and stress subscales and in seven studies for the DASS-21 Total Score (see Table 2). Coefficients ranged from good to excellent across samples for the DASS-D ($\alpha=0.72$ – 0.93), DASS-A ($\alpha=0.77$ – 0.90), DASS-S ($\alpha=0.70$ – 0.94), and DASS-T ($\alpha=0.83$ – 0.97). Overall, strong internal consistency was identified across scales.

Convergent Validity

To date, five studies have reported on the convergent validity of the DASS-21 when used with youth populations.

Table 2 Factor Structure and Internal Consistency of the DASS-21 Subscales and Total Score

Study	Method	Factors	Depression (α)	Anxiety (α)	Stress (α)	Total (α)
Anghel (2020)	EFA	D + A + S	.83-.86	.80-.82	.88-.90	–
Chin et al. (2019)	EFA & CFA	GF + D	–	–	–	–
Duffy et al. (2005)	CFA	GF + PA	–	–	–	–
Evans et al. (2022)	–	–	.92	.81	.81	–
Evans et al. (2021)	–	–	–	–	–	.93
Jovanović et al. (2021)	CFA	GF + D + A + S	.83	.79	.83	.91
Le et al. (2017)	CFA	GF + D + A + S	.72	.77	.70	–
Mellor et al. (2015)	CFA	D + A + S	.84	.75	.81	–
Moore et al. (2017)	CFA	GF + D + A + S	–	–	–	–
Naumova (2022)	CFA	GF + D + A + S	.93*	.90*	.90*	.97*
Norton (2007)	CFA	D + A + S	.83	.78	.87	–
Patrick et al. (2010)	EFA	GF	–	–	–	.83-.96**
Shaw et al. (2017)	CFA	GF	–	–	–	.94*
Silva et al. (2016)	EFA	D + A/S	.80	.80	.77	.88
Szabó (2010)	CFA	GF + D + A + S	–	–	–	–
Tully et al. (2009)	CFA	GF + D + A	.88	.79	.82	.93
Willemsen et al. (2011)	CFA	GF + D + A	–	–	–	–
Yap & Lee (2023)	EFA	D + A/S	.91	.81	.89	.95

Note. *omega coefficient; **Results reported separately by age and gender; A = anxiety; D = depression; GF = general factor; PA = physiological arousal S = stress

Depression Subscale. Four studies examined convergent validity by comparing the DASS-D subscale to other measures of depression or mental health. Evans et al. (2021) found strong correlations with the Center for Epidemiological Studies-Depression scale (CSED; Radloff, 1977) for both male (0.87) and female (0.86) secondary school students. Le et al. (2017) found the correlation with the Depression domain of the Duke Health Profile Adolescent Vietnamese validated version (ADHP-V; Parkerson et al., 1990) to also be statistically significant, though more moderate in strength (i.e., -0.58). Finally, both Jovanović et al. (2021) and Norton et al. (2007) found that the DASS-D was a significant predictor of scores on the Positive and Negative Affect Schedule-Positive Affect (PANAS-PA; Laurent et al., 1999; $\beta = -0.28$) and Beck Depression Inventory (BDI, Beck et al., 1996; $\eta^2 = 0.20$) respectively. 2007) (Table 3).

Anxiety Subscale. Four studies examined convergent validity by comparing the DASS-A subscale to other measures of anxiety or mental health. Evans et al. (2021) found moderate correlations with the GAD Screener (GAD7; Spitzer et al., 2006) for male (0.50) secondary school students, whereas correlations were stronger for female students (0.72). Le et al. (2017) found the correlation with the Anxiety domain of the ADHP-V (Parkerson et al., 1990) to also be statistically significant, though more moderate in strength (i.e., -0.50). Finally, Norton et al. (2007) found that the DASS-A was a stronger predictor of scores on the Beck Anxiety Inventory (BAI, Beck et al., 1988; $\eta^2 = 0.21$) than the PANAS-NA (Laurent et al., 1999; $\beta = -0.05$) when tested by Jovanović et al. (2021).

Stress Subscale. The stress subscale is intended to measure difficulty relaxing, tension, agitation, irritability, and negative affect. Three studies examined convergent

Table 3 Validity of the DASS-21 Subscales and Total Score

Study	Scale	DASS-A	DASS-S	DASS-T
Evans et al. (2021)	CESD: $r = 0.86^*$ (females) CESD: $r = 0.87^*$ (males)	GAD7: $r = .50^*$ (females) GAD7: $r = .72^*$ (males)		
Evans et al. (2022)				LSS: $r = .76^*$ (females) LSS: $r = .80^*$ (males)
Jovanović et al. (2021)	PANAS-PA: $\beta = -.28^*$ PANAS-NA: $\beta = .08^*$	PANAS-PA: $\beta = .03$ PANAS-NA: $\beta = -.05$	PANAS-PA: $\beta = .02$ PANAS-NA: $\beta = .07$	PANAS-PA: $\beta = -.43^*$ PANAS-NA: $\beta = .73^*$
Le et al. (2017)	ADHP-V Anxiety: $r = -.57^*$ ADHP-V Depression: $r = -.58^*$ ADHP-V Mental Health: $r = -.58^*$	ADHP-V Anxiety: $r = -.50^*$ ADHP-V Depression: $r = -.51^*$ ADHP-V General Health: $r = -.48^*$	ADHP-V Anxiety: $r = -.50^*$ ADHP-V Depression: $r = -.48^*$ ADHP-V General Health: $r = -.47^*$	ADHP-V Mental Health: $r = -.63^*$ ADHP-V General Health: $r = -.66^*$
Norton (2007)	BDI: $\eta^2 = .20^*$ BAI: $\eta^2 = .02^*$ PANAS-NA: $\eta^2 = .05^*$	BAI: $\eta^2 = .21^*$ BDI: $\eta^2 = .01^*$ PANAS-NA: $\eta^2 = .06^*$	BAI: $\eta^2 = .05^*$ BDI: $\eta^2 = .10^*$ PANAS-NA: $\eta^2 = .08^*$	

Note. $*p < .05$; ADHP-V = Duke Health Profile-Vietnamese; BAI = Beck Anxiety Inventory; BDI + Beck Depression Inventory; CESD = Center for Epidemiological Studies-Depression; GAD7 = Generalized Anxiety Disorder-7; LSS = Langner Symptom Survey; PANAS = Positive and Negative Affect Schedule

validity by comparing the DASS-21 stress subscale to measures of related constructs, including depression (e.g., ADHP-V, BDI), anxiety (e.g., ADHP-V, BAI), affect (e.g., PANAS-NA, PANAS-PA), and general health (e.g., ADHP-V). Moderate correlations were found with the ADHP-V Anxiety (-0.50), Depression (-0.48), and General Health (-0.47) subscales (Le et al., 2017). Norton (2007) found that the DASS-S significantly predicted scores on the BDI ($\eta^2=0.10$), PANAS-NA ($\eta^2=0.08$), and BAI ($\eta^2=0.05$). Jovanović et al. (2021) found that the DASS-S was not a significant predictor of scores on the PANAS-PA ($\beta=0.02$) and PANAS-NA ($\beta=0.07$).

Total Score. Three studies examined the convergent validity of the total score. Evans et al. (2021) found strong correlations with the Langner Symptom Survey, LSS; Langner, 1962) for both male (0.80) and female (0.76) secondary school students. Le et al. (2017) found correlations with the Mental Health and General Health domains of the ADHP-V (Parkerson et al., 1990) to also be statistically significant, though more moderate in strength (i.e., Mental Health = -0.63, General Health = -0.66). Finally, Jovanović et al. (2021) found the DASS-T to be a more significant predictor of scores on the PANAS (Laurent et al., 1999) PA ($\beta=-0.43$) and NA ($\beta=0.73$) than the individual subscales.

Test–Retest Reliability

One study reported evidence of test–retest reliability of the DASS-21. Silva et al. (2016) administered the DASS-21 on two occasions one week apart and found intraclass correlation coefficients to be very strong across the subscales (DASS-D=0.86; DASS-A=0.80; DASS-S=0.82).

Discussion

Although it has been suggested that the DASS-21 may be appropriate to use with youth in a progress monitoring context, much of the psychometric evidence has been obtained with adults to date. The purpose of this paper was therefore to review the growing literature pertaining to the psychometric properties of the DASS-21 when used specifically with youth populations to clarify the applied utility of the measure. The systematic review yielded 18 papers meeting inclusion criteria and helps to elucidate those areas in which psychometric evidence supporting use of the DASS-21 with youth populations is robust, as well as those in which additional inquiry appears warranted.

First and foremost, understanding the factor structure of the measure is critically important, as it provides guidance regarding how the items should be combined and summarized. Although the DASS-21 was designed to assess the three subscales of depression, anxiety, and stress in adults (Lovibond & Lovibond, 1995), studies utilizing the measure with youth participants have found varied factor structures (i.e., ranging from 1–4 factors). The greatest support (i.e., 31% of studies) was found for the hierarchical quadripartite model that involves a common negativity factor, as well as the three specific factors of depression, anxiety, and stress. It is notable, however, that several of the studies reviewed either found insufficient support for the three specific factors (supporting instead use of the total score; e.g., Patrick et al., 2010; Shaw et al., 2017) or only found support for the depression and anxiety scales (e.g., Silva et al., 2016; Tully et al., 2009). Related, Le et al. (2017) found that the DASS-D, DASS-A, and DASS-S correlated as strongly with measures of

the same internalizing domain as they did with measures of other internalizing domains. Taken together, these results seem to suggest that there may not be adequate differentiation between the subscales of the DASS-21. As such, the accumulated evidence seems to support use of the DASS-21 total score while further research is needed to understand the utility of the subscales.

A second notable finding was that few studies to date have established validity evidence based on relations to other variables when using the DASS-21 with youth. Understanding whether the DASS-21 scales are correlated with other measures of the same construct (as well as not correlated with measures of other constructs) is important, as it provides evidence that the measure is actually assessing what it purports to. Across domains, there appears to be the strongest convergent evidence for the total score (DASS-T). That is, three studies highlighted moderate to strong relationships between the DASS-T and overall measures of mental health and distress. When the depression and anxiety subscales were directly compared to extant measures of the same constructs (e.g., BDI, GAD7), the strength of these relationships was found to vary across these three studies ($r=0.50-0.87$). The range of these correlations is consistent, however, with evidence found with adult populations (e.g., Bibi et al., 2020; Oei et al., 2013). Although moderate associations have been identified between the stress subscale and other measures of mental (e.g., anxiety, depression) and general health, it is important to note that no studies to date have specifically examined correlations with established measures of stress.

Third, virtually no evidence was identified in support of the sensitivity to change of the DASS-21 when used with youth populations. Although reliability and validity evidence are relevant across all assessment purposes, sensitivity to change is uniquely relevant to progress monitoring. This indicator is important because if the DASS-21 is not sensitive to change, this means that it is unable to detect when meaningful changes occur as the result of treatment. The one study to report test–retest reliability found that subscale scores were highly consistent when the DASS-21 was administered one week apart; however, the study used a non-clinical youth sample who were not receiving treatment (Silva et al., 2016). Studies are therefore needed to understand whether the DASS-21 captures changes that occur as a function of treatment.

Finally, it is important to consider that differences in contextual and cultural factors may limit the generalizability of evidence on the psychometric properties of the DASS-21 to other youth populations. Although the current review found that structural validity has been examined across multiple, non-English speaking countries (i.e., Belgium, Brazil, Chile, China, Israel, Macedonia, Malaysia, Serbia, Singapore, Vietnam), additional evidence in support of reliability (e.g., internal consistency, test–retest) was not consistently presented and validity evidence was present in only two cases (i.e., Jovanovic et al., 2021; Le et al., 2017). As such, additional research is warranted to ensure that the DASS-21 operates similarly across international contexts.

Limitations and Recommendations for Future Research

The current review has limitations that should be acknowledged. First, the review excluded non-English papers, as well as unpublished work, which may have contributed relevant information. Second, most of the studies reviewed employed general, non-clinical samples. Although general samples are appropriate for screening purposes, it may be useful to investigate clinical samples of youth to provide support for use within progress monitoring.

Third, most of the studies that evaluated internal consistency (i.e., all but one) used Cronbach's alpha, which makes strong assumptions of unidimensionality and equal factor loadings. It was evident in the review of factor analysis studies that these assumptions were not always met.

Conclusions and Implications for Practice

Extant research has demonstrated that outcome measurement can improve the quality of care and improve the percentage of clients whose functioning improves during treatment (Blais et al., 2009). The DASS-21 has the potential to be a feasible means of monitoring youth functioning over time, given that it is both brief and freely available in the public domain. The majority of studies reporting psychometric properties in this review utilized the DASS-21 with mixed-gender groups of secondary school students, thus supporting its use with similar populations. Pending additional inquiry, however, practitioners are encouraged to utilize the total score, for which stronger evidence of reliability and validity exist. Additionally, future research focused on examining the sensitivity to change of DASS-21 data would help to support use with youth for clinical progress monitoring purposes.

Funding Open access funding provided by Northeastern University Library.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose

Human and animal rights As this study did not involve human participants, the University HSIRB confirmed that no ethical approval was required.

Informed Consent As this study did not involve human participants, informed consent was not applicable

Data Availability The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anghel, E. (2020). Longitudinal invariance and information of the Depression, Anxiety, and Stress Scales. *Journal of Clinical Psychology, 76*(10), 1923–1937.
- American Psychological Association Presidential Task Force on Evidence-Based Practice. Force, (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271–285.
- Bados, A., Solanas, A., & Andrés, R. (2005). Psychometric properties of the Spanish version of Depression, Anxiety and Stress Scales (DASS) [Propiedades psicométricas de la versión española de las Escalas de Depresión, Ansiedad y Estrés (DASS)]. *Psicothema, 17*(4), 679–683.

- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*(6), 893–897.
- Beck, A. T., Steer, R. A., & Brown, G. (1996). *Beck Depression Inventory-II*. Psychological Assessment. The Psychological Corporation.
- Bibi, A., Lin, M., Zhang, Z. C., & Margraf, J. (2020). Psychometric properties and measurement invariance of Depression, Anxiety, and Stress Scales (DASS-21) across cultures. *International Journal of Psychology, 55*, 916–925.
- Blais, M. A., Frank, R. G., Nierenberg, A. A., & Rauch, S. L. (2009). Measuring outcomes makes good sense. *Psychiatric Services, 60*(1), 112–112.
- Briesch, A. M. & Volpe, R. J. (2007). Important considerations in the selection of progress-monitoring measures for classroom behaviors. *School Psychology Forum, 1*(2), 59–74
- Cao, C., Liao, X., Gamble, J. H., Li, L., Jiang, X., Li, X., Griffiths, M. D., Chen, I., & Lin, C. (2023). Evaluating the psychometric properties of the Chinese Depression Anxiety Stress Scale for Youth (DASS-Y) and DASS-21. *Child and Adolescent Psychiatry and Mental Health, 17*, 106. <https://doi.org/10.1186/s13034-023-00655-2>
- Centers for Disease Control and Prevention. (2000). *Measuring healthy days*. Atlanta, Georgia: CDC.
- Chin, E. G., Buchanan, E. M., Ebesutani, C., & Young, J. (2019). Depression, anxiety, and stress: How should clinicians interpret the Total and subscale scores of the 21-item Depression, Anxiety, and Stress Scales? *Psychological Reports, 122*(4), 1550–1575
- Chorpita, B. F., Reise, S., Weisz, J. R., Grubbs, K., Becker, K. D., & Krull, J. L. (2010). Evaluation of the brief problem checklist: Child and caregiver interviews to measure clinical progress. *Journal of Consulting and Clinical Psychology, 78*, 526–536.
- Clara, I. P., Cox, B. J., & Enns, M. W. (2001). Confirmatory factor analysis of the Depression Anxiety and Stress Scales in depressed and anxious patients. *Journal of Psychopathology and Behavioral Assessment, 23*(1), 61–67.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316–336.
- Dart, E. H., Arora, P. G., Collins, T. A., & Doll, B. (2019). Progress monitoring measures for internalizing symptoms: A systematic review of the peer-reviewed literature. *School Mental Health, 11*(2), 265–275.
- Duffy, C. J., Cunningham, E. G., & Moore, S. M. (2005). Brief report: The factor structure of mood states in an early adolescent sample. *Journal of Adolescence, 28*(5), 677–680.
- Evans, L., Haerberlein, K., Chang, A., & Handa, P. (2022). An evaluation of the convergent validity of and preliminary cutoff scores for the DASS-21 Total score as a measure of distress in adolescents. *Current Psychology, 41*, 4283–4290.
- Evans, L., Haerberlein, K., Chang, A., & Handal, P. (2021). Convergent validity and preliminary cut-off scores for the Anxiety and Depression subscales of the DASS-21 in US adolescents. *Child Psychiatry & Human Development, 52*, 579–585.
- 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.
- Hustus, C. L., Owens, J. S., Volpe, R. J., Briesch, A. M., & Daniels, B. (2020). Treatment sensitivity of direct behavior rating—multi-item scales in the context of a daily report card intervention. *Journal of Emotional and Behavioral Disorders, 28*(1), 29–42. <https://doi.org/10.1177/1063426618806281>
- In-Albon, T., Meyer, A. H., & Schneider, S. (2013). Separation anxiety avoidance inventory-child and parent version: Psychometric properties and clinical utility in a clinical and school sample. *Child Psychiatry and Human Development, 44*, 689–697. <https://doi.org/10.1007/s10578-013-0364-z>
- Jovanović, V., Gavrilov-Jerković, V., & Lazić, M. (2021). Can adolescents differentiate between depression, anxiety and stress? Testing competing models of the Depression Anxiety Stress Scales (DASS-21). *Current Psychology, 40*(12), 6045–6056.
- Kelley, S. D., Vides de Andrade, A. R., Bickman *Jovanović, V., Gavrilov-Jerković, V., & Kendall, P. C., Compton, S. N., Walkup, J. T., Birmaher, B., Albano, A. M., Sherrill, J., Ginsburg, G., Rynn, M., McCracken, J., Gosch, E., Keeton, C., Bergman, L., Sakolsy, D., Iyengar, S., March, J., & Piacentini, J. (2010). Clinical characteristics of anxiety disordered youth. *Journal of Anxiety Disorders, 24*(3), 360–365.
- Kendall, P. C., Compton, S. N., Walkup, J. T., Birmaher, B., Albano, A. M., Sherrill, J., Ginsburg, G., Rynn, M., McCracken, J., Gosch, E., Keeton, C., Bergman, L., Sakolsy, D., Iyengar, S., March, J., & Piacentini, J. (2010). Clinical characteristics of anxiety disordered youth. *Journal of Anxiety Disorders, 24*(3), 360–365.
- Kovacs, M. (2011). *Children's depression inventory (2nd Edition)*. Multi-Health Systems.

- Kumar, K., Kumar, S., Mehrotra, D., Tiwari, S. C., Kumar, V., & Dwivedi, R. C. (2019). Reliability and psychometric validity of Hindi version of Depression, Anxiety and Stress Scale-21 (DASS-21) for Hindi speaking head neck cancer and oral potentially malignant disorders patients. *Journal of Cancer Research and Therapeutics*, *15*(3), 653–658.
- Langner, T. S. (1962). A twenty-two item screening score of psychiatric symptoms indicating impairment. *Journal of Health and Human Behavior*, *3*(4), 269–276.
- Laurent, J., Catanzaro, S. J., Joiner, T. E., Jr., Rudolph, K. D., Potter, K. I., Lambert, S., Osborne, L., & Gathright, T. (1999). A measure of positive and negative affect for children: Scale development and preliminary validation. *Psychological Assessment*, *11*(3), 326–338.
- Le, M. T. H., Tran, T. D., Holton, S., Nguyen, H. T., Wolfe, R., & Fisher, J. (2017). Reliability, convergent validity and factor structure of the DASS-21 in a sample of Vietnamese adolescents. *PLoS ONE*, *12*(7), e0180557.
- Lopez, B., Turner, R. J., & Saavedra, L. M. (2005). Anxiety and risk for substance dependence among late adolescents/young adults. *Journal of Anxiety Disorders*, *19*(3), 275–294.
- Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales (DASS)*. (2nd Ed.) Psychology Foundation of Australia.
- March, J. S. (2013). *Multidimensional anxiety scale for children (2nd edition)*. Multi-Health Systems.
- Mellor, D., Vinet, E. V., Xu, X., Mamat, N. H. B., Richardson, B., & Román, F. (2015). Factorial invariance of the DASS-21 among adolescents in four countries. *European Journal of Psychological Assessment*, *31*(2), 138–142.
- Merrell, K. W., & Gueldner, B. A. (2010). Preventative interventions for students with internalizing disorders: Effective strategies for promoting mental health in schools. In M. R. Shinn & H. M. Walker (Eds., pp. 729–823), *Interventions for Achievement and Behavior in a Three-Tier Model Including RTI* (3rd ed.). National Association of School Psychologists.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269.
- Moore, S. A., Dowdy, E., & Furlong, M. J. (2017). Using the Depression, Anxiety, Stress Scales–21 with US adolescents: An alternate models analysis. *Journal of Psychoeducational Assessment*, *35*(6), 581–598.
- Naumova, K. (2022). Dimensionality and reliability of the Depression Anxiety Stress Scales 21 among adolescents in Macedonia. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2022.1007594>
- Nelson, J. R., Benner, G. J., Lane, K., & Smith, B. W. (2004). Academic Achievement of K-12 Students With Emotional and Behavioral Disorders. *Exceptional Children*, *71*(1), 59–73. <https://doi.org/10.1177/001440290407100104>
- Ng, F., Trauer, T., Dodd, S., Callaly, T., Campbell, S., & Berk, M. (2007). The validity of the 21-item version of the Depression Anxiety Stress Scales as a routine clinical outcome measure. *Acta Neuropsychiatrica*, *19*(5), 304–310.
- Norton, P. J. (2007). Depression Anxiety and Stress Scales (DASS-21): Psychometric analysis across four racial groups. *Anxiety, Stress & Coping*, *20*, 253–265.
- Oei, T. P. S., Sawang, S., Goh, Y. W., & Mukhtar, F. (2013). Using the Depression Anxiety Stress Scale 21 (DASS-21) across cultures. *International Journal of Psychology*, *48*, 1018–1029.
- Parkerson, G. R., Jr., Broadhead, W. E., & Tse, C. K. J. (1990). The Duke Health Profile: A 17-item measure of health and dysfunction. *Medical Care*, *28*, 1056–1072.
- Patrick, J., Dyck, M., & Bramston, P. (2010). Depression Anxiety Stress Scale: Is it valid for children and adolescents? *Journal of Clinical Psychology*, *66*(9), 996–1007. Psychology Foundation of Australia (n.d.). *Depression Anxiety Stress Scales*. <http://www2.psy.unsw.edu.au/dass/>.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401.
- Reynolds, W. M. (2004). *Reynolds adolescent depression scale (2nd edition)*. PAR.
- Reynolds, C. R., & Richmond, B. O. (2008). *Revised children's manifest anxiety scale: Second edition (RCMAS-2)*. WPS.
- Shaw, T., Campbell, M. A., Runions, K. C., & Zubrick, S. R. (2017). Properties of the DASS-21 in an Australian community adolescent population. *Journal of Clinical Psychology*, *73*(7), 879–892.
- Silva, H. A. D., Passos, M. H. P. D., Oliveira, V. M. A. D., Palmeira, A. C., Pitanguí, A. C. R., & Araújo, R. C. D. (2016). Short version of the Depression Anxiety Stress Scale-21: Is it valid for Brazilian adolescents? *Einstein (são Paulo)*, *14*(4), 486–493.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097.

- Stratford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the minimum level of detectable change for the Roland-Morris questionnaire. *Physical Therapy, 76*, 359–365. <https://doi.org/10.1093/ptj/76.4.359>
- Szabó, M. (2010). The short version of the Depression Anxiety Stress Scales (DASS-21): Factor structure in a young adolescent sample. *Journal of Adolescence, 33*(1), 1–8.
- Tully, P. J., Zajac, I. T., & Venning, A. J. (2009). The structure of anxiety and depression in a normative sample of younger and older Australian adolescents. *Journal of Abnormal Child Psychology, 37*(5), 717–726.
- Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2000). Outcome questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*(2), 242–261.
- Walker, H. M., Nishioka, V. M., Zeller, R., Severson, H. H., & Feil, E. G. (2000). Causal factors and potential solutions for the persistent under-identification of students having emotional or behavioral disorders in the context of schooling. *Assessment for Effective Intervention, 26*, 29–39. <https://doi.org/10.1177/073724770002600105>
- Willemsen, J., Markey, S., Declercq, F., & Vanheule, S. (2011). Negative emotionality in a large community sample of adolescents: The factor structure and measurement invariance of the short version of the Depression Anxiety Stress Scales. *Stress and Health, 27*(3), e120–e128.
- Yap, A. U., & Lee, D. Z. R. (2023). Use of the Depression Anxiety Stress Scale-21 in young people with temporomandibular disorders: Reliability and dimensionality. *The Journal of Craniomandibular & Sleep Practice*. <https://doi.org/10.1080/08869634.2022.2158259>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.