

# Measurement Invariance of the GAD-7 and CESD-R-10 Among Adolescents in Canada

Isabella Romano,<sup>1</sup> BSc, Mark A. Ferro,<sup>1</sup> PhD, Karen A. Patte,<sup>2</sup> and Scott T. Leatherdale<sup>1</sup> PhD

<sup>1</sup>School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada, and <sup>2</sup>Department of Health Sciences, Brock University, St. Catharines, ON, Canada

All correspondence concerning this article should be addressed to Isabella Romano, BSc, School of Public Health Sciences, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. E-mail: iromano@uwaterloo.ca

Received 9 April 2021; revisions received 15 October 2021; accepted 17 October 2021

## Abstract

**Objective** The primary objective of our study was to assess measurement invariance (by grade and sex) of the 7-item Generalized Anxiety Disorder (GAD-7) and 10-item Center for Epidemiologic Studies Depression Revised (CESD-R-10) scales in a sample of adolescents in Canada. If measurement invariance was demonstrated, our secondary objective was to estimate differences in scale scores across these subgroups. **Methods** We used data from 59,052 adolescents in Year 7 (2018–19) of the COMPASS school-based study. Measurement invariance was tested within a multigroup confirmatory factor analysis framework. Differences in scale scores were estimated using mixed linear regression which accounted for school-level clustering and adjusted for relevant confounders. **Results** Both the GAD-7 and CESD-R-10 demonstrated strict measurement invariance by sex and grade in our sample. Mean scale scores were higher among adolescents in grade 12 compared to grade 9 ( $\beta_{\text{GAD-7}} = 0.91, p < .001$ ;  $\beta_{\text{CESD-R-10}} = 0.99, p < .001$ ) and among female adolescents compared to males ( $\beta_{\text{GAD-7}} = 3.36, p < .001$ ;  $\beta_{\text{CESD-R-10}} = 2.58, p < .001$ ). **Conclusions** Findings support the validity of the GAD-7 and CESD-R-10 for screening risk of generalized anxiety and depression among Canadian adolescents, and further indicate that differences observed in scale scores across subgroups reflect actual differences in risk for generalized anxiety and major depression, respectively.

**Key words:** adolescents; anxiety; depression; measurement invariance.

## Introduction

Approximately 12% of individuals will experience a major depressive episode during their adolescence, and nearly one-third of adolescents are estimated to experience an anxiety disorder (Merikangas, 2010). Compared to males, female adolescents are disproportionately impacted by internalizing disorders including generalized anxiety and major depression (Rosenfield & Mouzon, 2013; Van Droogenbroeck et al., 2018), and report experiencing higher levels of psychological distress (Boak et al., 2018). Male youth, on the other hand, are generally more likely to experience

externalizing disorders (Rosenfield & Mouzon, 2013). Older adolescent age is another risk factor contributing to differences in mental health (Boak et al., 2018; Kilburn et al., 2018), although trajectories are reportedly steeper among females than males (Fink et al., 2015; Van Droogenbroeck et al., 2018). To further understand sex- and age-related differences in adolescents' mental health, it is necessary that screening tools are appropriate for drawing valid conclusions regarding risk in research and pediatric psychology practice. Establishing validity is also vital to support screening efforts in general and specialty pediatric

clinics, where 50% of pediatric patients with physical illness may experience poor mental health (Butler et al., 2018) and upwards of 30% meet criteria for clinical psychiatric diagnoses such as generalized anxiety or major depression (Tegethoff et al., 2015).

The 7-item Generalized Anxiety Disorder (GAD-7 [Spitzer et al., 2006]) and the 10-item Center for Epidemiologic Studies Depression Revised (CESD-R-10 [Andresen et al., 1994]) scales represent two of the most widely used tools to screen for risk of depression and anxiety, respectively, within clinical and research contexts. Studies have reported use of these tools to screen for anxiety and depressive symptoms in pediatric health settings including among pediatric endocrinology (Benson et al., 2020; Moyer et al., 2019), neurology (Hellebrekers et al., 2019; Martinović et al., 2006), and organ transplant (McCormick et al., 2020) patients. The GAD-7 has demonstrated validity for use among general adult populations (Löwe et al., 2008; Hinz et al., 2017) and, while Radloff's (1977) original 20-item CESD has been validated extensively, fewer studies have focused on Andresen et al.'s (1994) shortened CESD-R-10. Björgvinsson et al. (2013), for example, found the CESD-R-10 to have strong psychometric properties for screening depression severity in a clinical sample.

However, there have been a paucity of studies assessing the psychometric properties of these tools among child and adolescent populations, despite being frequently used in research and within pediatric health settings (including specialty care clinics) to screen for anxiety and depressive symptoms. Two studies have explored the psychometric properties of the GAD-7 for screening for adolescent anxiety in the general population; first among a large sample of adolescents in Finland (Tiirikainen et al., 2019), followed by Adjorlolo (2019) who assessed validity among adolescents in Ghana. Findings from both studies supported the reliability, factorial validity, and construct validity of the GAD-7 in this age group. Among U.S. adolescents diagnosed with generalized anxiety disorder, the GAD-7 further demonstrated sufficient specificity and sensitivity for detecting symptoms (Mossman et al., 2017). With regard to the measurement of depressive symptoms, Bradley et al. (2010) investigated the factorial validity of the CESD-R-10 among a small community-based sample of adolescents in Canada, and their findings supported its use as a screening tool for adolescents in the community. Similar findings have been noted among a smaller sample of adolescents in France (Cartierre et al. 2011).

Relating to psychometric validation, *measurement invariance* refers to a statistical property indicating whether a given scale measures the same latent construct across subgroups within a sample (Brown, 2015). Establishing measurement invariance is a prerequisite

for making meaningful group-level comparisons (Brown, 2015), whereas measurement *non*-invariance is indicative of bias due to differences between subgroups in their interpretation of the items comprising a scale (van de Schoot et al., 2012). In the latter case, observed differences in scores may be artifactual and invalid. Non-invariance of tools used to screen for adolescent anxiety and depression can pose negative implications for pediatric psychology contexts; if scale items were to be interpreted differently by females and males, conclusions made might result in ill-informed screening efforts or mistargeted interventions.

Only a handful of studies have assessed measurement invariance of either the GAD-7 or the CESD-R-10 among pediatric samples, and none within a Canadian adolescent context. Only one previous study has confirmed invariance by sex of the GAD-7 among an adolescent sample (Adjorlolo, 2019), but findings from this Ghanaian sample may not generalize adequately to adolescents in Canada given observed discrepancies in the experience of mental disorders such as anxiety across cultures (Schreier et al., 2010). Existing evidence also supports full invariance of the CESD-R-10 by sex among French adolescents (Cartierre et al., 2011) and youth in sub-Saharan Africa (Kilburn et al., 2018). One previous version of the CESD-R-10 has been validated for use specifically among adolescents, albeit in the United States, and found no evidence of inequivalence for gender or sex within their sample (Haroz et al., 2014). No previous studies have examined measurement invariance of the GAD-7 or CESD-R-10 across age groups (e.g., secondary school grade) among adolescents.

## Objectives and Hypotheses

Using survey data from a large sample of adolescents across Canada, our primary objective was to test measurement invariance of the GAD-7 and CESD-R-10 by sex and grade. If measurement invariance was demonstrated, our secondary objective was to estimate the extent to which differences in GAD-7 and CESD-R-10 scores can be predicted by sex and grade among adolescents in Canada. In light of existing (albeit limited) available evidence, we hypothesized that the GAD-7 and CESD-R-10 would demonstrate full measurement invariance within our study sample of Canadian adolescents. In the presence of invariance, we further hypothesized that GAD-7 and CESD-R-10 scale scores would be higher among females and adolescents in older grades.

## Methods

### Study Design

We conducted this study using data from Year 7 (Y<sub>7</sub>[2018–19]) of the COMPASS study, a prospective

cohort study (2012–21) of adolescents attending secondary schools across Alberta, British Columbia, Ontario, and Quebec, Canada (Leatherdale et al., 2014). Schools were purposefully recruited based on permitted use of active information, passive consent data collection protocols, which are shown to support robust reporting of mental health among adolescents (Thompson-Haile et al., 2013). All adolescents attending participating secondary schools are eligible to participate. The COMPASS questionnaire (Cq) is a paper-and-pencil-based survey designed to be completed once annually during class time. In Y<sub>7</sub>, 74,501 participants across 136 schools ( $n_{AB} = 8$ ;  $n_{BC} = 15$ ;  $n_{ON} = 51$ ;  $n_{QC} = 52$ ) participated in COMPASS. We received approval from our institution's Office of Research Ethics and participating school boards. Additional details on the study design and methods are available in print (Leatherdale et al., 2014) and online ([www.compass.uwaterloo.ca](http://www.compass.uwaterloo.ca)).

### Instrument

The COMPASS mental health module (MH-M; Patte et al., 2017a, b) included the GAD-7 (Spitzer et al., 2006) to measure symptoms of generalized anxiety. Within the past 2 weeks, participants were asked to indicate the number of days they felt the following symptoms: 1) *feeling nervous, anxious, or on edge*, 2) *not being able to stop or control worrying*, 3) *worrying too much about different things*, 4) *trouble relaxing*, 5) *being so restless that it's hard to sit still*, 6) *becoming easily annoyed or irritable*, and 7) *feeling afraid as if something awful might happen*. Responses were scored on a 4-point Likert scale (0 = *not at all*, 1 = *several days*, 2 = *over half the days*, 3 = *nearly every day*) with sum scores ranging from 0 to 21 (higher scores indicating greater symptoms). Internal consistency in our sample was  $\alpha = 0.91$ .

The CESD-R-10 (Zhang et al., 2012) was included on the Cq MH-M (Patte et al., 2017a, b) as a measure of depressive symptoms. On a 4-point Likert scale (0 = *none or less than 1 day*, 1 = *1–2 days*, 2 = *3–4 days*, 3 = *5–7 days*), participants were asked to indicate how often they experienced each of the following in the past week: 1) *I was bothered by things that usually don't bother me*, 2) *I had trouble keeping my mind on what I was doing*, 3) *I felt depressed*, 4) *I felt that everything I did was an effort*, 5) *I felt hopeful about the future*, 6) *I felt fearful*, 7) *my sleep was restless*, 8) *I was happy*, 9) *I felt lonely*, and 10) *I could not get "going."* Individual sum scores ranged from 0 to 30 with higher scores representing greater symptomatology; items 5 and 8, which measure positive affect, are reverse-coded. Internal consistency across CESD-R-10 items within our study sample was  $\alpha = 0.83$ .

Participants responded to the question, "are you male or female" on the Cq and selected their sex from

*male* or *female* response options. Participants also indicated their school grade (9, 10, 11, 12) and age (in years), which were highly correlated ( $r = 0.85$ ). We relied on grade for the current analyses given its relevance to our study sample and correlation with age. Participants were also asked to indicate their available weekly spending/saving money (\$0, \$1–20, \$20–100, \$100+). Participants reported their ethno-racial identity, categorized as racialized (Black, Asian, Indigenous [First Nations, Métis, Inuit], Latin American/Hispanic, and/or other) or non-racialized (White) to more accurately describe race and ethnicity as a social construct, rather than as biological fact (Ross et al., 2020).

### Procedure

#### School-Level Variation

To account for potential variability introduced by clustering within schools, we first calculated the intra-class correlation coefficient (ICC) as an indicator of shared variation in participants' GAD-7 and CESD-R-10 sum scores. We measured the presence of within-school clustering for each measure ( $ICC_{GAD-7} = 0.036$ ,  $ICC_{CESD-R-10} = 0.025$ ) and adjusted for these effects within our analyses.

#### Measurement Invariance

Using a multigroup confirmatory factor analysis (CFA), our procedure for testing measurement invariance of the GAD-7 and CESD-R-10 incorporated four sequential steps requiring increasingly stringent equality constraints on between-group (i.e., males vs. females, grade 9 vs. 12) model parameters. First, we tested configural invariance to confirm consistency of the GAD-7 and CESD-R-10 factor structures between groups; no equality constraints were imposed on the configural model in this step (Byrne, 2012). Second, as a prerequisite for making valid group comparisons (Bollen, 1989), we established in the metric model whether factor loadings for each GAD-7 and CESD-R-10 item were equivalent between our groups. The third step verified whether mean item-level differences were fully explained by mean factor-level differences by testing for equivalent item intercepts in the scalar model (Brown, 2015). In the fourth step, a strict model was specified to test invariance across item residuals as a requisite for defensible item-level comparisons (Steinmetz, 2009). This four-step approach to systematically imposing constraints allowed us to identify which specific parameters contributed to model misfit and, in turn, to differences in interpretation (van de Schoot et al., 2012).

Criteria by which we established full measurement invariance were determined *a priori*. First was adequate model fit at each level of testing as indicated by at least two (Tompke et al., 2020; Tompke & Ferro,

2019) of the following: the comparative fit index (CFI), square root mean residual (SRMR), or root mean standard error of approximation (RMSEA). Established cut-points for these indices were CFI  $\geq 0.950$ ; SRMR  $\leq 0.080$ ; and RMSEA  $\leq 0.080$  (Brown, 2015). The second criterion required that changes in fit indices did not exceed cut-points in, again, at least two cases of the following:  $\Delta$ CFI  $\leq -0.010$ ,  $\Delta$ SRMR  $\geq 0.030$ , or  $\Delta$ RMSEA  $\geq 0.015$  (Chen, 2007). Given our large sample size, we did not rely on  $\chi^2$  goodness-of-fit and  $\Delta\chi^2$  as indices of model fit (Cheung & Rensvold, 2002). We conducted measurement invariance tests by adolescents' sex and grade in Mplus version 8.6 (Muthén & Muthén, 1998–2020) with categorical estimation (weighted least squares means and variance adjusted estimator).

### Differences in Sum Scores

If measurement invariance was established in the GAD-7 and/or CESD-R-10, we tested for differences in sum scores across sex and grade using mixed linear regression. We reported  $\beta$ -estimates from our model output alongside 95% confidence intervals (CIs), adjusted for participants' ethno-racial identity, province, and weekly spending money (as a proxy for individual-level socioeconomic status and/or part-time employment). We used SAS version 9.4 (SAS Institute, 2016) and accounted for school-level clustering using a multilevel model in PROC MIXED.

## Results

### Sample Characteristics

There were 59,052 adolescents in our analytic sample, among whom half were female (49.8%). Roughly one-third of adolescents reported a racialized identity (4.1% Black, 12.1% Asian, 2.9% Latin American/Hispanic, 15% other or Mixed) while a majority were non-racialized (White). Participants ranged in age from 12 to 19 years (mean = 15.6 years, SD = 1.2 years). The mean sum scores were 6.5 (SD = 5.7, range: 0–21) for the GAD-7 and 9.1 (SD = 6.1, range: 0–30) for the CESD-R-10.

Descriptive statistics for the GAD-7 and CESD-R-10 measures are presented in Table I. One-way ANOVA tests showed significant variation in both measures for sex and grade subgroups. *Post hoc* pairwise comparisons showed that females differed significantly from males in GAD-7 and CESD-R-10 mean scores at  $p < .0001$ . For grade, GAD-7 and CESD-R-10 mean scores were significantly different across subgroups at  $p < .0001$  except for participants in Grades 10 and 11. Additional descriptive statistics are provided in Supplementary File A (Tables S1 and S2).

### Measurement Invariance of the GAD-7 and CESD-R-10

The baseline CFA model demonstrated overall excellent fit for the GAD-7 (CFI = 0.980; SRMR = 0.026; RMSEA = 0.051 [90% CI 0.047–0.053]) in the total sample. The CESD-R-10 also demonstrated excellent fit (CFI = 0.984; SRMR = 0.028; RMSEA = 0.024 [90% CI 0.023–0.025]), once correlations were specified amongst residuals for items 3–5, 3–8, 3–9, 5–8, and 8–9. Supplementary File B (Table S3) presents fit statistics for all baseline models by subgroup.

Measurement invariance testing results of the GAD-7 and CESD-R-10 by sex and grade are shown in Table II. The configural models for sex and grade demonstrated excellent fit in the GAD-7 (sex: CFI = 0.978; SRMR = 0.027; RMSEA = 0.069 [90% CI 0.067–0.070]; grade: CFI = 0.978; SRMR = 0.028; RMSEA = 0.098 [90% CI 0.096–0.101]) and the CESD-R-10 (sex: CFI = 0.979; SRMR = 0.029; RMSEA = 0.035 [90% CI 0.034–0.036]; grade: CFI = 0.975; SRMR = 0.034; RMSEA = 0.046 [90% CI 0.044–0.048]). Equality constraints imposed on the factor loadings (metric model) did not result in a substantially worse fit for either scale, nor did equality constraints placed on the item intercepts (scalar model) or residuals (strict model). Full measurement invariance of the GAD-7 and CESD-R-10 was demonstrated by both sex and grade. Standardized, item-by-item factor loadings are available by subgroup in Table III.

### GAD-7 and CESD-R-10 Item Responses and Sum Scores

Adjusted likelihood of an average one-point increase in GAD-7 (Model I) and CESD-R-10 (Model II) sum scores, predicted by individual characteristics, are shown in Table IV. Compared to males, females reported higher average GAD-7 ( $\beta = 3.36$ ) and CESD-R-10 ( $\beta = 2.58$ ) sum scores. GAD-7 and CESD-R-10 scores generally increased with increasing school grade; compared to participants in grade 9, the likelihood of a one-point increase in GAD-7/CESD-R-10 sum score was increasingly higher for those in grade 10 ( $\beta_{\text{GAD-7}} = 0.49$ ;  $\beta_{\text{CESD-R-10}} = 0.55$ ), grade 11 ( $\beta_{\text{GAD-7}} = 0.64$ ;  $\beta_{\text{CESD-R-10}} = 0.75$ ), and grade 12 ( $\beta_{\text{GAD-7}} = 0.91$ ;  $\beta_{\text{CESD-R-10}} = 0.99$ ).

## Discussion

Using data collected among a large sample of Canadian adolescents enrolled in the COMPASS study, we sought to assess measurement invariance of the GAD-7 and CESD-R-10 by sex and grade. Our findings indicate the presence of strict measurement invariance within the GAD-7 and CESD-R-10, confirming that these screening tools equivalently measure symptoms of generalized anxiety and major

**Table 1.** Descriptive Comparisons of GAD-7 and CESD-R-10 Scores by Individual Characteristics, Among Y<sub>7</sub> COMPASS (2018–19) Participants

Measure	<i>n</i> (%)	GAD-7 <sup>a</sup> score			CESD-R-10 <sup>b</sup> score		
		Mean (SD)	<i>F</i>	<i>p</i>	Mean (SD)	<i>F</i>	<i>p</i>
Sex							
Male	29,637 (50.2)	4.9 (5.1)	5,114.9	<.0001	7.7 (5.4)	2,899.6	<.0001
Female	29,415 (49.8)	8.2 (5.8)			10.5 (6.3)		
Grade							
9	17,018 (28.8)	6.0 (5.5)	113.4	<.0001	8.5 (6.0)	90.4	<.0001
10	17,020 (28.8)	6.5 (5.7)			9.1 (6.0)		
11	15,766 (26.7)	6.7 (5.7)			9.3 (6.1)		
12 <sup>c</sup>	9,248 (15.7)	7.3 (6.0)			9.8 (6.1)		
Ethno-racial identity							
White	38,849 (65.9)	6.4 (5.6)	57.1	<.0001	8.8 (6.0)	88.7	<.0001
Black	2,418 (4.1)	5.9 (5.8)			9.2 (5.9)		
Asian	7,152 (12.1)	6.4 (5.3)			9.6 (5.6)		
Latin American/Hispanic	1,677 (2.9)	6.5 (5.7)			9.4 (6.0)		
Other <sup>d</sup> or Mixed	8,820 (15.0)	7.4 (6.1)			10.2 (6.5)		
Weekly spending money							
\$0 or 'don't know'	18,111 (30.7)	6.5 (5.7)	3.0	.0304	9.1 (6.7)	2.0	.1118
\$1–20	13,279 (22.5)	6.6 (5.7)			9.2 (6.0)		
\$21–100	14,480 (24.5)	6.5 (5.5)			9.0 (5.9)		
\$100+	13,182 (22.3)	6.7 (5.9)			9.1 (6.1)		
Province							
Alberta	3,223 (5.5)	7.1 (6.1)	209.1	<.0001	9.6 (6.3)	119.9	<.0001
British Columbia	10,068 (17.0)	6.6 (5.7)			9.5 (5.8)		
Ontario	29,953 (50.7)	7.0 (6.0)			9.4 (6.2)		
Quebec	15,808 (26.8)	5.6 (4.9)			8.3 (5.8)		
Total sample	59,052 (100.0)	6.1 (5.6)			9.1 (6.1)		

Note. SD = standard deviation.

<sup>a</sup>Higher scores indicate greater self-reported symptoms of anxiety.

<sup>b</sup>Higher scores indicate greater self-reported symptoms of depression.

<sup>c</sup>Note there are fewer participants in this category as there is no Grade 12 in Quebec.

<sup>d</sup>Includes participants who identified as Indigenous (First Nations, Métis, Inuit) on the Cq.

depression, respectively, across adolescent subgroups. While the GAD-7 and CESD-R-10 are commonly used in pediatric settings for research and clinical screening purposes, this study is the first to explicitly test and confirm measurement invariance of both screening tools among adolescents. Our findings provide robust evidence that further supports the validity of the GAD-7 and CESD-R-10 for use in pediatric populations.

We found the psychometric properties of the GAD-7 and CESD-R-10 items to be sufficiently generalizable to adolescents by sex or grade. Statistical differences observed in GAD-7 and CESD-R-10 scores across subgroups may therefore be representative of true differences in symptoms of anxiety and depression, respectively. We identified greater risk of generalized anxiety and major depression among females than males; this is consistent with evidence that female

adolescents are disproportionately impacted by internalizing disorders than males (Van Droogenbroeck et al., 2018), and typically report higher levels of psychological distress more generally (Boak et al., 2018). While the causal factors underlying sex differences in adolescent psychopathology are not fully understood, they are likely multifactorial and mediated by a variety of factors including pubertal differences in timing of cognitive development (van Beek et al., 2012) and associated gender roles (Patel et al., 2007; Rosenfield & Mouzon, 2013).

Our results also indicated that symptoms of anxiety and depression appear to increase with secondary school grade, which is also consistent with existing estimates (Beesdo et al., 2009; Boak et al., 2018). Previous research indicates that older adolescents tend to report greater symptoms of psychopathology than their younger counterparts (Boak et al., 2018; Kilburn

**Table II.** Measurement Invariance of the GAD-7 and CESD-R-10 Scales by Sex and Grade Among Y<sub>7</sub> COMPASS (2018–19) Participants

		$\chi^2$ (df)	CFI	SRMR	RMSEA (90% CI)	$\Delta\chi^2$ (df)	$\Delta$ CFI	$\Delta$ SRMR	$\Delta$ RMSEA
GAD-7	Sex <sup>a</sup>								
	Configural	4,756.7 (28)	0.978	0.027	0.069 (0.067–0.070)	–	–	–	–
	Metric	4,097.3 (34)	0.981	0.029	0.058 (0.056–0.059)	219.4 (6)	0.003	0.002	–0.011
	Scalar	3,120.6 (47)	0.986	0.029	0.043 (0.041–0.044)	119.6 (13)	0.005	0.000	–0.015
	Strict	5,094.5 (55)	0.976	0.058	0.051 (0.049–0.052)	1,879.3 (8)	–0.010	0.029	0.008
	Grade <sup>b</sup>								
	Configural	3,524.4 (28)	0.978	0.028	0.098 (0.096–0.101)	–	–	–	–
	Metric	3,027.6 (34)	0.981	0.028	0.083 (0.080–0.085)	49.6 (6)	0.003	0.000	–0.015
	Scalar	2,391.7 (47)	0.985	0.028	0.062 (0.060–0.064)	76.7 (13)	0.004	0.000	–0.021
	Strict	1,461.9 (55)	0.991	0.037	0.045 (0.043–0.046)	284.8 (8)	0.006	0.009	–0.017
CESD-R-10	Sex <sup>a</sup>								
	Configural	2,694.2 (60)	0.979	0.029	0.035 (0.034–0.036)	–	–	–	–
	Metric	2,449.8 (69)	0.981	0.038	0.031 (0.030–0.032)	547.0 (9)	0.002	0.009	–0.004
	Scalar	2,653.7 (88)	0.980	0.041	0.029 (0.028–0.030)	643.1 (19)	–0.001	0.003	–0.002
	Strict	5,803.3 (99)	0.955	0.059	0.040 (0.039–0.041)	2,296.2 (11)	–0.025	0.018	0.011
	Grade <sup>b</sup>								
	Configural	1,693.8 (60)	0.975	0.034	0.046 (0.044–0.048)	–	–	–	–
	Metric	1,416.8 (69)	0.980	0.037	0.039 (0.037–0.041)	227.6 (9)	0.005	0.003	–0.007
	Scalar	1,464.0 (88)	0.979	0.038	0.035 (0.033–0.037)	233.5 (19)	–0.001	0.001	–0.004
	Strict	1,672.2 (99)	0.976	0.043	0.035 (0.034–0.037)	344.9 (11)	–0.003	0.005	0.000

Note. CFI = comparative fit index; CI = confidence interval; *df* = degrees of freedom; RMSEA = root mean standard error of approximation; SRMR = square root mean residual. All  $\chi^2$  values were  $p < .001$ .

<sup>a</sup>Male versus female.

<sup>b</sup>Grade 12 versus 9.

et al., 2018); the incidence of anxiety and depressive disorders is thought to increase with age (Perou et al., 2013) as symptoms become increasingly consistent with DSM-5 criteria for adults (Mullen, 2018). Given our sample of adolescents aged 12–19 years, these findings also indicate an ability to detect age-related differences in anxiety and depressive symptoms among adolescents in general. Future research should confirm whether observed statistical differences at the population level may also equate to differences by sex and age in clinical pediatric contexts.

While the GAD-7 and CESD-R-10 screening tools have been previously validated across a number of clinical and non-clinical populations, our findings fill an important gap within the existing literature where evidence among pediatric populations was limited. Although Adjorlolo (2019) found measurement invariance of the GAD-7 among Ghanaian adolescents, we are unaware of any similar investigations among Canadian adolescents. As anxiety is not experienced ubiquitously across cultural contexts (e.g., there are cultural effects of individualism versus collectivism on social anxiety [Schreier et al., 2010]), findings from other studies may not generalize adequately to Canadian or North American adolescents; establishing validity of the GAD-7 for use in our population remains important for adolescent anxiety research and clinical screening practices. This also applies to previous research that has established measurement invariance of the CESD-R-10 among adolescents in France

(Cartierre et al., 2011) and Sub-Saharan Africa (Kilburn et al., 2018); cross-cultural differences in depressive disorders have been observed among pediatric samples (Stewart et al., 2012). Haroz et al. (2014) found a version of the CESD-R-10 to be invariant by gender and sex within their U.S. adolescent sample.

### Implications

This study provides practical and clinical implications by furthering the validity of the GAD-7 and CESD-R-10 as symptom screening tools among adolescents. Pediatric health researchers and clinicians relying on the GAD-7 and CESD-R-10 should note that these screening tools exhibited strict measurement invariance in our general Canadian adolescent sample—suggesting that anxiety and depressive symptoms can be equivalently measured across gender and grade with minimal risk of measurement bias. We encourage replication of our analyses using other child and adolescent samples to further investigate the psychometric properties, including measurement invariance, of the GAD-7 and CESD-R-10 for specific use in clinical pediatric settings.

### Strengths and Limitations

Our data—representing a robust sample of adolescents across Alberta, British Columbia, Ontario, and Quebec, Canada—constitute a primary strength of our study. COMPASS is the largest ongoing longitudinal cohort study of adolescent behavioral and mental

**Table III.** Item-Level GAD-7 and CESD-R-10 Factor Loadings (Standardized) by Sex and Grade Among Y<sub>7</sub> COMPASS (2018–19) Participants

Scale		Item	Factor loading (SE)			
			Female	Male	Grade 9	Grade 12
GAD-7	1	Feeling nervous, anxious, or on edge	0.84 (0.00)	0.86 (0.00)	0.86 (0.00)	0.89 (0.00)
	2	Not being able to stop or control worrying	0.91 (0.00)	0.92 (0.00)	0.93 (0.00)	0.94 (0.00)
	3	Worrying too much about different things	0.88 (0.00)	0.89 (0.00)	0.89 (0.00)	0.92 (0.00)
	4	Trouble relaxing	0.81 (0.00)	0.84 (0.00)	0.84 (0.00)	0.87 (0.00)
	5	Being so restless that it's hard to sit still	0.70 (0.01)	0.73 (0.01)	0.73 (0.01)	0.78 (0.01)
	6	Becoming easily annoyed or irritable	0.66 (0.01)	0.69 (0.00)	0.87 (0.01)	0.73 (0.01)
	7	Feeling afraid as if something awful might happen	0.76 (0.00)	0.79 (0.00)	0.77 (0.01)	0.81 (0.01)
CESD-R-10	1	I was bothered by things that usually don't bother me	0.67 (0.01)	0.71 (0.00)	0.71 (0.01)	0.73 (0.01)
	2	I had trouble keeping my mind on what I was doing	0.69 (0.01)	0.73 (0.01)	0.74 (0.01)	0.76 (0.01)
	3	I felt depressed	0.79 (0.00)	0.82 (0.00)	0.80 (0.01)	0.72 (0.01)
	4	I felt that everything I did was an effort	0.58 (0.01)	0.62 (0.01)	0.55 (0.01)	0.57 (0.01)
	5	I felt hopeful about the future	-0.06 (0.01)	-0.07 (0.01)	-0.02 (0.01)	-0.03 (0.01)
	6	I felt fearful	0.70 (0.01)	0.74 (0.00)	0.73 (0.01)	0.75 (0.01)
	7	My sleep was restless	0.59 (0.01)	0.63 (0.00)	0.63 (0.01)	0.65 (0.01)
	8	I was happy	-0.30 (0.01)	-0.33 (0.01)	-0.26 (0.01)	-0.28 (0.01)
	9	I felt lonely	0.74 (0.00)	0.77 (0.00)	0.75 (0.01)	0.77 (0.01)
	10	I could not get "going"	0.75 (0.01)	0.78 (0.00)	0.78 (0.01)	0.79 (0.01)

**Table IV.** Mixed Linear Regression Model Predicting GAD-7 and CESD-R-10 Scores Among Y<sub>7</sub> COMPASS (2018–19) Participants

Measure	Model I—GAD-7		Model II—CESD-R-10	
	B	95% CL	B	95% CL
Sex				
Male ( <i>ref.</i> )	0.00	–	0.00	–
Female	3.36*	3.35–3.36	2.58*	2.85–2.85
Grade				
9 ( <i>ref.</i> )	0.00	–	0.00	–
10	0.49*	0.47–0.48	0.55*	0.54–0.55
11	0.64*	0.63–0.64	0.75*	0.74–0.75
12	0.91*	0.90–0.91	0.99*	0.99–1.00

Note. CL = confidence limit; *ref.* = reference category; SE = standard error. All estimates are adjusted for race/ethnicity, province, and weekly spending money. Model I: Predicts the likelihood of a one-point increase in GAD-7 score. Model II: Predicts the likelihood of a one-point increase in CESD-R-10 score.

\**p* < .001.

health in Canada. Further strengths pertain to our robust analytic approach (i.e., full-information maximum likelihood and accounting for school-level clustering). Our study is not without limitations. First, it is possible that our large sample may partially account for differences identified in the GAD-7 and CESD-R-10. Though group differences we observed are statistically significant, further research is warranted to confirm whether they are in fact clinically significant. Second, we accounted for sex and not

participants' self-identified gender, as these data were not collected. Evidence suggests non-binary and transgender youth report experiencing disproportionate levels of psychopathology compared to cis-gender youth (Veale et al., 2017), thus additional work is needed to examine measurement invariance of the GAD-7 and CESD-R-10 across gender identity groups. Additionally, recognizing there are other social, structural, and systemic factors (e.g., racialization, discrimination) which can influence adolescents' mental

health experiences, future research should consider measurement biases within these broader contexts. To this point, it is imperative that future research consider the heterogeneity across race and ethnicity in recognition that grouping participants together into “other” and/or “mixed” identity groups may mask differences and fail to inform screening and intervention efforts. Third, we note these self-reported data are not nationally representative and generalizability of our findings to all adolescents in Canada is thus limited. However, COMPASS relies on purposive sampling procedures that contribute to large sample sizes, and use of active-information passive-consent data collection protocols supports robust youth self-report research (Chartier et al., 2008). Fourth, it should be noted that the GAD-7 and CESD-R-10 were examined in specific due to their availability as existing measures in the COMPASS host study; our findings of invariance cannot be generalized to other measures of the same constructs. Additional evidence is needed to confirm measurement invariance of other tools used to screen for symptoms of anxiety and depression, as well as for other mental disorders frequently observed among youth (e.g., externalizing disorders). We recommend similar tests of invariance be conducted for other measures used in pediatric psychology. Finally, it is a potential limitation that five residuals needed to be correlated in the CESD-R-10 to observe excellent fit of the CFA model. Although, it should be noted that specifying correlations among residuals does not usually have any negative impact on model results (Cole et al., 2007).

## Conclusion

Psychometrically valid screening tools are important for informing pediatric mental health research and supporting screening efforts in general and specialty pediatric settings. This study was the first to examine measurement invariance of the GAD-7 and CESD-R-10 in a large sample of general adolescents in Canada. We confirmed that the GAD-7 and CESD-R-10 were fully invariant across sex and grade subgroups. Consistent with previous research, our sample indicated higher symptoms of anxiety and depression among female adolescents and those in older school grades.

## Funding

The COMPASS study has been supported by a bridge grant from the CIHR Institute of Nutrition, Metabolism and Diabetes (INMD) through the “Obesity—Interventions to Prevent or Treat” priority funding awards (OOP-110788; awarded to S.T.L.), an operating grant from the CIHR Institute of Population and Public Health (IPPH) (MOP-114875; awarded to

S.T.L.), a CIHR project grant (PJT-148562; awarded to S.T.L.), a CIHR bridge grant (PJT-149092; awarded to K.A.P./S.L.), a CIHR project grant (PJT-159693; awarded to K.A.P.), and by a research funding arrangement with Health Canada (#1617-HQ-000012; contract awarded to S.T.L.). The COMPASS-Quebec project additionally benefits from funding from the Ministère de la Santé et des Services sociaux of the province of Quebec and the Direction régionale de santé publique du CIUSSS de la Capitale-Nationale.

## Supplementary Data

Supplementary data can be found at: <https://academic.oup.com/jpepsy>.

*Conflict of interest:* None declared.

## References

- Adjorlolo, S. (2019). Generalised anxiety disorder in adolescents in Ghana: Examination of the psychometric properties of the Generalised Anxiety Disorder-7 scale. *African Journal of Psychological Assessment, 1*(1), a10.
- Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D. *American Journal of Preventive Medicine, 10*(2), 77–84.
- Beesdo, K., Knappe, S., & Pine, D. S. (2009). Anxiety and anxiety disorders in children and adolescents: Developmental issues and implications for DSM-V. *Psychiatric Clinics, 3*(3), 483–524.
- Benson, J., Severn, C., Hudnut-Beumler, J., Simon, S. L., Abramson, N., Shomaker, L. B., Gulley, L. D., Taylor, A., Kelsey, M. M., Nadeau, K. J., Zeitler, P. S., Pyle, L., & Cree-Green, M. (2020). Depression in girls with obesity and polycystic ovary syndrome and/or type 2 diabetes. *Canadian Journal of Diabetes, 44*(6), 507–513.
- Björgvinsson, T., Kertz, S. J., Bigda-Peyton, J. S., McCoy, K. L., & Aderka, I. M. (2013). Psychometric properties of the CES-D-10 in a psychiatric sample. *Assessment, 20*(4), 429–436.
- Boak, A., Hamilton, H. A., Adlaf, E. M., Henderson, J. L., & Mann, R. E. (2018). *The mental health and well-being of Ontario students, 1991-2017: Detailed findings from the Ontario Student Drug Use and Health Survey (OSDUHS) (CAMH Research Document Series No. 47)*. Centre for Addiction and Mental Health.
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons, Inc.
- Bradley, K. L., Bagnell, A. L., & Brannen, C. L. (2010). Factorial validity of the Center for Epidemiological Studies Depression 10 in adolescents. *Issues in Mental Health Nursing, 31*(6), 408–412.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Butler, A., Van Lieshout, R. J., Lipman, E. L., MacMillan, H. L., Gonzalez, A., Gorter, J. W., Georgiades, K., Speechley, K. N., Boyle, M. H., & Ferro, M. A. (2018).



- Mental disorder in children with physical conditions: A pilot study. *BMJ Open*, 8(1), e019011.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Taylor & Francis Group, LLC.
- Cartierre, N., Coulon, N., & Demerval, R. (2011). Confirmatory factor analysis of the short French version of the Center for Epidemiological Studies of Depression Scale (CES-D10) in adolescents. *L'Encéphale*, 37(4), 273–277.
- Chartier, M., Stoep, A. V., McCauley, E., Herting, J. R., Tracy, M., & Lymp, J. (2008). Passive versus active parental permission: Implications for the ability of school-based depression screening to reach youth at risk. *The Journal of School Health*, 78(3), 157–164.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504.
- Cole, D. A., Ciesla, J. A., & Steiger, J. H. (2007). The insidious effects of failing to include design-driven correlated residuals in latent-variable covariance structure analysis. *Psychological Methods*, 12(4), 381–398.
- Fink, E., Patalay, P., Sharpe, H., Holley, S., Deighton, J., & Wolpert, M. (2015). Mental health difficulties in early adolescence: A comparison of two cross-sectional studies in England from 2009 to 2014. *The Journal of Adolescent Health*, 56(5), 502–507.
- Haroz, E. E., Ybarra, M. L., & Eaton, W. W. (2014). Psychometric evaluation of a self-report scale to measure adolescent depression: The CESDR-10 in two national adolescent samples in the United States. *Journal of Affective Disorders*, 158, 154–160.
- Hellebrekers, D. M., Lionarons, J. M., Faber, C. G., Klinkenberg, S., Vles, J. S., & Hendriksen, J. G. (2019). Instruments for the assessment of behavioral and psychosocial functioning in Duchenne and Becker muscular dystrophy; a systematic review of the literature. *Journal of Pediatric Psychology*, 44(10), 1205–1223.
- Hinz, A., Klein, A. M., Brähler, E., Glaesmer, H., Luck, T., Riedel-Heller, S. G., Wirkner, K., & Hilbert, A. (2017). Psychometric evaluation of the Generalized Anxiety Disorder Screener GAD-7 based on a large German general population sample. *Journal of Affective Disorders*, 210, 338–344.
- Kilburn, K., Prencipe, L., Hjelm, L., Peterman, A., Handa, S., & Palermo, T. (2018). Examination of performance of the Center for Epidemiological Studies Depression Scale Short Form 10 among African youth in poor, rural households. *BMC Psychiatry*, 18(1), 201.
- Leatherdale, S. T., Brown, K. S., Carson, V., Childs, R. A., Dubin, J. A., Elliott, S. J., Faulkner, G., Hammond, D., Manske, S., Sabiston, C. M., Laxer, R. E., Bredin, C., & Thompson-Haile, A. (2014). The COMPASS study: A longitudinal hierarchical research platform for evaluating natural experiments related to changes in school-level programs, policies and built environment resources. *BMC Public Health*, 14(1), 331.
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274.
- Martinović, Ž., Simonović, P., & Djokić, R. (2006). Preventing depression in adolescents with epilepsy. *Epilepsy & Behavior*, 9(4), 619–624.
- McCormick, A. D., Schumacher, K. R., Zamberlan, M., Uzark, K., Yu, S., Lowery, R., Rottach, N., & Cousino, M. K. (2020). Generalized and specific anxiety in adolescents following heart transplant. *Pediatric Transplantation*, 24(1), e13647.
- Merikangas, K. R., He, J-p., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Survey Replication–Adolescent Supplement (NCS-A). *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(10), 980–989.
- Mossman, S. A., Luft, M. J., Schroeder, H. K., Varney, S. T., Fleck, D. E., Barzman, D. H., Gilman, R., DelBello, M., & Strawn, J. R. (2017). The generalized anxiety disorder 7-item scale in adolescents with generalized anxiety disorder: Signal detection and validation. *Annals of Clinical Psychiatry*, 29(4), 227–234A.
- Moyer, D. N., Connelly, K. J., & Holley, A. L. (2019). Using the PHQ-9 and GAD-7 to screen for acute distress in transgender youth: Findings from a pediatric endocrinology clinic. *Journal of Pediatric Endocrinology & Metabolism*, 32(1), 71–74.
- Mullen, S. (2018). Major depressive disorder in children and adolescents. *Mental Health Clinician*, 8(6), 275–283.
- Muthén L. K., & Muthén, B. O. (1998–2020). *Mplus User's Guide* (8th edn). Muthén & Muthén.
- Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *The Lancet*, 369(9569), 1302–1313.
- Patte, K. A., Bredin, C., Henderson, J., Elton-Marshall, T., Faulkner, G., Sabiston, C., Battista, K., Ferro, M., Cole, A., Doggett, A., Godin, K., Aleyan, S., Butler, A., & Leatherdale, S. T. (2017a). *Development of a mental health module for the COMPASS system: Improving youth mental health trajectories. Part 2: Pilot test and focus group results*, 4, 3. University of Waterloo. www.compass.uwaterloo.ca Retrieved 26 October 2021.
- Patte, K. A., Bredin, C., Henderson, J., Elton-Marshall, T., Faulkner, G., Sabiston, C., Battista, K., & Leatherdale, S. T. (2017b). *Development of a mental health module for the COMPASS system: Improving youth mental health trajectories. Part 1: Draft development and design*, 4, 2. University of Waterloo. www.compass.uwaterloo.ca Retrieved 26 October 2021.
- Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., Hedden, S. L., Crosby, A. E., Visser, S. N., Schieve, L. A., Parks, S. E., Hall, J. E., Brody, D., Simile, C. M., Thompson, W. W., Baio, J., Avenevoli, S., Kogan, M. D., Huang, L. N., & Centers for Disease Control. (2013). Mental health surveillance among children—United States, 2005–2011. *Morbidity and*

- Mortality Weekly Report. Surveillance Summaries* (Washington, D.C.: 2002), 62 Suppl 2(2), 1–35.
- 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.
- Rosenfield, S., & Mouzon, D. (2013). Gender and mental health. In C. S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health. Handbooks of Sociology and Social Research* (pp. 277–296). Springer.
- Ross, P. T., Hart-Johnson, T., Santen, S. A., & Zaidi, N. L. B. (2020). Considerations for using race and ethnicity as quantitative variables in medical education research. *Perspectives on Medical Education*, 9(5), 318–323.
- Schreier, S.-S., Heinrichs, N., Alden, L., Rapee, R. M., Hofmann, S. G., Chen, J., Oh, K. J., & Bögels, S. (2010). Social anxiety and social norms in individualistic and collectivistic countries. *Depression and Anxiety*, 27(12), 1128–1134.
- 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.
- Steinmetz, H., Schmidt, P., Tina-Booh, A., Wiczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality & Quantity*, 43(4), 599–616.
- Stewart, S. M., Simmons, A., & Habibpour, E. (2012). Treatment of culturally diverse children and adolescents with depression. *Journal of Child and Adolescent Psychopharmacology*, 22(1), 72–79.
- Tegethoff, M., Belardi, A., Stalujanis, E., & Meinschmidt, G. (2015). Association between mental disorders and physical diseases in adolescents from a nationally representative cohort. *Psychosomatic Medicine*, 77(3), 319–332.
- Tiirikainen, K., Haravuori, H., Ranta, K., Kaltiala-Heino, R., & Marttunen, M. (2019). Psychometric properties of the 7-item Generalized Anxiety Disorder Scale (GAD-7) in a large representative sample of Finnish adolescents. *Psychiatry Research*, 272, 30–35.
- Tompke, B. K., & Ferro, M. A. (2019). Measurement invariance and informant discrepancies of the KIDSCREEN-27 in children with mental disorder. *Applied Research in Quality of Life*, 16, 891–910.
- Tompke, B. K., Tang, J., Oltean, I. I., Buchan, M. C., Reaume, S. V., & Ferro, M. A. (2020). Measurement invariance of the WHODAS 2.0 across youth with and without physical or mental conditions. *Assessment*, 27(7), 1490–1501.
- Thompson-Haile, A., Bredin, C., & Leatherdale, S. T. (2013). Rationale for using an active-information passive-consent permission protocol in COMPASS. COMPASS Technical Report Series, 1(6). University of Waterloo. www.compass.uwaterloo.ca Retrieved 26 October 2021.
- van Beek, Y., Hessen, D. J., Hutteman, R., Verhulp, E. E., & Van Leuven, M. (2012). Age and gender differences in depression across adolescence: Real or 'bias'? *Journal of Child Psychology and Psychiatry* 53(9), 973–985.
- van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486–492.
- Van Droogenbroeck, F., Spruyt, B., & Keppens, G. (2018). Gender differences in mental health problems among adolescents and the role of social support: Results from the Belgian health interview surveys 2008 and 2013. *BMC Psychiatry*, 18(1), 6.
- Veale, J. F., Watson, R. J., Peter, T., & Saewyc, E. M. (2017). The mental health of Canadian transgender youth compared with the Canadian population. *The Journal of Adolescent Health*, 60(1), 44–49.
- Zhang, W., O'Brien, N., Forrest, J. I., Salters, K. A., Patterson, T. L., Montaner, J. S. G., Hogg, R. S., & Lima, V. D. (2012). Validating a shortened depression scale (10 item CES-D) among HIV-positive people in British Columbia. *PLoS One*, 7(7), e40793.