# The Center for Epidemiological Studies Depression Scale: Measurement and structural invariance across ratings of older adult men and women

Rapson Gomez *, Suzanne McLaren

*Federation University Australia, Australia*

## ARTICLE INFO

## ABSTRACT

The current study examined the measurement and structural invariance of the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977) across ratings provided by older adult men ($n = 573$) and women ($n = 605$). Ratings were modeled in terms of the original four-factor oblique model: a simple structure with correlated factors for Depressed Affect, Positive Affect, Somatic Symptoms, and Interpersonal Difficulties. Multiple-group confirmatory factor analysis supported full measurement and structural invariance, and no sex difference for the four latent factors. These findings indicate good support for measurement and structural invariance of CES-D ratings across older adult men and women. The psychometric and practical implications of the findings are discussed.

© 2014 Published by Elsevier Ltd.

## 1. Introduction

High levels of depression symptoms in older adults have serious negative consequences, including increased burden of physical illness, impaired functioning, and risk of suicide (Fiske, Wetherell, & Gatz, 2009). Thus we need reliable and valid measures for screening depression symptoms in older adults. The Center for Epidemiological Studies Depression Scale (CES-D) is a 20-item self-rating questionnaire, developed by Radloff (1977) that is used widely in research and clinical settings for screening depressive symptoms in adults, including older adults (Ros et al., 2011). The current study used multiple-group confirmatory factor analysis (CFA) to examine the measurement and structural invariance of ratings of the CES-D items provided by older adult men and women. Additionally, it examined the differences in latent mean scores across these groups.

The original theoretical structure of the CES-D is a four-factor oblique model, with factors for Depressed Affect (DA), (lack of) Positive Affect (PA), Somatic Symptoms and Retarded Activity (SS), and Interpersonal Difficulties (ID) (see Fig. 1; Radloff, 1977). Although alternate factor models for the CES-D have been proposed, two separate meta-analyses of factor analyses studies of the CES-D have concluded most support for the original four-factor model (Kim, DeCoster, Huang, & Chiriboga, 2011; Shafer, 2006).

Currently there are CFA studies showing support for this structure in older adults (Foley, Reed, Mutran, & DeVellis, 2002; Hertzog, Van Alstine, Usala, Hultsch, & Dixon, 1990; Ros et al., 2011).

Despite having four separate factors, when scoring the CES-D, most studies have used the total score based on all items (Edwards, Cheavens, Heiy, & Cukrowicz, 2010). Some recent studies with older adults have however shown that the different CES-D scales scores are differentially associated with relevant external variables (e.g., St. John, Tyas, & Montgomery, 2013; Yu, Li, Cuijpers, Wu, & Wu, 2012). For instance, Yu et al. found that while age was associated positively with SS, it was associated negatively with DA and ID, and it had no relation with the total score. Also, while family support was the strongest predictor for DA, health status was the strongest predictor for SS. Findings such as these highlight the relevance, usefulness and importance of using the CES-D scales in studies involving older adults.

If the CES-D scale scores are to be used with confidence in future studies with older adults, then it is critical that there is good support for the psychometric properties of the four-factor oblique model in this group. A psychometric property that could foster more confidence in the use of the CES-D factor scales in this group would be support for measurement invariance for the CES-D in terms of the four-factor oblique model across men and women. Measurement invariance refers to groups reporting the same observed scores when they have the same level of the underlying trait (Reise, Widaman, & Paugh, 1993). Invariance would mean that for the groups being compared, the measure in question has the same measurement and scaling properties. If there is weak or no

---

* Corresponding author at: School of Health Sciences, Federation University Australia, University Drive, Mt Helen, PO Box 663, Ballarat, Victoria 3353, Australia.
  E-mail address: rapson.gomez@federation.edu.au (R. Gomez).

support for invariance, then it follows that the groups in question cannot be justifiably compared in terms of observed scores, as the same observed scores for the groups do not reflect the same levels of the underlying trait. When applied to the CES-D, support for measurement invariance for sex would mean that men and women have the same observed scores for the same level of the latent trait. If the majority of items in the scale show measurement invariance, the observed scale scores can be compared directly across the groups, as the scores across the groups are not confounded by differences in measurement and scaling properties.

For CES-D ratings from older adults, there are reasons to suspect that they could lack measurement invariance. Item response theory (IRT) based studies involving older adults have consistently shown lack of invariance for the item "I had crying spells", with women endorsing higher ratings for the same level of depression (Cole, Kawachi, Maller, & Berkman, 2000; Gelin & Zumbo, 2003; Yang & Jones, 2007; Yang, Tommet, & Jones, 2009). The Yang et al. study also showed lack of measurement invariance for "I felt like a failure", "People were unfriendly" and "I felt that people disliked me". For all three items, men endorsed higher ratings for the same level of depression as women.

A powerful method for examining invariance is the multiple-group CFA approach. Assuming that the indicator-ratings are treated as ordered-categorical (as in the CES-D), this approach can test for configural invariance (same overall factor structure), item fac-tor loadings invariance (same strengths of the associations of items with the first-order factors), invariance for item thresholds (equivalency for responses for the different item categories), and error variances invariance (equivalency in the error variances of the items). When there is support for invariance for item factor loadings and thresholds, structural invariance or invariance for latent variance and covariance can be tested. In addition, the equivalency for the factor mean scores can be computed after removing the items that lack invariance.

Using a multiple-group CFA approach, in a study by Stommel et al. (1993) that included some older adults (percentage not reported), non-invariance was found for the factor loadings of the "crying spell" (higher for women) and "I talked less than usual" (higher for men) for a three-factor CES-D model that excluded the two ID items. O'Rourke (2005) also examined measurement invariance in a group of adult men and women that included some older adults (percentage not reported). This was examined for a higher order factor model, in which the lower order factors were the four original CES-D factors. The study found lack of invariance for the factor loadings for the items on "I felt that I could not shake off the blues even with help from my family or friends" (higher for men), "I felt lonely" (higher for men), and "I enjoyed life" (higher for women). Currently, no study has examined measurement invariance across sex for a group of exclusive older adult men and women in terms of the original four-factor model.

For the original CES-D four-factor oblique model, the current study used multiple-group CFA approach to examine measurement (configural, factor loadings, intercepts and error variances) and structural (factor variances and covariances) invariance across the ratings of older adult men and women. Additionally, the study examined the differences in latent mean scores across the groups.

## 2. Method

### 2.1. Participants

A convenience sample of 1178 older adults participated in the study. The sample consisted of 573 men aged from 65 to 96 years of age ($M = 74.98$, $SD = 7.48$) and 605 women aged from 65 to 98 years ($M = 75.01$, $SD = 7.65$). The men and women did not differ significantly on age, $t(1176) = -.06$, $p > .05$. As can be seen from Table 1, most men were married, and had completed secondary school or equivalent. Almost half of the women were married and most had completed secondary school or equivalent. There was a significant association between gender and marital status, $\chi^2(3, N = 1168) = 14.35$, $p = .002$, with more men being married ($n = 313$; standardized adjusted residuals statistic or HAR = 2.6, $p < .01$), and more women being divorced ($n = 95$; HAR = 3.0, $p < .01$) than that expected by chance ($n$ for married = 597, and $n$ for divorced = 151). There was no relationship between gender and highest level of education, $\chi^2(2) = 3.74$, $p > .05$. Men
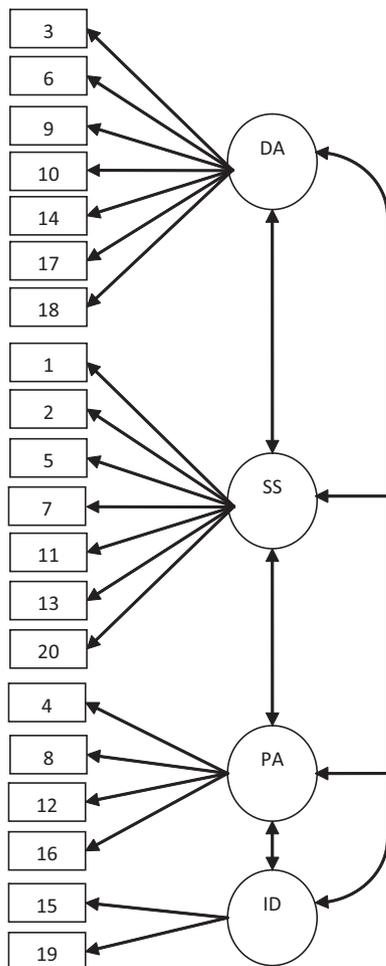


**Fig. 1.** Schematic path diagram of the 4-factor oblique CES-D model. *Note.* The residuals are not shown in the figure. DA = Depressed Affect; PA = Positive Affect; SS = Somatic Symptoms; ID = Interpersonal Difficulties. The numbers in the rectangles refer to the item numbers in the CES-D.

**Table 1**
Relationship status and highest level of education for participants.

| Variable | Men | | Women | | Total | |
|---|---|---|---|---|---|---|
| | $n$ | % | $n$ | % | $n$ | % |
| Relationship status | | | | | | |
| Married | 313 | 54.6 | 284 | 46.9 | 597 | 50.7 |
| Single | 55 | 9.6 | 46 | 7.6 | 101 | 8.6 |
| Divorced | 56 | 9.8 | 95 | 15.7 | 151 | 12.8 |
| Widowed | 149 | 26.0 | 180 | 29.8 | 329 | 27.9 |
| Highest level of education | | | | | | |
| Primary School | 87 | 15.2 | 73 | 12.1 | 160 | 13.6 |
| Secondary or Trade School | 401 | 70.0 | 453 | 74.9 | 854 | 72.5 |
| University Degree | 85 | 14.8 | 79 | 13.0 | 164 | 13.9 |