



Published in final edited form as:

*Psychol Assess.* 2014 December ; 26(4): 1347–1355. doi:10.1037/a0037608.

## A Longitudinal Examination of the Measurement Properties and Predictive Utility of the Center for Epidemiologic Studies Depression Scale Among North American Indigenous Adolescents

**Brian E. Armenta,**

Department of Sociology, University of Nebraska—Lincoln

**Kelley J. Sittner Hartshorn,**

Department of Sociology, University of Nebraska—Lincoln; Department of Sociology, Oklahoma State University

**Les B. Whitbeck,**

Department of Sociology, University of Nebraska—Lincoln

**Devan M. Crawford,** and

Department of Sociology, University of Nebraska—Lincoln

**Dan R. Hoyt**

Department of Sociology, University of Nebraska—Lincoln

### Abstract

We examined the longitudinal measurement properties and predictive utility of the Center for Epidemiologic Studies Depression Scale (CES-D) from early to late adolescence among a sample of North American Indigenous youths. Participants were 632 North American Indigenous adolescents ( $n = 632$ ; 50.3% girls;  $M$  age at baseline = 11.11 years) participating in an 8-year, 8-wave longitudinal study. Via in-person interviews, participants completed the CES-D at Waves 1, 3, 5, and 7, and the major depressive disorder (MDD) module of the Diagnostic Interview Schedule for Children at Waves 1, 4, 6, and 8. Confirmatory factor analyses indicated that responses to the CES-D were similarly explained by 2-, 3-, and 4-factor models, as well as a 1-factor model with correlations between the error variances for the positively worded items. Longitudinal measurement equivalence analyses indicated full structural (i.e., factor structure), metric (i.e., factor loadings), and scalar (i.e., observed item intercepts) equivalence for each factor structure. Substantive analyses showed that the CES-D was significantly associated with MDD both concurrently and prospectively, although these effects were smaller than might be expected. Finally, the CES-D negative affect and somatic complaints subscales were the strongest and most consistent predictors of MDD. Among our sample of North American Indigenous youths, the measurement properties of the CES-D were stable from early to late adolescence. Moreover, somatic difficulties and depressed affect were the strongest predictors of MDD.

## Keywords

depressive symptoms; longitudinal measurement equivalence; Native Americans; Canadian First Nations; adolescents

Depressive symptoms have been linked to negative psychosocial outcomes among Indigenous populations (i.e., Native Americans and Canadian First Nations people; e.g., LaFromboise, Medoff, Lee, & Harris, 2007; Walls, Chapple, & Johnson, 2007). For both clinical and empirical reasons, accurate assessment of depressive symptoms among this population is of the utmost importance. This is especially the case for adolescents, who experience rapid developmental (i.e., physical and psychological) and social (e.g., expanding social network) changes that may increase their susceptibility to depression (Erikson, 1968). There has been disagreement, however, regarding how to best assess depressive symptoms among Indigenous youths (Thrane, Whitbeck, Hoyt, & Shelley, 2004), and there is a lack of evidence regarding the utility of depressive symptoms measures in predicting major depressive disorders (MDDs) among this population. Furthermore, there is no evidence that measures of depressive symptoms function similarly (i.e., show measurement equivalence) among Indigenous youths as they progress through adolescence. The present study addresses these issues by examining (a) the factor structure of the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), (b) the longitudinal measurement properties (i.e., configural, metric, and scalar equivalence) of the CES-D from early to late adolescence, and (c) the utility of the CES-D for concurrently and prospectively predicting diagnoses of MDD among a sample of Indigenous youths who share the same tribal affiliation.<sup>1</sup>

## The Center for Epidemiologic Studies Depression Scale

The CES-D (Radloff, 1977) is one of the most widely used measures of depressive symptoms and has been included in a number of studies with Indigenous samples (e.g., Beals, Manson, Keane, & Dick, 1991; Chapleski, Lamphere, Kaczynski, Lichtenberg, & Dwyer, 1997; Dick, Beals, Keane, & Manson, 1994; Manson, Ackerson, Dick, Baron, & Fleming, 1990). The CES-D was developed to assess symptoms associated with major depression, specifically, “depressed mood, feelings of guilt and worthlessness, feelings of helplessness and hopelessness, psychomotor retardation, loss of appetite and sleep disturbances” (Radloff, 1977, p. 386). Correspondingly, the CES-D includes 20 statements regarding experiences indicative of major depression, such as “I felt lonely” and “I felt depressed” (see the Appendix for all 20 items). Participants are asked to indicate the frequency with which they have had each experience during the *past week*, using the response options of (0) *rarely or none of the time* (less than 1 day), (1) *some or a little of the time* (1–2 days), (2) *occasionally or a moderate amount of time* (3–4 days), and (3) *most or all of the time* (5–7 days).

<sup>1</sup>Throughout this article, we use the term *Indigenous* to refer to Native peoples of North America, including Native Americans in the United States and First Nations people in Canada. As discussed in the Method section, our specific sample includes both Native American and First Nations youths who share the same cultural heritage and identify as members of a single cultural group.

The initial validation of the CES-D was based on two randomly selected normative adult samples and a sample of adults with severe depression who were seeking psychiatric treatment (Radloff, 1977). Principle components analysis (with varimax rotation) suggested a four-factor solution, with factors representing (a) *depressed affect* (e.g., “I felt sad”), (b) *somatic complaints* (e.g., “I could not get ‘going’”), (c) *interpersonal difficulties* (e.g., “I felt that people disliked me”), and (d) *positive affect* (e.g., “I was happy”; reverse-coded to reflect a lack of positive affect). This factor structure was found to be largely consistent for Black and White Americans, individuals of varying incomes, normative and severely depressed individuals, and individuals in early, middle, and late adulthood. Despite showing a four-factor structure, Radloff (1977) argued that the 20 items could appropriately be used to create composite depressive symptoms scores, owing to the high internal consistency of the entire scale (i.e., coefficient alphas ranging from .85 to .90).

Since being introduced into the literature, the CES-D has been used in thousands of studies. In line with Radloff's (1977) suggestion, scholars most often use the CES-D items to compute single, composite depressive symptoms scores. The exact factor structure of this measure, however, has been a point of contention. Analyses conducted with various samples, including individuals from several ethnic and national backgrounds ranging in ages from early adolescence to late adulthood, suggest that a unidimensional conceptualization of depressive symptoms may be inaccurate (cf., Edwards, Cheavens, Heiy, & Cukrowicz, 2010; Wood, Taylor, & Joseph, 2010). Empirical evidence suggests, rather, that responses to the CES-D items may be better explained by a two-, three-, or four-factor solution. The four-factor structure includes the separate components originally identified by Radloff, namely, depressed affect, somatic complaints, interpersonal difficulties, and positive affect (e.g., Crockett, Randall, Shen, Russell, & Driscoll, 2005; Nguyen, Kitner-Triolo, Evans, & Zonderman, 2004). The three-factor solution includes interpersonal difficulties, positive affect, and a combined depressed affect/somatic complaints factor (e.g., Chapleski et al., 1997; Russell, Crockett, Shen, & Lee, 2008). Finally, the two-factor solution includes factors representing depressive symptoms (i.e., depressed affect, somatic complaints, and interpersonal difficulties) and positive affect (e.g., Edman et al., 1999; Miller, Markides, & Black, 1997).

## The Factor Structure of the CES-D Among North American Indigenous Individuals

Although evidence regarding the most appropriate CES-D factor structure has been mixed in general, the same is not true for Indigenous adolescents in particular. Indeed, studies indicate that among Indigenous early (Dick et al., 1994) and late (Manson et al., 1990) adolescents, the 20-item CES-D is most accurately represented by three factors, including interpersonal difficulties, positive affect, and a factor composed of the depressed affect and somatic complaints items. The authors of these studies utilized an exploratory approach in analyzing their data, specifically, principle components analysis. Although useful for the initial stages of scale development, exploratory analyses are data driven and potentially problematic inasmuch as they are based on the interrelations among items. Consequently, items that do not tap into the intended theoretical construct(s) may alter the overall factor

structure of a measure. Confirmatory factor analysis (CFA) provides a more accurate alternative to principle components analysis because it allows one to specify and evaluate theoretically based factor structures.

Using CFA, Chapleski et al. (1997) also found that the three-factor structure best explained responses to the CES-D among Indigenous individuals, although this sample included only elderly participants (see also Beals et al., 1991). Whether or not the three-factor structure generalizes to Indigenous adolescents using a confirmatory approach remains uncertain. In order to address this issue, the *first goal* of the present study was to identify the most appropriate factor structure of the CES-D among Indigenous adolescents using confirmatory analyses. Moreover, the present study relies on longitudinal data to examine whether the factor structure remains consistent from early to late adolescence (i.e., *longitudinal configural equivalence*). A lack of longitudinal configural equivalence from early to late adolescence would indicate that the structure of the CES-D, as well as the presumed underlying (i.e., latent or hypothetical) structure of depression, varies across developmental periods.

Establishing longitudinal configural equivalence for a measure represents a critical first step in determining whether a construct, as assessed by a given measure, is similar over time. Configural equivalence does not, however, provide evidence that the items of a measure similarly tap into the latent construct (or constructs), nor does it provide evidence that the individual items assess similar levels of the latent construct (or constructs) over time. That is, longitudinal configural equivalence does not speak to the equality of factor loadings (i.e., *longitudinal metric equivalence*) or observed item intercepts (i.e., *longitudinal scalar equivalence*) across time. If a measure is not found to hold longitudinal metric equivalence, attempts to aggregate the items of that measure will inaccurately represent the underlying latent construct across time and thus render statistical relations uncertain at best. If a measure is not found to hold longitudinal scalar equivalence, mean comparisons using aggregate scale scores may lead to inaccurate conclusions owing to measurement bias (Vandenberg & Lance, 2000). For the sake of brevity, we use the term *measurement equivalence* to refer to the combination of configural, metric, and scalar equivalence and refer to specific aspects of measurement equivalence where necessary.

Although studies suggest that depressive symptoms, as assessed with the CES-D, fluctuate during adolescence (e.g., Ge, Lorenz, Conger, Elder, & Simons, 1994), only one study has considered whether the CES-D holds equivalent measurement properties over time for adolescents. Specifically, using a three-wave design, Motl, Dishman, Birnbaum, and Lytle (2005) showed that the 20-item CES-D held metric equivalence among (primarily European American) early adolescents as they moved from seventh to eighth grade. No studies, however, have considered whether longitudinal metric equivalence holds across the entire adolescent period or if longitudinal scalar equivalence holds during any adolescent period. Moreover, no studies have considered the longitudinal measurement equivalence of the CES-D among Indigenous youths. In order to address these issues, the *second goal* of the present study was to examine whether the CES-D holds metric and scalar equivalence among Indigenous youths at four time points, beginning in early adolescence and ending in late adolescence.

## Predictive Utility of the CES-D

Determining that a scale holds sound measurement properties is critical for accurate assessment in both clinical and research contexts. Evidence of acceptable measurement properties, however, does not speak to the predictive utility of a measure. Inasmuch as the CES-D assesses symptoms associated with major depression, then CES-D scores should predict whether or not individuals meet criteria for an MDD. We are aware of no studies that have examined the predictive utility of the CES-D in this manner among Indigenous youths. Thus, the *third goal* of the present study was to examine the utility of the CES-D in predicting diagnoses of MDD.

## Summary and Hypotheses

To summarize, there were three goals to the present study. First, we sought to identify the most appropriate factor structure for the CES-D among a sample of Indigenous adolescents. To this end, we examined the fit of previously identified CES-D factor structures, ranging from one to four factors. Given that existing evidence strongly favors a three-factor structure for Indigenous adolescents (Dick et al., 1994; Manson et al., 1990), we expected the three-factor structure to provide the best fit to our data.

Importantly, two recent studies have revisited the possibility that responses to the items of the CES-D may be adequately accounted for by a single factor. Specifically, Edwards et al. (2010) and Wood et al. (2010) observed that the positive affect items, which ostensibly represent the reverse of depressive symptoms, may introduce a measurement artifact that can mask the plausibility of a single-factor structure. When this measurement artifact is taken into account, either by allowing the error variances for the positive affect items to correlate (Wood et al., 2010) or by including a method factor for the positive affect items (Edwards et al., 2010), a one-factor model adequately accounts for responses to the CES-D items. We thus also considered a single-factor structure in which the error variances for the positive affect items were allowed to correlate with one another.

Second, we examined the stability of the CES-D measurement properties among our sample as they progressed through adolescence. To this end, we examined the configural, metric, and scalar equivalence of the CES-D at four time points, beginning in early adolescence and ending in late adolescence (i.e., 10–18 years of age). In line with our first hypothesis, we expected the three-factor structure to hold similarly across the four waves of our study (i.e., demonstrate longitudinal configural equivalence). Moreover, we saw no strong theoretical or empirical reason to expect metric (i.e., factor loadings) or scalar (i.e., observed item intercepts) nonequivalence throughout adolescence. We thus hypothesized that the CES-D would demonstrate longitudinal metric and scalar equivalence for our sample.

Third, we considered the utility of the CES-D for predicting MDD in our sample. To this end, we examined the associations between the CES-D and MDD both concurrently and prospectively. Given that the CES-D assesses symptoms associated with major depression, we hypothesized that the CES-D would, in general, significantly predict MDD. To be comprehensive, we considered each of the previously identified CES-D subscales (based on the differential factor structures) in relation to MDD. Given the lack of consideration of the

CES-D subscales in relation to MDD, we made no strong a priori predictions regarding their differential predictive utility. There are reasons to believe, however, that some of the subscales may better predict MDD than others in our sample. First, the *Diagnostic and Statistical Manual of Mental Disorders* fourth edition (*DSM-IV*; American Psychiatric Association, 1994) and fifth edition (*DSM-5*; American Psychiatric Association, 2013) both acknowledge that somatic difficulties may be the primary presenting symptom for major depression among many (ostensibly non-Western) cultures. In addition, interpersonal and community connectedness is a central aspect of many Indigenous cultures (e.g., Bennett, 1994). Thus, the somatic complaints and interpersonal difficulties subscales may be especially strongly related to MDD diagnoses.

## Method

### Recruitment and Procedure

The data for the present article were drawn from an eight-wave longitudinal study examining culturally specific resilience and risk factors among a North American Indigenous population, which was designed in partnership with three U.S. reservations and five Canadian First Nations reserves. The reservations and reserves share a common cultural tradition and language with minor regional variations in dialects. The specific cultural group represents one of the most populous Indigenous groups in the United States and Canada. As part of confidentiality agreements, the names of the cultural group and participating sites are not provided, nor are any attempts made to make comparisons between the reservations/reserves.

At each site, tribal advisory boards were responsible for advising the research team on questionnaire development and handling personnel difficulties. The interviewers and site coordinators all were approved by advisory boards and were either enrolled tribal members or, in a very few cases, spouses of enrollees. Interviewers were trained prior to each wave concerning methodological guidelines of personal interviewing and protection of human subjects.

Prior to data collection, each participating reservation/reserve provided us a list of all families with tribally enrolled children ages 10–12 years who lived on or proximate to (within 50 miles) the reservation or reserve. An attempt to contact all families was made in order to achieve a representative sample of the communities. Families were formally recruited for the study through personal interviewer visits, during which the families were presented with a culture-specific traditional gift and an overview of the project. For those families who agreed to participate (79.4%), both the target adolescent and at least one adult caretaker were interviewed once per year over an 8-year period. Data were collected between 2002 (Wave 1) and 2009 (Wave 8; wave numbers correspond to the specific year of data collection), and no new participants were recruited beyond the initial year of data collection. For each wave of the study, participating families were given \$40 for each participant (i.e., adolescent and caregiver or caregivers) as compensation. The study was approved by and conducted in compliance with our institutional review board.

## Participants and Measures

Analyses for the present article are based on adolescent CES-D data collected during four waves of the study (i.e., Waves 1, 3, 5, and 7) and MDD data collected during four waves of the study (i.e., Waves 1, 4, 6, and 8). The initial sample included 632 youths ( $M$  age = 11.11,  $SD = .83$ ; 50.3% girls), of whom 95.4% completed Wave 3, 93.0% completed Wave 4, 93.5% completed Wave 5, 93.5% completed Wave 6, 86.2% completed Wave 7, and 82.9% completed Wave 8.

Participants completed a battery of measures at each wave, including the CES-D (Radloff, 1977) at Waves 1, 3, 5, and 7, and the Diagnostic Interview Schedule for Children–IV (DISC-IV; Shaffer, Schwab-Stone, Fisher, & Cohen, 1993) at Waves 1, 4, 6, and 8. The CES-D includes 20 items to assess four aspects of depressive symptoms, including depressed affect (seven items), somatic complaints (seven items), interpersonal difficulties (two items), and positive affect (four items; see the Appendix for all 20 items). The DISC-IV is a structured interview intended for administration by trained interviewers who may or may not have formal clinical training. We report information based on the MDD module of the DISC-IV. Standard computer algorithms were used to determine prior month and prior year MDD. Prior month MDD was used for the concurrent analysis at Wave 1, and prior year MDD was used for the prospective analyses at the remaining waves (described below).

Descriptive statistics for the various potential CES-D subscale scores (summed across the relevant items) and frequencies of MDD diagnoses across the waves are reported in Tables 1 and 2, respectively.

## Analytic Strategy and Preliminary Analyses

We conducted our longitudinal measurement analyses using Mplus Version 6.1 (Muthén & Muthén, 1998–2010). All models were estimated using full information maximum likelihood with the expectation maximization algorithm in order to account for missing data. The chi-square test of model fit becomes increasingly biased with increases in sample size, owing to the inclusion of sample size in the calculation of this statistic (Bollen, 1989). Thus, chi-square values are reported for each model but were not used for model fit evaluation. Rather, we evaluated the fit of the models using two alternative fit indices, the root-mean-square error of approximation (RMSEA) and standardized root-mean-square residual (SRMR). According to Hu and Bentler (1999), a model provides a *good fit* to the data with an RMSEA value close to or less than .06 and an SRMR value close to or less than .08. Following the suggestions of Little, Card, Slegers, and Ledford (2007; cf. Little, 1997), we used changes RMSEA (ARMSEA) values for model fit comparisons. According to these scholars, an appreciable drop in model fit is indicated by an RMSEA *increase* greater than .01.

Initial CFAs were estimated for each wave separately in order to examine the fit of the five alternative factor structures, which are shown in Figure 1. For identification purposes, the factor loading and observed item intercept for a single item for each factor were fixed to 1 and 0, respectively. Similar to previous studies (e.g., Rivera-Medina, Caraballo, Rodríguez-Cordero, Bernal, & Dávila-Marrero, 2010), the loading for the item “everything I did was an

effort” (somatic complaints) was not statistically significant in most of the models and was thus excluded from further analyses. Fit indices for the CFA models by wave using the remaining 19 items are shown in Table 3.

The one-factor model provided a good fit to the data for Wave 1 but a poor fit to the data for Waves 3, 5, and 7. For all waves, the four-factor, three-factor, two-factor, and one-factor with correlated errors models provided a good fit to the data. For all waves, the four-factor, three-factor, two-factor, and one-factor with correlated errors models provided a roughly equivalent fit to the data, as indicated by the RMSEA values. The one-factor model provided a poorer fit to the data than the remaining factor structures across all waves. For our tests of measurement equivalence, we thus excluded the one-factor model (without correlated errors) from consideration.

The tests of longitudinal measurement equivalence consisted of three steps. First, the target factor structure was estimated simultaneously for the four waves, with factor loadings and observed item intercepts allowed to estimate freely for each wave (*unconstrained model*). The error variances for like items were allowed to correlate across the waves in order to account for temporal stability in item specificities (Vandenberg & Lance, 2000). Second, to examine metric equivalence (i.e., equivalent factor loadings), a model in which the factor loadings for like items were constrained to be equal across the waves (*constrained loadings model*) was compared to the unconstrained model. Finally, to examine scalar equivalence (i.e., equivalent observed item intercepts), a model in which the intercepts for like items were constrained to be equal across the waves (*constrained intercepts model*) was compared to the constrained loadings model.

After testing for longitudinal measurement equivalence, point-biserial correlations were computed to examine the association between each possible CES-D subscale and MDD. We examined the association between the CES-D at Wave 1 and *prior month* MDD at Wave 1, the CES-D at Wave 3 and *prior year* MDD at Wave 4, the CES-D at Wave 5 and *prior year* MDD at Wave 6, and the CES-D at Wave 7 and *prior year* MDD at Wave 8.

## Results

As already reported, the two-factor, three-factor, four-factor, and one-factor with correlated error variances models provided a good fit to the data at each wave. As shown in Table 4 (*unconstrained models*), the longitudinal CFA models for each of the factor structures with factor loadings and observed item intercepts allowed to estimate freely also provided a good fit to the data. The same was the case for the constrained loadings models and the constrained intercepts models. Importantly, for each factor structure, the constrained loadings models did not result in an appreciable drop in model fit compared to the unconstrained models, and the constrained intercepts models did not result in an appreciable drop in model fit compared to the constrained loadings models, as indicated by RMSEA values that were less than .01. These results indicate that each of the factor structures considered held fully equivalent measurement properties across the waves.



## Associations With Major Depressive Disorder

The point-biserial correlations between each of the possible CES-D subscales and MDD are reported in Table 5. With four exceptions, each of the possible CES-D subscales was positively and significantly associated with MDD. The exceptions were the nonsignificant associations between interpersonal difficulties and MDD at Wave 1, positive affect at Wave 3 and MDD at Wave 4, interpersonal difficulties at Wave 7 and MDD at Wave 8, and positive affect at Wave 7 and MDD at Wave 8. At all waves, the magnitude of the associations for the subscales that included (a) all of the CES-D items, (b) the negative affect items only, (c) the somatic complaints items only, (d) the negative affect and somatic complaints items combined, and (e) the negative affect, somatic complaints, and interpersonal difficulties items combined were similar. The associations for the interpersonal difficulties items only and positive affect items only were notably weaker in magnitude.

## Discussion

We examined the factor structure and longitudinal measurement (configural, metric, and scalar) equivalence of the CES-D among a sample of North American Indigenous youths as they progressed through adolescence, and the utility of the various potential CES-D subscales (based on the factor structures; see Figure 1) in predicting MDD using seven waves of data. Drawing on previous research with North American Indigenous adolescents (Beals et al., 1991; Chapleski et al., 1997; Dick et al., 1994; Manson et al., 1990), we hypothesized that a three-factor structure, consisting of interpersonal difficulties, positive affect, and a combination of depressed affect and somatic complaints, would best explain responses to the CES-D in our sample. This hypothesis was not supported; specifically, the four-factor, three-factor, two-factor, and one-factor with correlated errors (for the positive affect items) models provided a roughly equivalent fit to the data. This was the case when the models were estimated separately for each wave and simultaneously across waves. Our hypothesis that the CES-D would hold equivalent measurement properties across the waves was supported for each of the models considered, which excluded only the one-factor model (without correlated errors).

We also considered the predictive utility of the CES-D by examining prospective (Wave 1) and longitudinal (Waves 3, 5, and 7) associations between the various CES-D subscales and past month (Wave 1) or past year (Waves 4, 6, and 8) diagnoses of MDD. We predicted that, in general, the CES-D would be significantly associated with MDD. This hypothesis was supported, as the CES-D total scores were significantly associated with MDD across all waves considered. Based on the *DSM-IV* (American Psychiatric Association, 1994) and *DSM-5* (American Psychiatric Association, 2013), which suggest that somatic difficulties may be the primary presenting symptom for major depression among many (ostensibly non-Western) cultures, we suggested that somatic difficulties may be especially strongly related to MDD. This hypothesis was partially supported, in that the association between the somatic complaints subscale and MDD was roughly equivalent to the association between the full CES-D scale, the negative affect items only, the combination of the negative affect and somatic complaints items, and the combination of the negative affect, somatic complaints, and interpersonal difficulties items.

Given the centrality of communalism within many Indigenous cultures (e.g., Bennett, 1994), we also suggested that interpersonal difficulties may be especially strongly related to MDD. This hypothesis was not supported. Specifically, when considered alone, the interpersonal difficulties subscale was either nonsignificantly associated with MDD (Waves 1 and 7) or more weakly associated with MDD than other subscales (Waves 3 and 5). Moreover, when considered alone, the positive affect subscale was either nonsignificantly associated with MDD (Waves 3 and 7) or more weakly associated with MDD than other subscales (Waves 1 and 5).

Our results allow us to draw three conclusions. First, among our sample of Indigenous adolescents, responses to the CES-D were adequately and similarly explained by the previously identified four-factor (e.g., Crockett et al., 2005; Nguyen et al., 2004), three-factor (e.g., Chapleski et al., 1997; Russell et al., 2008), and two-factor (e.g., Edman et al., 1999; Miller et al., 1997) models (see Figure 1). Moreover, once the method effect resulting from the inclusion of the four positive affect items was taken into consideration, the one-factor model also provided an adequate and similar account of the data compared to the two-, three-, and four-factor models. Given the scientific goal of parsimony, these findings indicate that the CES-D is best explained by a single underlying construct, at least among our sample. Moreover, our findings at least suggest that the previously identified multifactor structures of the CES-D may be the result of a measurement artifact, specifically, the method effect introduced by the inclusion of the four positive affect items (Edwards et al., 2010; Wood et al., 2010). Further replication, however, is warranted.

Second, for our sample, the measurement properties (i.e., factor loadings and observed item intercepts) of the CES-D held constant from early to late adolescence. This was true regardless of the factor structure considered. Our study represents the first effort to consider the measurement equivalence of the CES-D across the entire adolescent period, and our findings suggest that the items of the CES-D assess the same underlying construct (i.e., depressive symptoms) and do so in a similar way throughout the entire period of adolescence. Such evidence is necessary in order to draw valid conclusions regarding the correlates of and mean differences in depressive symptoms (as assessed by the CES-D) from early to late adolescence (Vandenberg & Lance, 2000).

Third, with the exception of the interpersonal difficulties and positive affect components of the CES-D, among our sample the CES-D total scores and the additional subscale scores functioned similarly in predicting MDD. If assessment time is of concern, our results suggest that it may be reasonable to rely on the negative affect and somatic difficulties items to assess depressive symptoms among North American Indigenous youths. In addition, given that the item “I felt everything I did was an effort” appears to be inappropriate for assessing somatic complaints in our sample, scholars may consider dropping this item. Importantly, scholars have shown that this item performs poorly in other samples (e.g., Puerto Ricans in Puerto Rico; Rivera-Medina et al., 2010). At minimum, we urge scholars who are using the CES-D to consider the usefulness of this item and clinicians who use the CES-D for screening purposes to be aware of the potential inadequacy of this item.

One additional note regarding the use of the CES-D among Indigenous populations warrants discussion. Specifically, whether depressive symptoms as conceptualized with the CES-D are adequately indicative of problematic symptomatology among North American Indigenous populations requires some consideration. To be certain, the CES-D was developed as a screening tool, rather than a diagnostic instrument, and relies on a different metric (i.e., continuous) than MDD (i.e., dichotomous). Nonetheless, the magnitude of the associations between the CES-D and MDD was small. This raises questions regarding the usefulness of the CES-D for this population. Indeed, some scholars have suggested that the dominant, heavily Western conceptualization of depression may not accurately reflect North American Indigenous experiences that are associated with maladaptive functioning. For example, based on her ethnographic work with the Flathead people of Wisconsin, O'Neill (2004) suggested that experiences of loneliness characterized by the sense of cultural degradation and historical discontinuity may be a better gauge of pathology than the Western notion of depression. It thus seems at least plausible that the underlying roots of pathology may be distinct from the traditional Western notions of depression (Manson, 1995) and perhaps across North American Indigenous populations who share some degree of similar beliefs and practices (relative to Western beliefs and practices) and who have been similarly subjected to Western colonization (cf. Allen, 1998).

The culturally grounded notion of loneliness may prove to be equally important as or more important than the dominant Western notions of depressive symptoms for understanding the sources of maladaptive functioning among North American Indigenous populations. Although the interpersonal difficulties items of the CES-D hint at loneliness, it may be more useful to focus on broader relational difficulties, such as feeling disconnected or rejected from one's family, community, and/or culture. Scholars are strongly urged to consider these and other culture-specific conceptualizations of problematic symptomatology as they further attempt to understand the antecedents to psychosocial functioning among North American Indigenous individuals.

In summary, the results of the study reported in this article provide evidence that the measurement properties of the CES-D remain stable throughout adolescence and that the CES-D is useful in predicting MDD, although perhaps not as strongly as could be expected. With the exception of the interpersonal difficulties and positive affect items considered as separate subscales, the remaining combination of CES-D items (i.e., negative affect only, somatic complaints only, and the combinations between negative affect, somatic complaints, interpersonal difficulties, and positive affect) functioned similarly in predicting MDD. This may not be the case for other important indicators of adaptive (e.g., positive school adjustment) and maladaptive (e.g., suicidal ideation) functioning, and scholars are urged to consider the potential differential effects of the CES-D subscales when focusing on outcomes other than MDD. Finally, we have highlighted some potential cultural issues with assessing depressive symptoms among members of Indigenous groups using the CES-D, which we hope will stimulate additional investigations into the nature of problematic symptomatology among this population.

## Acknowledgments

The research reported in this article was supported by National Institute on Drug Abuse Grant DA13580, National Institute of Mental Health Grant MH67281, and National Institute on Alcohol Abuse and Alcoholism Grant AA020299 awarded to Les B. Whitbeck.

## Appendix

### Items From the Center for Epidemiologic Studies Depression Scale

#### Negative Affect

1. I felt depressed.
2. I felt sad.
3. I had crying spells.
4. I thought my life had been a failure.
5. I felt fearful.
6. I felt that I could not shake off the blues even with help from my family or friends.
7. I felt lonely.

#### Somatic Complaints

1. I could not get “going.”
2. My sleep was restless.
3. I was bothered by things that usually don't bother me.
4. I did not feel like eating; my appetite was poor.
5. I had trouble keeping my mind on what I was.
6. I talked less than usual.
7. I felt that everything I did was an effort.

#### Interpersonal Difficulties

1. People were unfriendly.
2. I felt that people disliked me.

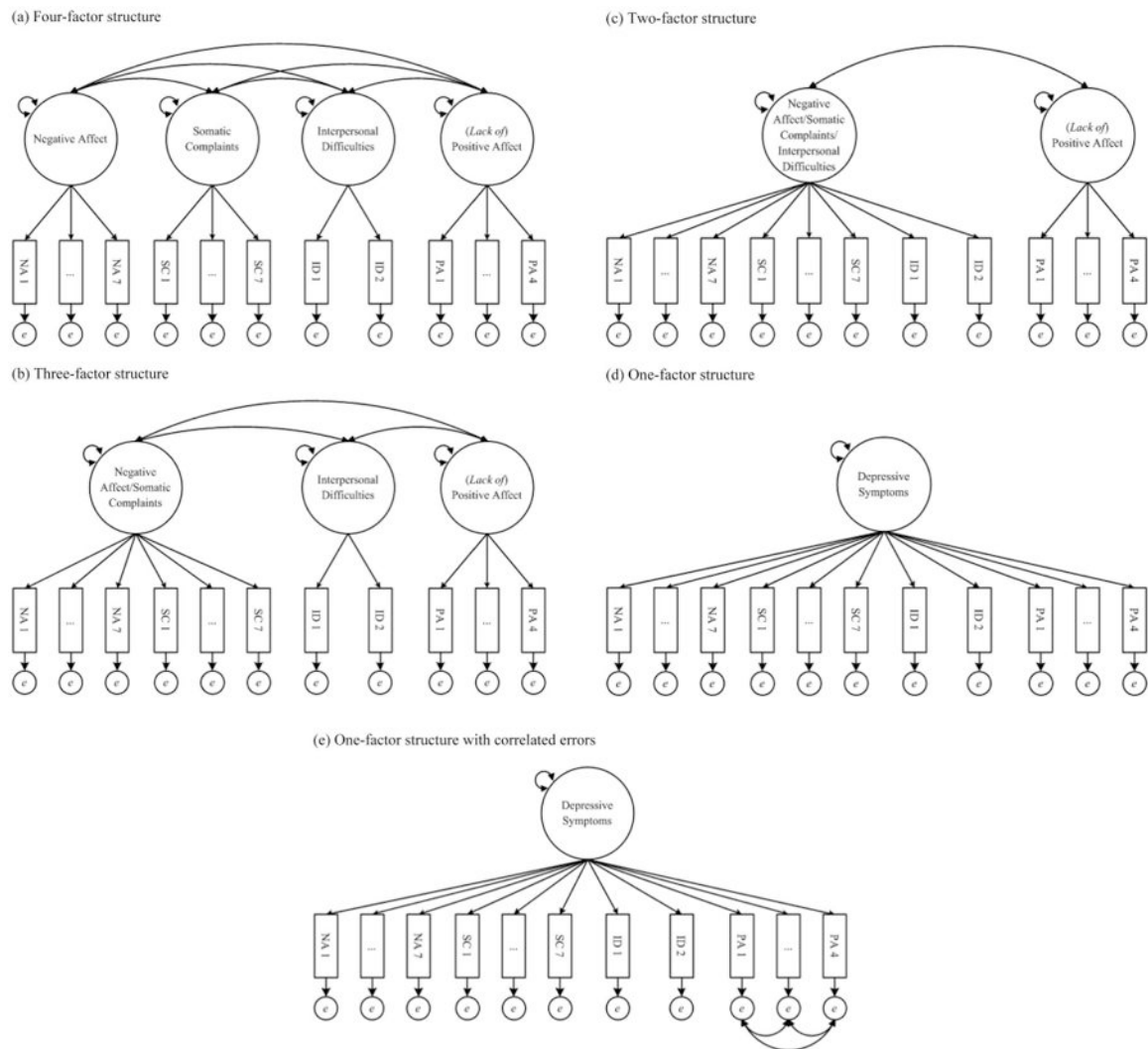
#### Positive Affect

1. I was happy.
2. I enjoyed life.
3. I felt hopeful about the future.
4. I felt I was just as good as other people.

## References

- Allen J. Personality assessment with American Indians and Alaska Natives: Instrument considerations and service delivery style. *Journal of Personality Assessment*. 1998; 70:17–42.10.1207/s15327752jpa7001\_2
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4th. Washington, DC: Author; 1994.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th. Washington, DC: Author; 2013.
- Beals J, Manson SM, Keane EM, Dick RW. Factorial structure of the Center for Epidemiologic Studies–Depression Scale among American Indian college students. *Psychological Assessment*. 1991; 3:623–627.10.1037/1040-3590.3.4.623
- Bennett, S. The American Indian: A psychological overview. In: Lonner, WJ.; Malpass, R., editors. *Psychology and culture*. Boston, MA: Allyn & Bacon; 1994. p. 35-39.
- Bollen, KA. *Structural equations with latent variables*. New York, NY: Wiley; 1989.
- Chapleski EE, Lamphere JK, Kaczynski R, Lichtenberg PA, Dwyer JW. Structure of a depression measure among American Indian elders: Confirmatory factor analysis of the CES-D Scale. *Research on Aging*. 1997; 19:462–485.10.1177/0164027597194004
- Crockett LJ, Randall BA, Shen Y, Russell ST, Driscoll AK. Measurement equivalence of the Center for Epidemiological Studies Depression Scale for Latino and Anglo adolescents: A national study. *Journal of Consulting and Clinical Psychology*. 2005; 73:47–58.10.1037/0022-006X.73.1.47 [PubMed: 15709831]
- Dick RW, Beals J, Keane EM, Manson SM. Factorial structure of the CES-D among American Indian adolescents. *Journal of Adolescence*. 1994; 17:73–79.10.1006/jado.1994.1007
- Edman JL, Danko GP, Andrade N, McArde JJ, Foster J, Glipa J. Factor structure of the CES-D (Center for Epidemiologic Studies Depression Scale) among Filipino-American adolescents. *Social Psychiatry and Psychiatric Epidemiology*. 1999; 34:211–215.10.1007/s001270050135 [PubMed: 10365627]
- Edwards MC, Cheavens JS, Heiy JE, Cukrowicz KC. A reexamination of the factor structure of the Center for Epidemiologic Studies Depression Scale: Is a one-factor model plausible? *Psychological Assessment*. 2010; 22:711–715.10.1037/a0019917 [PubMed: 20822284]
- Erikson, EH. *Identity: Youth and crisis*. New York, NY: Norton; 1968.
- Ge X, Lorenz FO, Conger RD, Elder GH, Simons RL. Trajectories of stressful life events and depressive symptoms during adolescence. *Developmental Psychology*. 1994; 30:467–483.10.1037/0012-1649.30.4.467
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6:1–55.10.1080/10705519909540118
- LaFromboise TD, Medoff L, Lee CC, Harris A. Psychosocial and cultural correlates of suicidal ideation among American Indian early adolescents on a Northern Plains reservation. *Research in Human Development*. 2007; 4:119–143.10.1080/15427600701481020
- Little TD. Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*. 1997; 32:53–76.10.1207/s15327906mbr3201\_3
- Little, TD.; Card, NA.; Slegers, DW.; Ledford, EC. Representing contextual effects in multiple-group MACS models. In: Little, TD.; Bovaird, JA.; Card, NA., editors. *Modeling contextual effects in longitudinal studies*. Mahwah, NJ: Erlbaum; 2007. p. 121-147.
- Manson SM. Culture and major depression: Current challenges in the diagnosis of mood disorders. *Psychiatric Clinics of North America*. 1995; 18:487–501. [PubMed: 8545263]
- Manson SM, Ackerson LM, Dick RW, Baron AE, Fleming CM. Depressive symptoms among American Indian adolescents: Psychometric characteristics of the Center for Epidemiologic Studies Depression Scale (CES-D). *Psychological Assessment*. 1990; 2:231–237.10.1037/1040-3590.2.3.231

- Miller TQ, Markides KS, Black SA. The factor structure of the CES-D in two surveys of elderly Mexican Americans. *Journals of Gerontology: Series B Psychological Sciences and Social Sciences*. 1997; 52:S259–S269.10.1093/geronb/52B.5.S259
- Motl RW, Dishman RK, Birnbaum AS, Lytle LA. Longitudinal invariance of the Center for Epidemiological Studies-Depression Scale among girls and boys in middle school. *Educational and Psychological Measurement*. 2005; 65:90–108.10.1177/0013164404266256
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén; 1998–2010.
- Nguyen HT, Kitner-Triolo M, Evans MK, Zonderman AB. Factorial invariance of the CES-D in low socioeconomic status African Americans compared with a nationally representative sample. *Psychiatry Research*. 2004; 126:177–187.10.1016/j.psychres.2004.02.004 [PubMed: 15123397]
- O'Neill TD. Culture and pathology: Flathead loneliness revisited. *Culture, Medicine and Psychiatry*. 2004; 28:221–230.10.1023/B:MEDI.0000034412.24293.bf
- Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977; 1:385–401.10.1177/014662167700100306
- Rivera-Medina CL, Caraballo JN, Rodríguez-Cordero ER, Bernal G, Dávila-Marrero E. Factor structure of the CES-D and measurement invariance across gender for low-income Puerto Ricans in a probability sample. *Journal of Consulting and Clinical Psychology*. 2010; 78:398–408.10.1037/a0019054 [PubMed: 20515214]
- Russell ST, Crockett LJ, Shen Y, Lee S. Cross-ethnic invariance of self-esteem and depression measures for Chinese, Filipino, and European American adolescents. *Journal of Youth and Adolescence*. 2008; 37:50–61.10.1007/s10964-007-9231-1
- Shaffer D, Schwab-Stone M, Fisher PW, Cohen P. The Diagnostic Interview Schedule for Children—Revised Version (DISC-R): I. Preparation, field testing, interrater reliability, and acceptability. *Journal of the American Academy of Child & Adolescent Psychiatry*. 1993; 32:643–650.10.1097/00004583-199305000-00023 [PubMed: 8496128]
- Thrane LE, Whitbeck LB, Hoyt DR, Shelley MC. Comparing three measures of depressive symptoms among American Indian adolescents. *American Indian and Alaska Native Mental Health Research*. 2004; 11:20–42.10.5820/aian.1103.2004.20 [PubMed: 15536587]
- Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*. 2000; 3:4–70.10.1177/109442810031002
- Walls ML, Chapple CL, Johnson KD. Strain, emotion, and suicide among American Indian youth. *Deviant Behavior*. 2007; 28:219–246.10.1080/01639620701233100
- Wood AM, Taylor PJ, Joseph S. Does the CES-D measure a continuum from depression to happiness? Comparing substantive and artifactual models. *Psychiatry Research*. 2010; 177:120–123.10.1016/j.psychres.2010.02.003 [PubMed: 20207424]



**Figure 1.** Alternative Center for Epidemiologic Studies Depression Scale factor structures. a: Four-factor structure. b: Three-factor structure. c: Two-factor structure. d: One-factor structure. e: One-factor structure with correlated errors. NA = negative affect; SC = somatic complaints; ID = interpersonal difficulties; PA = (lack of) positive affect;  $e$  = error.

**Table 1**  
**Descriptive Statistics for the Center for Epidemiologic Studies Depression Scale (CES-D) by Wave**

CES-D scale/subscale	All waves:scale range	Wave 1 (n = 673)			Wave 3 (n = 627)			Wave 5 (n = 605)			Wave 7 (n = 569)		
		M	SD	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$
Full scale	0-57	11.17	8.17	.85	12.10	8.55	.87	12.59	8.68	.88	11.14	7.89	.87
NA	0-21	3.25	3.49	.66	3.28	3.74	.70	3.29	3.80	.71	2.89	3.45	.72
SC	0-18	3.55	3.14	.56	3.81	3.31	.55	4.44	3.54	.57	4.07	3.27	.57
ID	0-6	1.44	1.40	.70	1.43	1.23	.77	1.65	1.32	.79	1.49	1.27	.78
PA	0-12	2.96	2.71	.79	3.61	2.80	.84	3.23	2.68	.86	2.70	2.52	.85
NA + SC	0-39	6.80	6.10	.85	7.09	6.55	.89	7.74	6.84	.90	6.96	6.24	.89
NA + SC + ID	0-45	8.22	6.94	.86	8.50	7.24	.89	9.36	7.59	.90	8.44	6.97	.89

Note. NA = negative affect; SC = somatic complaints; ID = interpersonal difficulties; PA = positive affect.



**Table 2**  
**Frequencies For Past Month (Wave 1) And Past Year (Waves 4, 6, And 8) Major Depressive Disorder**

Wave	<i>n</i>	% of sample
Wave 1–past month ( <i>n</i> = 673)	8	1.2
Wave 4–past year ( <i>n</i> = 585)	48	8.2
Wave 6–past year ( <i>n</i> = 590)	27	4.6
Wave 8–past year ( <i>n</i> = 523)	8	1.5

**Table 3**  
**Model Fit for Within-Wave Confirmatory Factor Models**

Wave/model	$\chi^2$	<i>df</i>	RMSEA	SRMR
Wave 1				
Four factors	296.17	146	.039	.041
Three factors	304.28	149	.039	.041
Two factors	325.32	151	.041	.043
One factor	522.23	152	.060	.062
One factor with correlated errors	316.41	146	.042	.041
Wave 3				
Four factors	346.67	146	.047	.050
Three factors	394.89	149	.051	.052
Two factors	425.19	151	.054	.054
One factor	639.41	152	.072	.069
One factor with correlated errors	371.94	146	.050	.045
Wave 5				
Four factors	332.21	146	.046	.050
Three factors	378.09	149	.050	.051
Two factors	420.01	151	.054	.055
One factor	664.06	152	.075	.071
One factor with correlated errors	344.82	146	.047	.042
Wave 7				
Four factors	384.51	146	.054	.058
Three factors	456.75	149	.060	.060
Two factors	492.31	151	.063	.062
One factor	760.36	152	.084	.079
One factor with correlated errors	394.91	146	.055	.046

*Note.* RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

**Table 4**  
**Model Fit for Longitudinal Measurement Equivalence Analyses**

Model	$\chi^2$	<i>df</i>	RMSEA	SRMR
Four factors				
Unconstrained	3,630.23	2,540	.025	.042
Constrained loadings	3,738.67	2,858	.026	.045
Constrained intercepts	3,981.76	2,630	.028	.046
Three factors				
Unconstrained	3,851.54	2,594	.027	.043
Constrained loadings	3,960.81	2,642	.027	.046
Constrained intercepts	4,240.26	2,690	.029	.048
Two factors				
Unconstrained	4,031.72	2,632	.028	.045
Constrained loadings	4,178.29	2,684	.029	.049
Constrained intercepts	4,465.16	2,735	.031	.050
One factor with correlated errors				
Unconstrained	3,983.70	2,630	.028	.048
Constrained loadings	4,114.56	2,684	.028	.051
Constrained intercepts	4,442.66	2,738	.030	.053

*Note.* RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

**Table 5**  
**Point-Biserial Correlations**

CES-D scale/subscales	Wave 1 CES-D → Wave 1 MDD ( <i>n</i> = 670)	Wave 3 CES-D → Wave 4 MDD ( <i>n</i> = 569)	Wave 5 CES-D → Wave 6 MDD ( <i>n</i> = 566)	Wave 7 CES-D → Wave 8 MDD ( <i>n</i> = 486)
Total scale	.13***	.18***	.22***	.13**
Negative affect	.11**	.18***	.19***	.13**
Somatic complaints	.13***	.19***	.21***	.17***
Interpersonal difficulties	.04	.11*	.13**	.09
Positive affect (reverse-scored)	.08*	.05	.11**	-.03
Negative affect/somatic complaints	.13***	.20***	.22***	.16***
Negative affect/somatic complaints/interpersonal difficulties	.12**	.20***	.22***	.16***

*Note.* CES-D = Center for Epidemiologic Studies Depression Scale; MDD = major depressive disorder.

\*  
 $p < .05$ .

\*\*  
 $p < .01$ .

\*\*\*  
 $p < .001$ .