Psychiatry Research 179 (2010) 107–112

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

An examination of the measurement adequacy of the CES-D among African American women family caregivers

Philip A. Rozario [a,*], Natasha Menon [b,1]

[a] Adelphi University School of Social Work, One South Avenue Box 701, Garden City, NY 11530, United States
[b] Fordham University, Graduate School of Social Service, 113 West 60th Street, New York, NY 10023, United States

ARTICLE INFO

Article history:
Received 27 May 2009
Received in revised form 2 June 2010
Accepted 23 June 2010

Keywords:
Depression
Mental Health
Ethnicity
Confirmatory factor analysis

ABSTRACT

The Center for Epidemiologic Studies Depression Scale (CES-D) has been used extensively in community-based surveys to describe and explain the prevalence of depression in the general population. Yet, questions have been raised regarding its adequacy for use among ethnic minority because of its factor variance. Employing a within-gender and race approach, we test the validity of the CES-D for use among a sample of African American women family caregivers. Using data from a cross-sectional community sample of 521 urban and rural African American women family caregivers, this study examines the dimensionality of the CES-D by testing four different measurement models through confirmatory factor analyses. Among the four measurement models tested using Weighted Least Squares estimation, our findings support previous research that has identified four dimensions in the CES-D: depressed affect, positive affect, somatic complaints, and interpersonal relations for our sample. Additionally, a three-factor (somatization) model and a four-factor model were shown to be equivalent. Implications for further measurement and model testing, and the use of the CES-D for research among African American women caregivers are discussed.

1. Introduction

Developed to measure depressive symptoms in the general population, the Center for Epidemiologic Studies Depression Scale (CES-D) has been used extensively by researchers in community-based surveys to describe and explain the prevalence of depression (Radloff, 1977; Long-Foley et al., 2002; Perreira et al., 2005). The CES-D has high internal consistency (Callahan and Wolinsky, 1994; Radloff, 1977), and is generally assumed to assess the same underlying symptoms of depression among different groups (Aneshensel et al., 1983). Despite its extensive use in epidemiological and other community-based studies, some researchers have questioned its adequacy for use among different racial/ethnic groups because of its measurement variance across different racial and gender groups (Long-Foley et al., 2002; Perreira et al., 2005).

For example, Callahan and Wolinsky (1994) found considerable variance in depressive rates, as measured by the CES-D, for different race–gender combinations of older primary care patients. In their study, they suggested that significant differences among these groups might be due to three possible explanations: high non-response rates because of their lower functioning sample, socio-cultural differences, and/or measurement artifact, i.e., the wording of items on the scale. These explanations highlight issues related to different aspects of the lack of measurement equivalence (or invariance) that may be due to biases in the conceptualization of the construct of interest, methodological strategies, or administration of the measurement (Byrne and Watkins, 2003). Further, in a study using a convenience sample of older African Americans, ages 59 to 96 years, Long-Foley et al. (2002), found further evidence that items of the CES-D loaded differently on four factors in comparison to previous works. Although these studies illuminate some of the problems inherent in the dimensionality of the scale, the generalizability of these findings are limited because these researchers relied on convenience samples that included African Americans.

Several researchers have found the four-factor structure to be consistent across different samples. In their examination of three different samples (two of which were convenience samples of community-dwelling African American men and women with a history of drug use and an exclusively White sample, as well as a stratified random sample comprising a nationally representative sample of African Americans), Nguyen et al. (2004) found that the four-factor structure of the CES-D was robust across all three samples. Still, Nguyen et al. (2004) reported that they did not find evidence of factor loading invariance across the subgroups in their sample and argued that the factors might not be measuring the same facets of depression. Similarly, Blazer et al. (1998) found support for the four-factor structure model for a representative sample of older adults in North Carolina. However, in their study using data from a nationally representative sample of youth ages 12–20 years, Perreira et al. (2005) found that the CES-D was not

* Corresponding author. Tel.: +1 517 877 4419.
E-mail addresses: rozario@adelphi.edu (P.A. Rozario), menon@fordham.edu (N. Menon).
[1] Tel.: +1 212 636 7784.

0165-1781/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2010.06.022

psychometrically equivalent across different racial–ethnic groups, that consisted of White, Black, Asian, Hispanic youth who were either immigrants or native-born. They posited that these cross-ethnic differences might be due to issues of respondents' unfamiliarity with the CES-D and social desirability.

In addition to the multifaceted nature of depression, its measurement is made more complex from a cross-cultural standpoint in that researchers recognize that depression is sometimes manifested and expressed differently across various cultural groups (Iwata et al., 2002). Indeed, some researchers argue that African Americans are more likely to somaticize their depression than their White counterparts (Blazer et al., 1998). As such, researchers have to exercise caution in making cross-cultural comparisons because in comparisons of psychological constructs like depression "scores are interpretable only in light of evidence that the meaning and dimensional structure of the construct… as well as the items comprising the measuring instrument are group-equivalent" (Byrne et al., 2009, p. 95).

Our sole focus on African American women caregivers stems from the observation that depression is disproportionately represented among women and non-whites (Barbee, 1992) and that most researchers conceptualize negative outcomes of caregiving as depressive symptoms and many operationalize the symptoms with the CES-D (Pinquart and Sörensen, 2003). In her review of depressive symptoms among African American women, Barbee (1992) emphasized the importance for researchers to consider contextual factors such as racial and gender status when examining depression among African American women. By limiting our analyses to African American women caregivers, we are better able to address the diversity that exists within one racial and gender group (Rozario et al., 2008). In employing a within-gender and race approach, we test the validity of the CES-D for use among a group of African American women family caregivers.

2. Methods

2.1. Study participants

Data came from a cross-sectional study of a community sample of 521 urban and rural African American women family caregivers of African American elders (over 65 years). The data collection for the original study, the Black Rural and Urban Caregivers Mental Health and Functioning Study, was conducted between July 1999 and August 2002 in urban and rural locations in Missouri with the approval of the Institutional Review Board of Washington University in St. Louis.

The sample of caregivers was obtained using reverse screening methodology on a list of Medicare enrollees who were African Americans, 65 years or older, and residing in the St. Louis metropolitan area or in seven rural Southeast counties. The researchers used stratified random sampling method to obtain a list of African American elders from the list of urban Medicare enrollees, while they selected everyone from the list of rural Medicare enrollees. Because of the smaller pool of elderly Black Medicare enrollees in the rural counties, the principal investigator and her team decided to screen all 1994 enrollees to achieve a pool of 285 eligible caregivers. For the rural caregivers, the researchers achieved a response rate of almost 93%. They yielded an overall response rate of 88%.

Trained interviewers screened the identified enrollees to determine if they met the inclusion criteria of self-identification as either African American, Black, Negro, or Colored; being 65 years or older; and receiving unpaid help from an African American female caregiver with at least one activity of daily living, or one instrumental activity of daily living, or decision making. The screeners asked eligible elders to provide the contact information of up to two unpaid African American women who provide them with help. The elders identified ninety-five percent of their caregivers as helping them the most, while the remaining caregivers were secondary helpers. A second screening was done with the caregivers to verify if they met the criteria for inclusion in the study, which were being African American women and providing unpaid help to their elder relatives.

Upon receiving written consent, in-home interviewers conducted interviews that lasted approximately 2.5 h using computerized assisted personal interview. Caregivers were compensated $15 for their participation upon completion of the interviews. Further details of the sample selection are reported elsewhere (see Chadiha et al., 2004).

2.2. Analysis

Comprising 20 items, the original CES-D Scale measures the respondent's self-reported current level of depressive symptoms, with an emphasis on the affective dimension of depressive symptoms (Radloff, 1977). Sample statements include "You did not feel like eating; your appetite was poor," "You had trouble keeping your mind on what you were doing," "You felt depressed," and "You felt everything was an effort." Responses ranged from 0 = rarely or none of the time to 3 = most or all the time. Four items were reversed scored and a summative score was obtained for each respondent. Since we were interested in a-priori testing of and comparability with previously identified measurement models in Perreira et al.'s (2005) study of CES-D, we excluded two items, "Your sleep was restless" and "You had crying spells," from our factor analyses. In their secondary data analyses, Perreira et al. (2005) excluded these items because the dataset that they relied on, ADD-Health, did not include these items in their questionnaire.

For the psychometric analyses, we conducted descriptive analyses ($n = 521$), including item-by-item description and item-total correlations. Then we performed confirmatory factor analyses (CFA) on four different measurement models to examine the dimensionality of the CES-D. The four structural models were: a) a single-factor model, b) a three-factor model with Depressed and Positive Affect, Somatic Complaints, and Interpersonal Relations as factors (Model 1: Positive Affect), c) another three-factor model with Depressed Affect and Somatization, Positive Affect, and Interpersonal Relations as factors (Model 2: Somatization), and d) a four-factor model with Depressed Affect, Positive Affect, Somatic Complaints, and Interpersonal Relations as factors (see Table 1). We chose a confirmatory approach over an exploratory approach given the theoretical and a-priori nature of the relationships between the observed and latent variables. We used the maximum likelihood (ML) estimation method and the weighted least squares (WLS) estimation method to analyze the factor structure. We used ML to reproduce and compare Perreira et al.'s results with our sample of African American female caregivers. However, as individual CES-D items were ordinal in measurement, we re-estimated the measurement models using WLS, and compared the results. The WLS is an appropriate estimation method for use when the data are ordinal in nature, and data are not assumed to be normally distributed (Byrne, 1998; Schumacker and Lomax, 2004). We inputted a polychoric correlation and an asymptotic covariance matrix to conduct the WLS analyses. All analyses were performed using LISREL 8.80 and PRELIS 2.0 (Jöreskog and Sörbom, 2007).

The CFA involves the use of structural equations to estimate the relationships between observed variables and latent variables. The use of the structural equations framework also permits the estimation of measurement error simultaneously. A total of 18 observed variables were used to estimate different measurement models as detailed by Perreira et al. (2005). All four measurement models were evaluated through a series of nested CFA. As recommended by Bentler (2007), all models were evaluated for fit using multiple criteria namely the Chi-square statistic, the Chi-square/degree of freedom ratio, the Root Mean Square Error of Estimation (RMSEA), the 90% Confidence Interval of the RMSEA, the Comparative Fit Index (CFI), and the Goodness-of-fit Index (GFI). We also used the model Akaike Information Criterion (AIC) to compare alternative models (Kline, 1998). All variables were observed for deviations from normality and listwise deletion was used for missing data.

3. Results

3.1. Description of the sample

The average age of our sample of African American women caregivers was 53.8 years (S.D. = 15.05 years, range 19–92 years). The majority of caregivers were daughter caregivers (57%) followed by wife caregivers (23%). About 55% of the caregivers reported that they were widowed, and only 29% reported that they were married. Our sample reported receiving an average of 12.3 years (S.D. = 2.74 years) in formal education. Income wise, the median household income was $18,500 (mean = $24,000, S.D. = $18,444). Our sample was almost evenly divided between urban (49.1%) and rural (50.9%) caregivers. Urban caregivers reported a significantly higher mean annual income, $29,531 (S.D. = 20,237) than their rural counterpart, $17,780 (S.D. = 13,440).

The mean CES-D score for our sample was 8.99 (S.D. = 8.32). The Cronbach's alpha for the CES-D for this sample was high ($\alpha = 0.83$). Using the cutoff of 16, we found that slightly more than 18% of the caregivers were at risk of clinical depression. Indeed, this is considerably a higher prevalence rate of depression than that of Blazer et al.'s epidemiological findings among older African Americans (9.5%). In our bivariate analyses, we did not find any statistically significant difference in the CES-D scores between rural and urban caregivers. Table 2 presents the item-by-item descriptive analyses of the CES-D Scale. The responses of African American caregivers were skewed towards less depressive symptoms in our sample, which is similar to Long-Foley et al.'s (2002) findings of a sample of older African Americans. On average, about 72% of the sample reported rare or no "negative" symptoms (or most of the time for "positive" symptoms) for all items. This is a lower percentage than that of Long-Foley et al.'s community-based sample of older African Americans.

3.2. Confirmatory factor analysis

Based on Perreira et al.'s conceptual framework, we first performed CFA on the four measurement models using ML estimation.