# Gender Differences in the Assessment of Depression in American Indian Older Adults: The Strong Heart Study

**Celestina Barbosa-Leiker, PhD**[a], **Ekaterina Burduli, PhD**[a], **Randi Arias-Losado, MS**[a], **Clemma Muller, PhD**[b,c], **Carolyn Noonan, MS**[b,c], **Astrid Suchy-Dicey, PhD**[b,c], **Lonnie Nelson, PhD**[a,c], **Steven P. Verney, PhD**[d], **Thomas J. Montine, MD, PhD**[e], **Dedra Buchwald, MD**[b,c]

[a]Washington State University, College of Nursing

[b]Washington State University, Elson S. Floyd College of Medicine

[c]Washington State University, Institute for Research and Education to Advance Community Health

[d]University of New Mexico, Department of Psychology

[e]Stanford University, Department of Pathology

## Abstract

The validation of the assessment of depression across ethnic groups is critical yet deficient for American Indian (AI) adults. Therefore, we assessed the psychometric properties of the Center for Epidemiological Studies-Depression (CES-D) in AI elders and tested differences in depression constructs between gender. Participants were 817 AI adults (68% women), mean age 73.2 years (SD = 6.1, range 64 to 95) for women and 72.6 years (SD = 5.3, range 65 to 90) for men., in the Cerebrovascular Disease and Its Consequences in American Indians Study. We evaluated the factor structure of the 20-item and 12-item CES-D and tested measurement invariance between gender. Results demonstrated poor fit for the 20-item CES-D and partial gender measurement invariance of the 12-item CES-D. AI female elders had significantly higher depression levels than AI male elders on the depressed affect subscale, the somatic symptoms subscale, and the wellbeing (reverse-coded) subscale. Further replication is needed, and we recommend future psychometric work with the 12-item CES-D with AI elders.

### Keywords

depression; American Indian; aging

## Introduction

Depression is related to cardiovascular disease (CVD) (Musselman, Evans, & Nemeroff, 1998; Nemeroff, Musselman, & Evans, 1998), stroke (Hackett, Yapa, Parag, & Anderson, 2005), and Alzheimer's Disease and Related Dementia (ADRD) (Diniz, Butters, Albert,

**Corresponding Author:** Celestina Barbosa-Leiker, PhD, 412 E. Spokane Falls Blvd., Spokane, WA 99202-2131, PH 509-324-7477, celestina@wsu.edu.

Dew, & Reynolds, 2013). American Indian (AI) adults have higher prevalence of CVD (Centers for Disease Control and Prevention, 2014), stroke (Zhang et al., 2008), and ADRD-related comorbidities (Zhang et al., 2008) compared to non-Hispanic White counterparts. AI adults also have disproportionately high rates of mental health issues compared to the general US population (Heart, Chase, Elkins, & Altschul, 2011). Contributing factors such as oppression, historical trauma, and adverse events have been linked to mental health issues experienced by AI adults (American Psychological Association Division of Diversity and Health Equity, 2017). However, research with AI adults in tribes within the Northern Plains and the Southwest found lower prevalence of lifetime major depressive disorder compared to the general US population (Beals et al., 2005). In the general US population, women had a higher prevalence of depressive disorder compared to men (Kuehner, 2003); in contrast, no gender differences in prevalence of depressive disorders have been identified within AI tribes in the Northern Plains or the Southwest (Beals et al., 2005). Identifying differences in depressive symptomology in elder AI men and women may assist with the assessment of depressive disorders in this population.

The Center for Epidemiological Studies-Depression (CES-D) assesses occurrence of depressive symptoms in the general adult population (Radloff, 1977) and is commonly used to measure depressive symptoms in various community and clinical samples. Numerous psychometric models of the CES-D have been tested between adult samples, most positing 4-factor models (including a single higher-order model) consisting of depressed affect, well-being, somatic symptoms, and inter-personal relations. For example, measurement invariance of the 4-factor model was found between older men and women, with no differences in the latent means across gender (Gomez & McLaren, 2015). Using the Item Response Theory framework, researchers have also identified sources of item-response bias in the CES-D. Cole et al. found that two interpersonal items were higher in Black vs. White respondents when matched on overall depressive symptoms, and the "crying spells" item was higher in women vs. men when matched on overall depressive symptoms (Cole, Kawachi, Maller, & Berkman, 2000).

Few studies have established the factor structure of the CES-D in AI samples, and even less in AI elders, or older adults. One study examining the factor structure of the CES-D in older AI adults found that the somatic factor was not identified and the original wellbeing factor was separated into two factors (Barón, 1990). In a study of AI adults from the Northwest Coast (Somervell et al., 1992, 1993), the factor structure differed from the original (Radloff, 1977) factor structure, with no clear division between the depressed affect and somatic factors.

An abbreviated 12-item version of the CES-D was proposed in a study of older AI adults from the Great Lake region and was tested against the 20-item CES-D (Chapleski, Lamphere, Kaczynski, Lichtenberg, & Dwyer, 1997). Results indicated that a 3-factor model of the 12-item scale was the best fitting model, where the factors represented depressive affect, well-being, and somatization. Measurement invariance of the hypothesized CES-D structures between residential areas was also tested, and results indicated that only the12-item version had a statistically similar factor structure and factor loadings between

residential strata (urban vs. rural vs. reservation). Authors concluded that the shortened, 12-item scale, comprised items more theoretically relevant for the AI samples.

More recently, the 20-item and 12-item CES-D scales were tested in a sample of older AI adults in the Southeast region of the U.S. (Schure & Goins, 2017). While their findings determined high internal reliability and concurrent validity of the full 20-item and the 12-item CES-D scale scores, the 12-item scale resulted in a 1-factor model, as opposed to a 3-factor model (Chapleski et al., 1997). While the CES-D was conceptualized as a 20-item scale (Radloff, 1977) and psychometric properties of the full scale have been tested in AI populations (Barón, 1990; Somervell et al., 1992, 1993), Chapleski et al. (1997) and Schure and Goins (2017) have tested psychometric properties of the 12-item CES-D in AI populations with contradictory findings across studies. Thus, previous research warrants tests of psychometric properties of the 20-item and 12-item CES-D in American Indian older adults. Therefore, while foundational psychometric work has been done on the CES-D with AI elders, more work is needed as no studies have been able to replicate previous factor structures within AI samples. In addition, the CES-D factor models have never been tested for measurement invariance between AI men and women which is needed to compare depression between AI men and women. We therefore tested previous factor structures of the CES-D in AI elders and conducted tests of measurement invariance between men and women. Finally, we tested latent means of the CES-D factor(s) between gender to determine gender differences in depressive symptomology in AI elders.

## Method

### Sampling.

The Cerebrovascular Disease and Its Consequences in American Indians (CDCAI) Study (A. M. Suchy-Dicey et al., 2016) is the largest cohort of aging AIs from 13 tribal communities across three major geographic regions, including the Northern Plains (North and South Dakota), Southern Plains (Oklahoma), and Southwest (Arizona), with examinations of detailed cognitive function and neuroimaging. More than 817 survivors (N=554 women; N=263 men) of the parent study (Strong Heart Study) were recruited, enrolled, and participated in neurological and neuropsychological evaluations in 2010–2013 (A. M. Suchy-Dicey et al., 2016). Mean age was 73.2 years ($SD$ = 6.1, range 64 to 95) for women and 72.6 years ($SD$ = 5.3, range 65 to 90) for men. All visits involved a physical clinical examination, and an extensive questionnaire, including demographic information, medical and traumatic history, and clinical comorbidities (diabetes, hypertension, dyslipidemia, kidney disease, vascular disease). All participants provided written informed consent. This secondary data analysis includes data from the 2010–2013 data collection phase because all participants had the opportunity to complete the CES-D during this study phase. A detailed description of the design and recruitment methods for the CDCAI Study are described elsewhere (A. M. Suchy-Dicey et al., 2016). This study received approvals by five Tribal Review Boards or Tribal Councils, five academic or medical institutional review boards (IRBs), and three regional Indian Health Service IRBs. The primary institution IRB approval was through the University of Washington; the present study also received IRB approval from Washington State University.

### Variables.

Depression was assessed via the CES-D. This self-report scale consists of 20 items and was then reduced to a 12-item scale after data collection (Chapleski et al., 1997; Schure & Goins, 2017). The items are rated based on frequency of occurrence during the past week using a 4-point Likert-type scale. The anchors are 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). The scoring of the items on the Well-Being construct were reversed scored for all analyses (descriptive statistics and factor analyses). Note that the original wording for a CES-D item was "I could not get going"; this item was amended in the CDCAI to "I felt like I couldn't do what I needed to do", in order to accommodate cultural and linguistic confusion with the idiomatic phrase "get going."

### Data Analysis.

Descriptive statistics included means and standard deviations, and skewness and kurtosis statistics were used to assess assumptions of normality. Frequency statistics were calculated for all CES-D items for the total sample and separately for men and women. Full-information maximum likelihood estimation was used for missing data. Confirmatory factor analyses (CFA) first tested the hypothesized factor models of the CES-D for the total sample and for men and women. Based on previous 20-item CES-D research, we tested a 1-factor higher-order model with 4 first-order factors, a 4-factor model, and a 3-factor model. The 4-factor model consisted of seven items loading on Depressed Affect, four items loading on Well-Being, seven items loading on Somatic Symptoms, and two items loading on Inter-personal factor. The 3-factor model consisted of 14 items loading on Depressed Affect, four items loading on Well-Being, and two items loading on Inter-personal factor (the seven items that generally load on Somatic Symptoms in the 4-factor model are added to Depressed Affect in the 3-factor model). In addition, using the 12-item scale, we tested 3-factor and 1-factor models. The 3-factor, 12-item model consisted of five Depressed Affect items, four Somatic Symptoms items, and three Well-Being items. Overall model fit was evaluated using the robust comparative fit index (CFI; study criterion 0.900), the robust root mean square error of approximation (RMSEA; study criterion 0.080), and the robust standardized root mean square residual (SRMR; study criterion 0.080), and pattern and statistical significance of factor loadings were also evaluated (Brown, 2006). We also assessed factor correlations and internal consistency via SEM-based coefficient alphas in the best-fitting model.

Next, a series of multiple-group analyses was used to assess measurement invariance between gender. Configural invariance examined if there is an equal factor structure between the groups; metric invariance tested if the relationships of CES-D items were equivalent for like items between gender (i.e., constrained factor loadings); scalar invariance constrained the intercepts of like items to be equal between groups to test whether differences in means of the items were due to differences in the construct. This last step evaluated whether observed values of the CES-D items were equivalent between gender when there was a constant level of depression. After the demonstration of scalar invariance between gender was shown, mean differences between gender can then be attributed to true differences in the construct. While $\chi^2$ difference tests are often used to compare nested models (e.g.,

measurement invariance models), these can be influenced by sample size (Chen, Sousa, & West, 2005). Therefore, for tests of measurement invariance, a change in CFI between comparison and nested models of −0.010 in addition to a change in RMSEA of 0.015 or a change in SRMR of 0.030 (for loading invariance) and 0.010 (for intercept invariance) was used (Chen, 2007). Once measurement invariance was demonstrated, latent means were tested between men and women. Mplus version 8 was used for all analyses (Muthén & Muthén, 1998–2018).

## Results

### Baseline factor models.

Skewness and kurtosis statistics indicated normal distributions between items with the exception of the "dislike" item which was slightly skewed (skewness statistic = 3.27) and peaked (kurtosis statistic = 11.70). All response categories were endorsed. As can be seen in Table 1, factor models using the full 20-item CES-D (the higher-order 1-factor model with 3 lower-order factors, the 4-factor model, and the 3-factor model) did not fit the total sample. However, the 3-factor model of the 12-item CES-D demonstrated good fit for the total sample ($\chi^2(51) = 187.88$, CFI = .920, SRMR = .047, RMSEA = .057). This model, with the addition of a correlated error between the "sad" with "crying" items, resulted in good fit for both women ($\chi^2(50) = 94.86$, CFI = .964, SRMR = .042, RMSEA = .040) and men ($\chi^2(50) = 77.36$, CFI = .942, SRMR = .055, RMSEA = .046). Factor loadings for women and men were statistically significant indicating each item explained a significant amount of variance in the latent construct (Table 2). Factor correlations were as follows: Depressed Affect with Somatic Symptoms $r = .88$, $p < .001$, Depressed Affect with Well-Being $r = .27$, $p < .001$, Somatic Symptoms with Well-Being, $r = .14$, $p = .017$. We then proceeded to use only the 12-item 3-factor model of the CES-D for invariance testing between women and men.

### Measurement Invariance Between Groups.

The fit for tests of measurement invariance between gender are presented in Table 3. Results indicated that the configural invariance had good fit, however metric invariance did not demonstrate good fit and modification indices indicated that the "crying" item constraint should be released between gender. Once that constraint was released, the metric invariance model showed adequate fit between gender and there was not a decrement in fit when going from the configural invariance model to metric invariance model. Lastly, the scalar invariance model showed adequate fit between gender and there was not a decrement in fit when going from the metric invariance models to scalar invariance model. SEM-based coefficient alphas for the total sample were as follows: total scale = .72, Depressed Affect = 0.78, Somatic Symptoms = .66, and Well-Being = .71. A similar pattern of results was found for women: total scale = .74, Depressed Affect = 0.81, Somatic Symptoms = .66, and Well-Being = .71. SEM-based coefficient alphas for men were lower: total scale = .65, Depressed Affect = 0.69, Somatic Symptoms = .64, and Well-Being = .69.

### Latent means analysis.

With measurement invariance of the 3-factor CES-D model demonstrated between gender, latent means analysis was conducted to determine differences in depression between groups.

Results indicated that AI female elders had significantly higher depression levels than AI male elders on the depressed affect subscale mean score ($z = 0.246$, $p = .002$), the somatic symptoms subscale mean score ($z = 0.297$, $p = .001$), and the wellbeing (reverse-coded) subscale mean score ($z = 0.163$, $p = .05$).

## Discussion

Higher-order 1-factor, 4-factor, and 3-factor models were tested with the 20-item CES-D and resulted in poor fit in AI elders. In examining the abbreviated 12-item factor model posited by researchers studying AI elders (Chapleski et al., 1997; Schure & Goins, 2017), we found a good-fit for the 3-factor model and not the 1-factor model, in line with one psychometric study of AI elders (Chapleski et al., 1997) and not another (Schure & Goins, 2017), demonstrating the need to acknowledge differences across tribal communities and to conduct research with AI communities across the nation. Similar to Chapleski et al. (1997) and Schure and Goins (2017), internal consistency in the present study indicated a reliable total score for the total sample. However, the Somatic Symptoms scale was < .70 (.64 to .66) across the samples, and reliability across the total score and subscales was questionable for men (.64 to .69), indicating the need for future psychometric assessment. The pattern of factor correlations indicated a strong relationship between Depressed Affect with Somatic Symptoms, with weaker but significant correlations between Depressed Affect with Well-Being and Somatic Symptoms with Well-Being. This pattern of factor correlations was also found by Chaplenski et al. (1997).

When testing measurement invariance between gender, we needed to add in a correlated error between the items about crying and being sad, indicating that there is a relationship between those two items that may not be related to the construct of depression. Tests of measurement invariance indicated that the factor models were equivalent between women and men, except for the item about crying. That constraint needed to be released between gender due to the much smaller relationship between endorsement of crying with depression for men compared to women. This corresponds to item-response bias found for this item in Black and White adults, where the item about crying was higher in women vs. men when matched on overall depressive symptoms (Cole et al., 2000). While the item about crying was non-invariant across gender, we do not endorse removing that item from CES-D analyses as future studies may find that changing gender norms at the societal level may lead to this item reflecting depression for men and women, equally.

Female AI elders reported significantly higher levels on the depressed affect and somatic symptoms, and lower levels of wellbeing subscales, than AI male elders. This is in line with national reports of gender differences in depression prevalence but in contrast to the lack of CES-D latent mean differences in a non-AI sample of older adults (Gomez & McLaren, 2015) and previous research on lifetime depressive disorder with AI adults in Northern Plains and Southwest tribes (Beals et al., 2005). However, as the CES-D measures past week depressive symptomology and not lifetime depressive disorder, our results indicate gender differences in symptomology in AI elders and cannot determine gender differences in depression diagnoses. Our work highlights higher levels of depressive symptomology in AI women, compared to AI men, in a population that has disproportionately high rates

of mental health issues compared to the general US population (Heart, Chase, Elkins, & Altschul, 2011). As depression is linked to illnesses that are prevalent in AI populations (CVD, stroke, and ADRD-related comorbidities), we must consider gender differences in culturally tailored prevention and treatment programs with AI adults, with focused attention on depressive symptomology for AI older women.

The parent study was conducted among residents of 13 tribal and AI communities (Suchy-Dicey et al., 2016), therefore results may not be generalizable to people living on other tribal lands and regions, or to the more than 70% of AI adults who reside in urban areas (Indian Health Services, 2018). Additional limitations are the self-report nature and inherent limitations of a symptom survey attempting to capture a construct as complex as depression, as well as the assessment of gender solely as a binary variable. Given the variability across tribal communities within the United States and across Indigenous populations in Northern America, we advocate for additional psychometric research on the CES-D. While we have assessed convergent validity of the 20-item CES-D with age, gender, education, income, marital status, alcohol use, smoking, hypertension, diabetes, stroke, processing speed, general cognition, verbal fluency, and motor function in the CDCAI study (Suchy-Dicey et al., 2020), we have not assessed convergent and discriminant validity of the 12-item CES-D scale. We propose this as future work to be done after approval from the tribal councils. As the CDCAI study and the Strong Heart Study represent the largest cohort of aging American Indians, we offer additional evidence for using the 12-item CES-D for American Indian adults in other research studies with similar samples, but suggest additional psychometric work be undertaken (internal consistency, test-retest reliability, convergent and discriminant validity, invariance across age bands, longitudinal measurement invariance, etc.). Additional work on assessing measurement invariance across socioeconomic strata, and the intersection between socioeconomic status, gender, and race, will be critical to ensure that depression is assessed equivalently across groups.

While a 10-item version of the CES-D has shown to be a valid and reliable measure of depression in the general population (Andresen, Malmgren, Carter, & Patrick, 1994; Mohebbi et al., 2018), and with older populations (Cheng & Chan, 2005), this short form has not been analyzed in American Indian samples. Further research is needed to determine if the 10-item CES-D is a valid and reliable measure of depression in American Indian older adults. As the Patient Health Questionnaire (PHQ-9) (Kroenke, Spitzer, & Williams, 2001) is more commonly used in clinical practice, of note is that 5 of the items in the PHQ-9 are similar to items in the 12-item CES-D (items about being depressed, change in appetite, trouble with sleep, lack of energy, and difficulty concentrating). The 12-item CES-D may offer information of presence (number of items endorsed) and severity of depression symptomology (summed scored of the 12-item CES-D), however, additional research is needed to identify clinically relevant cut-scores.

## Conclusion

We tested the psychometric properties of the CES-D in AI elders and tested differences in depression constructs between gender. Results indicate that while the 20-item CES-D did not fit the data, the 12-item CES-D demonstrated a valid factor structure and reliable total

score with AI elders. Female AI elders had higher levels on the depressed affect and somatic symptoms, and lower levels of wellbeing, than AI male elders. Healthcare providers must understand cultural factors in the AI elder's experience when assessing depression. Future work should continue to test the psychometric properties of psychological scales with AI adults in order to determine mean scores and test for differences between groups and assess change over time.

## Funding

## References

American Psychological Association Division of Diversity and Health Equity. (2017). Mental Health Disparities: American Indians and Alaska Natives.

Andresen EM, Malmgren JA, Carter WB, & Patrick DL (1994). Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). American Journal of Preventative Medicine, 10(2), 77–84.

Barón AE, Manson SM, Ackerson LM, & Brenneman DL (1990). Depressive symptomatology in older American Indians with chronic disease: Some psychometric considerations. In Zich CCAJM (Ed.), Depression in primary care: Screening and detection (pp. 217–231). Florence, KY, US: Taylor & Frances/Routledge.

Beals J, Novins DK, Whitesell NR, Spicer P, Mitchell CM, & Manson SM (2005). Prevalence of mental disorders and utilization of mental health services in two American Indian reservation populations: mental health disparities in a national context. American Journal of Psychiatry, 162(9), 1723–1732. doi: 10.1176/appi.ajp.162.9.1723

Brown TA (2006). Confirmatory factor analysis for applied research. New York: Guilford Press.

Centers for Disease Control and Prevention. (2014). Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2012 RetrievedSeptember 19, 2019, from http://www.cdc.gov/nchs/data/series/sr_10/sr10_260.pdf

Chapleski EE, Lamphere JK, Kaczynski R, Lichtenberg PA, & Dwyer JW (1997). Structure of a depression measure among American Indian elders: Confirmatory factor analysis of the CES-D scale. Research on Aging, 19(4), 462–485.

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural Equation Modeling-a Multidisciplinary Journal, 14(3), 464–504.

Chen FF, Sousa KH, & West SG (2005). Testing measurement invariance of second-order factor models. Structural Equation Modeling-a Multidisciplinary Journal, 12(3), 471–492.

Cheng ST, & Chan AC (2005). The Center for Epidemiologic Studies Depression Scale in older Chinese: thresholds for long and short forms. International Journal of Geriatric Psychiatry, 20(5), 465–470. doi: 10.1002/gps.1314 [PubMed: 15852439]

Cole SR, Kawachi I, Maller SJ, & Berkman LF (2000). Test of item-response bias in the CES-D scale: experience from the New Haven EPESE Study. Journal of Clinical Epidemiology, 53(3), 285–289. doi: 10.1016/S0895-4356(99)00151-1 [PubMed: 10760639]

Diniz BS, Butters MA, Albert SM, Dew MA, & Reynolds CF 3rd. (2013). Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies. British Journal of Psychiatry, 202(5), 329–335. doi: 10.1192/bjp.bp.112.118307

Gomez R, & McLaren S (2015). The Center for Epidemiological Studies Depression Scale: Measurement and structural invariance across ratings of older adult men and women. Personality and Individual Differences, 75, 130–134. doi: 10.1016/j.paid.2014.11.005

Hackett ML, Yapa C, Parag V, & Anderson CS (2005). Frequency of depression after stroke: a systematic review of observational studies. Stroke, 36(6), 1330–1340. [PubMed: 15879342]

Heart MY, Chase J, Elkins J, & Altschul DB (2011). Historical trauma among Indigenous Peoples of the Americas: concepts, research, and clinical considerations. Journal of Psychoactive Drugss, 43(4), 282–290. doi: 10.1080/02791072.2011.628913

Indian Health Services. (2018). Urban Indian Health Program, from https://www.ihs.gov/newsroom/factsheets/uihp/

Kroenke K, Spitzer RL, & Williams JB (2001). The PHQ-9: validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. doi: 10.1046/j.1525-1497.2001.016009606.x [PubMed: 11556941]

Kuehner C (2003). Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. Acta Psychiatrica Scandinavica 108(3), 163–174. [PubMed: 12890270]

Mohebbi M, Nguyen V, McNeil JJ, Woods RL, Nelson MR, Shah RC, … Group AI (2018). Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. General Hospital Psychiatry, 51, 118–125. doi: 10.1016/j.genhosppsych.2017.08.002 [PubMed: 28890280]

Musselman DL, Evans DL, & Nemeroff CB (1998). The relationship of depression to cardiovascular disease: epidemiology, biology, and treatment. Archives of General Psychiatry 55(7), 580–592. doi: 10.1001/archpsyc.55.7.580 [PubMed: 9672048]

Muthén LK, & Muthén BO (1998–2018). Mplus User's Guide (Eight Edition ed.). Los Angeles: Muthén & Muthén.

Nemeroff CB, Musselman DL, & Evans DL (1998). Depression and cardiac disease. Depression and Anxiety, 8 Suppl 1, 71–79.

Radloff LS (1977). The CES-D scale: A self report depression scale for research in the general population. Applied Psychological Measurement, 1, 385–401.

Schure M, & Goins RT (2017). Psychometric Examination of the Center for Epidemiologic Studies Depression Scale with Older American Indians: The Native Elder Care Study. American Indian and Alaska Native Mental Health Research, 24(3), 1–13. doi: 10.5820/Aian.2403.2017.1

Somervell PD, Beals J, Kinzie JD, Boehnlein J, Leung P, & Manson SM (1992). Use of the Ces-D in an American-Indian Village. Culture Medicine and Psychiatry, 16(4), 503–517.

Somervell PD, Beals J, Kinzie JD, Boehnlein J, Leung P, & Manson SM (1993). Criterion validity of the Center for Epidemiologic Studies Depression Scale in a population sample from an American Indian village. Psychiatry Research, 47(3), 255–266. [PubMed: 8372163]

Suchy-Dicey A, Verney SP, Nelson LA, Barbosa-Leiker C, Howard BA, Crane PK, & Buchwald DS (2020). Depression Symptoms and Cognitive Test Performance in Older American Indians: The Strong Heart Study. Journal of the American Geriatrics Society, doi: 10.1111/jgs.16434

Suchy-Dicey AM, Shibata D, Best LG, Verney SP, Longstreth WT Jr., Lee ET, … Buchwald D (2016). Cranial Magnetic Resonance Imaging in Elderly American Indians: Design, Methods, and Implementation of the Cerebrovascular Disease and Its Consequences in American Indians Study. Neuroepidemiology, 47(2), 67–75. doi: 10.1159/000443277 [PubMed: 27603047]

Zhang Y, Galloway JM, Welty TK, Wiebers DO, Whisnant JP, Devereux RB, … Yeh J (2008). Incidence and risk factors for stroke in American Indians. Circulation, 118(15), 1577–1584. [PubMed: 18809797]

## Public Significance Statement

An abbreviated 12-item Center for Epidemiological Studies-Depression scale assesses depression in American Indian female and male elders. Female AI elders had higher levels on the depressed affect and somatic symptoms, and lower levels of wellbeing, than AI male elders.

**Table 1**

Model fit of hypothesized CES-D Models for the total sample

| Model | $\chi^2$ | $df$ | CFI | SRMR | RMSEA |
|---|---|---|---|---|---|
| 20-item | | | | | |
| 1-Factor Higher-Order | 541.20[*] | 166 | .885 | .054 | .053 |
| 4-Factor | 530.69[*] | 164 | .888 | .052 | .052 |
| 3-Factor | 604.25[*] | 167 | .866 | .056 | .057 |
| 12-item | | | | | |
| 3-Factor | 187.88[*] | 51 | .920 | .047 | .057 |
| 1-Factor | 528.22[*] | 54 | .723 | .090 | .104 |

*Note.* $\chi^2$ = Chi-square. *df* = Degrees of Freedom. CFI = Comparative Fit Index. SRMR = Standardized Root Mean Square Residuals. RMSEA = Root Mean Square Error Approximation.

[*]
*p* .05.

**Table 2**

Standardized factor loadings (standard error) for the 3-factor 12-item CES-D between gender

| | Women N=554 | Men N=263 |
|---|---|---|
| **Depressed Affect** | | |
| During the past week I felt depressed | .79 (.03) | .73 (.08) |
| During the past week I felt sad | .72 (.04) | .50 (.08) |
| During the past week I felt lonely | .73 (.04) | .72 (.07) |
| During the past week I felt that I could not shake the blues even with help from my family or friends | .65 (.05) | .58 (.08) |
| During the past week I had crying spells | .54 (.05) | .20 (.09) |
| **Somatic Symptoms** | | |
| During the past week I did not feel like eating; my appetite was poor | .49 (.05) | .43(.09) |
| During the past week I felt like I couldn't do what I needed to do | .62 (.04) | .50 (.09) |
| During the past week my sleep was restless | .56 (.05) | .60 (.06) |
| During the past week I had trouble keeping my mind on what I was doing | .62 (.04) | .72 (.07) |
| **Well-being** | | |
| During the past week I felt hopeful about the future | .54 (.05) | .40 (.12) |
| During the past week I was happy | .82 (.05) | .91 (.20) |
| During the past week I enjoyed life | .67 (.05) | .61 (.15) |

**Table 3**

Invariance of the 3-factor, 12-item CES-D model between gender

| Model | $\chi^2$ | df | CFI | SRMR | RMSEA |
|---|---|---|---|---|---|
| 1. Configural invariance | 175.19[*] | 100 | .958 | .046 | .043 |
| 2. Metric Invariance **A** | 226.22[*] | 112 | .936 | .083 | .050 |
| 3. Metric Invariance **B** | 197.03[*] | 111 | .952 | .066 | .044 |
| 4. Scalar Invariance [†] | 212.66[*] | 119 | .947 | .067 | .044 |
| **Model Comparison** | $\chi^2$ | df | CFI | SRMR | RMSEA |
| 1 vs. 2 | +51.03 | 12 | −.022 | +.037 | +.007 |
| 1 vs. 3 | +21.84 | 11 | −.006 | +.020 | +.001 |
| 3 vs. 4 | +15.63 | 8 | −.005 | +.001 | .000 |

Note. $\chi^2$=Chi-square. df=Degrees of Freedom. CFI=Comparative Fit Index. SRMR=Standardized Root Mean Square Residuals. RMSEA=Root Mean Square Error Approximation. $\chi^2$ reflects the difference in $\chi^2$ values between nested models.

[*] p .05. Metric Invariance model **A**: correlated error between Sad and Crying items only. Metric Invariance model **B**: correlated error between Sad with Crying items and Crying item released from constraints between gender.

[†=] Scalar invariance model includes a correlated error between Sad and Crying items as well as Crying item released from constraints between gender.