Available online at www.sciencedirect.com
SCIENCE DIRECT®
Psychiatry Research 126 (2004) 177–187
PSYCHIATRY RESEARCH
www.elsevier.com/locate/psychres

Factorial invariance of the CES-D in low socioeconomic status African Americans compared with a nationally representative sample

Ha T. Nguyen[a,*], Melissa Kitner-Triolo[a], Michele K. Evans[b], Alan B. Zonderman[a,b]

[a]*Cognition Section, Laboratory of Personality and Cognition, National Institute on Aging, National Institutes of Health, Gerontology Research Center, 5600 Nathan Shock Drive, Baltimore, MD 21224, USA*
[b]*Health Disparities Section, Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA*

Received 6 February 2003; received in revised form 12 January 2004; accepted 11 February 2004

Abstract

This study examined the factor structure of the Center for Epidemiologic Studies Depression Scale (CES-D) in low socioeconomic status African Americans ($n=426$). Confirmatory factor analysis indicated that the four factors—(1) depressed affect, (2) positive affect, (3) somatic complaints, and (4) interpersonal problems—of the CES-D scale previously found in the general population were supported in this sample. These results were cross-validated in other nationally representative samples of African-American participants ($n=988$) and Caucasians ($n=666$), and the four-factor structure of the CES-D scale replicated in all three groups in this study. An alternative model was also tested with the factors depressed affect and somatic complaints combined as a single factor, a finding often reported in minority groups. Results indicated a significantly poorer fit for the three-factor model compared with the four-factor model for all three groups. In addition, higher loading differences were significantly evident between African-American and Caucasian groups, while higher loading similarities were found between the two African-American groups. These findings provide further evidence of measurement equivalency of the CES-D scale in samples with differential characteristics including race and socioeconomic status.
© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Blacks; Ethnicity; Depression; Factor structure; Measurement equivalency

1. Introduction

The maintenance of mental health plays an important role for the preservation of quality of life, including physical, social and cognitive functioning. However, epidemiologic studies report a greater risk of major depression for women (5–9%) compared with men (2–3%) (Pearlin, 1989; Blazer et al., 1991; Rhee et al., 1999). Higher levels of depressive symptoms are also reported among minority groups. Specifically, African Americans are reported as having generally higher prevalence rates of depressive symptoms than Cau-

*Corresponding author. Tel.: +1-410-558-8072; fax: +1-410-558-8674.
E-mail address: nguyenha@grc.nia.nih.gov (H.T. Nguyen).

0165-1781/04/$ - see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.psychres.2004.02.004

casians (Jones-Webb and Snowden, 1993; McBarnette, 1996). Other studies, however, provided no evidence to support differences in the rates of depressive symptoms between African Americans and Caucasians. These community-based studies found the same proportion of African Americans as Caucasians scoring above the cut-off point for clinically significant depressive symptoms on the Center for Epidemiological Studies Depression Scale (CES-D) (Murrell et al., 1983; Berkman et al., 1986).

Research on ethnic differences in depression evidences much variety and many inconsistencies, and these inconsistencies could be due to a lack of validation of depressive surveys among population subgroups. Cultural and social differences among minority populations may result in a disagreement about the conceptualization of depressive symptoms. For example, compared with Caucasians, African Americans tend to incorporate physical complaints into their answers of questions designed to assess affective symptoms (Brown et al., 1996). Given such evidence, depressive symptoms may be under-reported or exacerbated in African Americans when applying depressive diagnoses or standard measures that have been validated among Caucasians. Thus, additional research is needed to assess the sensitivity, reliability, and validity of depressive instruments for use in population subgroups, particularly in African Americans. The present study examines the validity of the CES-D factor structure in African Americans.

Developed by the National Institute of Mental Health Center for Epidemiologic Studies, the CES-D is a 20-item inventory that has been widely used in assessing depressive symptomatology in community and population-based studies. Radloff (1977) examined the factor structure of the CES-D and identified the following four factors: (1) depressed affect, (2) positive affect, (3) somatic complaints, and (4) interpersonal problems. Radloff's four-factor structure has been extensively replicated and is widely accepted in epidemiologic studies of depression in predominantly Caucasian populations (Weissman et al., 1977; Hertzog et al., 1990; Knight et al., 1997). The factor structure of the CES-D in ethnic minorities, however, has not been extensively validated, and there is little consensus about its underlying structure among available studies. Roberts (1980) examined the measurement adequacy of the CES-D and found the same general factor structure of responses among Caucasians, African Americans, and Mexican Americans. In contrast, Chapleski et al. (1997), using confirmatory factor analysis, did not identify a four-factor model in American Indians. A shortened 12-item CES-D scale was more useful for assessing depressive symptoms in their study of American Indians. Three studies examined the factor structure of the CES-D scale in Hispanic Americans, and the results were not consistent. Liang et al. (1989) replicated the four-factor model identified by Radloff (1977) and also, using only 12 items, found support for a three-factor (i.e. somatic complaints, depressed affect, and positive affect) model across three generations of Mexican Americans. Using the Hispanic NHANES data, Guarnaccia et al. (1989) identified a slightly different three-factor model, with somatic complaints and depressive affect factors combined as one factor rather than two distinct factors. These authors suggested that conceptual distinctions between somatic complaints and depressive affect might not exist for Mexican Americans, Cuban Americans, and Puerto Ricans. Rather, Hispanic groups tend to use both somatic complaints and depressive affect to communicate their depressive feelings and psychological distress. Using a different data set, Miller et al. (1997) found a more parsimonious two-factor model provided a better fit to their data on elderly Mexican Americans.

Among studies examining the factor structure of the CES-D in African Americans, Blazer et al. (1998) reported evidence of Radloff's four-factor structure for older whites and African Americans. One of the items ('fearful'), however, did not significantly load on any factor and was eventually excluded from the analysis. Foley et al. (2002) generally did not confirm Radloff's four-factor model in older African Americans. Using exploratory factor analysis, they found no distinction between somatic complaints and depressed affect, and they identified one new factor, 'social well-being', that has not been reported in the general population.

Table 1
Description of study groups

	HANDLS $n=428$	NHEFS $n=988$	BLSA $n=666$
Sampling design	Convenience	Stratified probability	Convenience
Date of data	2000–2001	1981 wave	1995–2002
Male (%)	188 (44%)	294 (30%)	256 (38%)
Age range (mean)	18–92 (50)	33–86 (53)	19–92 (53)
Education (mean)	12 years	9–12 years*	16 years

* Education was coded as a categorical variable in the original data set; the median is given here.

While previous studies provide evidence for a four-factor model of the CES-D in some minority groups, the evidence they offer is far from conclusive. Mindful of these considerations, the present study replicates and attempts to confirm the work begun by Radloff (1977) in determining the factor structure of the CES-D in low socioeconomic status (SES) African Americans. That is, can the four-factor structure of the CES-D scale be replicated across groups with different characteristics such as race and socioeconomic status, or sampling designs including samples of convenience and stratified probability samples? We examined the factor structure of the CES-D in the Healthy Aging in Nationally Diverse Longitudinal Samples study (HANDLS) of low SES African Americans, a sample of convenience. We additionally tested the stability of the measurement model in other samples of African-American participants from the National Health and Nutrition Examination Survey I Epidemiologic Follow-up Study (NHEFS) and Caucasians from the Baltimore Longitudinal Study of Aging (BLSA). Since research on the measurement structure of the CES-D has not been consistently pursued in minority populations, cross-validation of the model was considered necessary to (a) rule out the possibility of random and measurement error; and (b) validate scores based on those structures that are used in assessing depressive symptoms in African Americans.

2. Methods

2.1. Study participants

The sample for this study includes participants from the HANDLS, NHANES, and BLSA studies. The HANDLS pilot project was a sample of convenience examining health disparities among low SES status African-American men and women living in Baltimore. Participants were 426 African-American men and women between the ages of 18 and 92 years (see Table 1 for group characteristics). Among the sample, 60% of men and 40% women were current or previous drug users. Participants were paid. Between October 2000 and December 2001, study participants received a variety of medical and psychological tests, including an assessment of depressive symptoms. All participants were given the CES-D scale as part of the larger HANDLS project to study depressive symptoms. The CES-D scale was self-administered except in situations where low literacy skills interfered with self-administration. For participants with low literacy skills, trained psychometric technicians read the CES-D questions and recorded participant responses.

For purposes of cross-validation, other groups of African Americans and Caucasians with complete CES-D scores were extracted from the NHEFS and BLSA studies. The NHEFS data subset included 988 African Americans, and these subjects were matched against HANDLS participants on factors including gender, age, and education level (within 5 years) to avoid possible differences due to sample selection biases. Medical risk factor and psychological data including the CES-D scale were collected as part of the NHEFS study, a stratified probability survey of the adult, non-institutionalized, civilian population of the United States. The sampling design of the NHEFS study ensured that the subjects who participated were stratified probability samples of the US pop-

ulation. Consequently, participants who completed the CES-D scale were a representative sample of the United States. A full description of the rationale, methods, and subject characteristics of this study can be found elsewhere (Cornoni-Huntley et al., 1983). The Baltimore Longitudinal Study of Aging (BLSA) is a longitudinal study conducted by the Intramural Research Program of the National Institute on Aging (NIA) (Shock et al., 1984). Participants are community-dwelling volunteers who return approximately every 2 years to the Gerontology Research Center of the NIA for comprehensive medical, physiological, and neuropsychological evaluations. For this study, we extracted 666 Caucasian participants who completed every item on the CES-D scale between the years 1995 and 2002. The cohort consists of predominantly upper-middle class, white professionals with mean educational levels that are higher than participants from either the HANDLS or NHEFS surveys (see Table 1 for a description of study groups).

2.2. Measure

The CES-D scale is a 20-item inventory developed by the National Institute of Mental Health Center for Epidemiological Studies to assess the frequency and severity of depressive symptoms in the past week (Radloff, 1977). The inventory has been extensively validated (Lewinsohn et al., 1988) and is widely accepted in epidemiologic studies of depression in general populations. The CES-D scale correlates strongly with other self-reported depression inventories and it correlates with variables related closely to clinical diagnoses of depression. Scores for clinically depressed patients are much higher than those for normal subjects (Weissman et al., 1977), and a standard cutoff score of 16 has been defined as indicating depressive symptoms. This score identifies a large proportion of individuals with major depressive disorders, but largely in non-minority samples (Eaton and Kessler, 1981). Radloff (1977) and others (Weissman et al., 1977; Hertzog et al., 1990; Knight et al., 1997) identified four different factors of the CES-D scale in the general population. The factors were labeled (1) depressed affect, (2) positive affect, (3) somatic complaints, and (4) interpersonal problems. Table 2 lists the items associated with the four factors.

Table 2
CES-D scale items and abbreviation

Item	Abbreviation
Factor 1: Somatic complaints	
1. I was bothered by things that usually do not bother me.	Bothered
2. I did not feel like eating; my appetite was poor.	Appetite
3. I had trouble keeping my mind on what I was doing.	Mind
4. I felt that everything I did was an effort.	Effort
5. My sleep was restless.	Sleep
6. I talked less than usual.	Talk
7. I could not get going.	Going
Factor 2: Depressive affect	
8. I felt that I could not shake off the blues even with help from my family or friends.	Blues
9. I felt depressed.	Depressed
10. I felt lonely.	Lonely
11. I had crying spells.	Crying
12. I felt sad.	Sad
13. I thought my life had been a failure.	Failure
14. I felt fearful.	Fearful
Factor 3: Positive affect	
15. I felt that I was just as good as other people.	Good
16. I felt hopeful about the future.	Hopeful
17. I was happy.	Happy
18. I enjoyed life.	Enjoyed
Factor 4: Interpersonal problems	
19. People were unfriendly.	Unfriendly
20. I felt that people disliked me.	Dislike

2.3. Analysis

Approximately 0.02% of the HANDLS participants had one or two CES-D items missing. Because of the small percentage, cohort by gender by education means were substituted for missing data within the group. We used confirmatory factor analysis to examine the factor structure of the CES-D scale across the HANDLS, NHEFS, and BLSA data sets. Restricted factor analysis using LISREL 8.52 (Jöreskog and Sorböm, 2001) software was employed to replicate the CES-D factor structure described by Radloff (1977) and previous research (Hertzog et al., 1990; Knight et al., 1997)

(see Table 2 for factor descriptions). A widely used method of covariance structure analysis, LISREL, can be helpful in the evaluation and selection of models since it is capable of testing simultaneous multiple covariance matrices by imposing equality constraints across groups. LISREL can estimate a set of models that specify the relationship of observed variables to underlying latent variables presumed to account for them. At the same time, it separates imperfect model specification and measurement errors from reliable variance and structural constraints imposed by the model (Byrne, 1998). We also used Jöreskog and Sorböm's (2001) PRELIS program to convert data into polychoric correlation matrices, with asymptotic covariance matrices to adjust for multivariate non-normality to be analyzed in LISREL.

In addition, tests of structural invariance often represent different types of invariance and constitute a hierarchy. For example, sets of parameters are usually tested in an orderly sequence of logical steps and in an increasingly restrictive manner. For the present study, the following questions were asked: (1) Is the number of factors the same across groups? (2) Are the factor loading coefficients equivalent across groups? (3) Are tests of error variances/covariances equivalent across groups? (4) Are factor variances/covariances equivalent across groups? The process of determining equivalence of measurement and structural parameters was terminated at the level where non-equivalence occurred. For example, if the number of factors was the same across groups, then model testing proceeded to constraining factor loadings across groups. If non-equivalencies in factor loadings were found, particularly if they represented different groups, then error terms and factor variances/covariances were also expected to differ between groups (Bentler, 1992; Brown et al., 1996; Byrne, 1998).

To examine the extent to which our hypothesized confirmatory model fit, the following fit indices were used to evaluate model fit in the current study. The chi-square (χ^2) statistic was used to test the closeness of fit between the unrestricted sample covariance matrix and the restricted covariance matrix. Due to its sensitivity to sample size, non-normality and model complexity, the χ^2 test may not have been realistic to use as an index of fit. However, probability values associated with χ^2 tests are reported in the current study to assess the extent to which a respecified model exhibits improvement in fit. To evaluate model improvement, the difference in χ^2 ($\Delta\chi^2$) between two nested models was examined (Bollen, 1989; Byrne, 1998). Other fit indices including the RMSEA (root mean square error of approximation), GFI (goodness-of-fit index), and CFI (comparative fit index) are also reported. The RMSEA takes into account the error of approximation in the population, and has recently been recognized as one of the most informative criteria in covariance structure modeling. RMSEA values less than 0.05 indicate a good fit, 0.06–0.08 represent reasonable fit, 0.08–0.10 indicate mediocre fit, and those above 0.10 indicate poor fit. GFI compares the postulated model with no model at all. GFI values range from zero to 1.00, with values closer to 1.00 representing a good fit. CFI is independent of sample size, and takes the complexity of the model into account and compares the hypothesized model with the independence model. The upper limit of CFI is 1.00, with a value >0.90 indicating a good fit to the data (for complete reference to these fit indices, see Bentler, 1992; Browne and Cudeck, 1993; Byrne, 1998).

3. Results

Prior to testing for invariance across the three groups, confirmatory factor analysis was conducted to establish baseline models for the HANDLS group. The baseline model was specified with the four factors (see Table 2) as described by Radloff (1977) and previous studies (Hertzog et al., 1990; Knight et al., 1997). All indicators were found to have significant proportions of variance or significant loadings, and the fit statistics for this baseline model were consistent with one that fit well ($\chi^2=336.88$, d.f.$=164$, RMSEA$=0.05$, GFI$=0.97$, CFI$=0.95$). The accepted baseline model was then extended to the NHEFS and BLSA groups for purposes of cross-validation. The first test of invariance assessed whether the CES-D was best described by a four-factor structure for all three

groups. At this level, no equality constraints were specified across groups, and invariance was modeled to test for an equivalent number of factors that best represented the data for the HANDLS, NHEFS, and BLSA groups. All indicators were found to have significant loadings. The magnitudes of the fit indices were impressive, given the complexity of the model and as indicated by the fit indices ($\chi^2 = 1150.66$, d.f. = 492, RMSEA = 0.04, GFI = 0.97, CFI = 0.97). These results were considered optimal in representing the data for all three groups. Table 3 lists loading descriptions and factor correlations for the three groups. In line with previous research that found no distinction between somatic and depressive symptoms, factor correlations between somatic complaints and depressive affect appeared to be high across groups in the current study. As such, a second model was derived that combined these two factors as one. This three-factor model was then tested across groups. A decrease in model fit was indicated by an increase chi-square values ($\Delta\chi^2 = 85.03$, Δd.f. = 9, $P < 0.005$), though other fit indices remained relatively the same (RMSEA = 0.046, GFI = 0.96, CFI = 0.96). In addition, the unstandardized loadings substantially departed from the loadings obtained under the four-factor model and differentially shifted among the three groups. For example, items including 'lonely', 'sad', 'crying', and 'failure' became less related to its own factor for the NHEFS group, but the opposite was observed for the BLSA group. Furthermore, the magnitude of several error terms for items including 'blues', 'depressed', and 'enjoyed' decreased substantially and became non-significant, which could indicate that these residuals were being overestimated (Byrne, 1998; Jöreskog and Sorböm, 2001). Based on these findings, the CES-D scale was not well represented by the hypothesized three-factor model and was best described by a four-factor model for all three groups.

At a minimum, our results indicated that the number of factors and the pattern of factor loadings were equivalent across all three groups. At this level of testing, however, our model provided a less constrained test and required that simple structure of the factor patterns remained the same across groups. In seeking further evidence of measurement equivalency of the CES-D scale in samples with differential characteristics, we tested to see if the magnitudes of the factor loadings were equal across HANDLS, NHEFS, and BLSA groups. In testing this hypothesis, the four-factor model was re-specified with equality constraints placed for all three groups on all factor loadings. Results indicated a statistically significant reduction in fit for this model ($\Delta\chi^2 = 366.94$, Δd.f. = 32, $P < 0.01$). Further analyses were undertaken to identify parameters causing differences. To approximately gauge the differences in the parameter estimates, confidence intervals around the individual factor loadings were examined to see which specific loadings were significantly different between groups (i.e. HANDLS vs. NHEFS, HANDLS vs. BLSA, NHEFS vs. BLSA). Higher loading differences were significantly evident between African-American and Caucasian groups, while loading similarities were significantly more tenable between the two African-American groups. Table 4 presents loadings that are similar or different between groups. Since the magnitude of factor loadings was not the same across groups, model testings were terminated at this level and no further invariance tests, including error variances and factor variances/covariances, were conducted.

4. Discussion

The present study demonstrated that the same four-factor model found in Caucasian populations could be replicated in the low SES African-American HANDLS sample. A strength of this study is the ability to examine the hypothesized factor model and validate such results on other independent samples that were different in sampling design, race, SES, and time of measurement. An adequate fit for a four-factor factor model was shown to hold not only across the three groups, but most items were reliable with standardized loadings greater than 0.50. Results indicate that the items are related to the hypothesized factors and provide the interpretation of the CES-D scale in terms of four symptom dimensions including somatic complaints, depressed affect, positive affect, and interpersonal problems. Since the findings are robust in each group and therefore suggest similar inter-

Table 3
Unstandardized solutions for CES-D items in HANDLS, NHEFS, and BLSA

Item	Factor 1			Factor 2			Factor 3			Factor 4		
	Somatic complaints			Depressive affect			Positive affect			Interpersonal problems		
	HANDLS	NHEFS	BLSA	HANDLS	NHEFS	BLSA	HANDLS	NHEFS	BLSA	HANDLS	NHEFS	BLSA
Loadings												
1. Bothered*	1.00	1.00	1.00									
2. Appetite	0.86	1.03	1.03									
3. Mind	0.91	1.13	0.99									
4. Effort	0.60	0.99	1.32									
5. Sleep	1.01	1.07	0.75									
6. Talk	0.84	0.98	0.89									
7. Going	1.24	1.13	1.27									
8. Blues*				1.00	1.00	1.00						
9. Depressed				0.94	1.07	1.03						
10. Lonely				0.98	1.03	0.88						
11. Crying				0.97	1.04	0.82						
12. Sad				1.05	1.06	0.98						
13. Failure				0.92	0.98	0.97						
14. Fearful				0.94	0.98	0.83						
15. Good*							1.00	1.00	1.00			
16. Hopeful							1.04	0.94	1.22			
17. Happy							1.30	1.20	1.35			
18. Enjoyed							1.22	1.20	1.40			
19. Dislike*										1.00	1.00	1.00
20. Unfriendly										1.34	1.15	1.17
Interfactor correlations												
Factor 1	1.00	1.00	1.00									
Factor 2	0.61	0.61	0.62	1.00	1.00	1.00						
Factor 3	0.36	0.20	0.37	0.48	0.31	0.56	1.00	1.00	1.00			
Factor 4	0.41	0.48	0.42	0.55	0.64	0.62	0.34	0.23	0.42	1.00	1.00	1.00

* These items were fixed to 1.00 for model identifying purposes (Jöreskog and Sorböm, 2001).

Table 4
Equality of item loadings between groups: HANDLS vs. NHEFS, HANDLS vs. BLSA, and NHEFS vs. BLSA

Item	HANDLS vs. NHEFS		HANDLS vs. BLSA		NHEFS vs. BLSA	
F1: Somatic complaints						
1. Bothered*	1.00	1.00	1.00	1.00	1.00	1.00
2. Appetite	0.85	1.01	0.87	1.02	(1.03	1.03)
3. Mind	0.88	1.12	(0.96	0.96)	1.10	1.01
4. Effort	0.57	0.99	0.61	1.30	0.97	1.33
5. Sleep	(1.04	1.04)	1.01	0.74	1.05	0.75
6. Talk	0.84	0.96	(0.87	0.87)	(0.93	0.93)
7. Going	(1.15	1.15)	(1.26	1.26)	1.10	1.29
F2: Depressive affect						
8. Blues*	1.00	1.00	1.00	1.00	1.00	1.00
9. Depressed	0.95	1.06	0.95	1.02	(1.05	1.05)
10. Lonely	(1.01	1.01)	0.99	0.87	1.01	0.89
11. Crying	(1.01	1.01)	0.99	0.81	1.02	0.83
12. Sad	(1.06	1.06)	1.06	0.97	1.04	0.99
13. Failure	(0.96	0.96)	(0.96	0.96)	(0.98	0.98)
14. Fearful	(0.97	0.97)	0.96	0.83	0.96	0.84
F3: Positive affect						
15. Good*	1.00	1.00	1.00	1.00	1.00	1.00
16. Hopeful	(0.96	0.96	(1.16	1.16)	0.99	1.15
17. Happy	(1.23	1.23)	(1.34	1.34)	(1.28	1.28)
18. Enjoyed	(1.20	1.20)	(1.35	1.35)	(1.32	1.32)
F4: Interpersonal problems						
19. Dislike*	1.00	1.00	1.00	1.00	1.00	1.00
20. Unfriendly	1.36	1.14	1.36	1.17	(1.17	1.17)

Note: items in brackets indicate loading similarities between groups.

* These items were fixed to 1.00 for model identifying purposes (Jöreskog and Sorböm, 2001).

pretations across groups, the latent variable model of CES-D would allow the examination of symptom domain differences according to self-report ethnicity and SES. For example, Baker et al. (1996) found somatic complaints more prominent than reported mood changes in detecting the presence of depression in African Americans. Thus, the distribution and prevalence of somatization symptoms may be a stronger marker of depression than other symptom clusters in assessing depression among African Americans (Baker et al., 1996). Furthermore, an estimation of latent variable scores could be useful in providing treatment information, since symptom clusters may differentially respond to treatment among different ethnic groups (Angel and Thoits, 1987; McBarnette, 1996). Hence, the dimensions of the CES-D may be more informative than a total CES-D score in assessing depressive symptoms across different ethnic groups.

To compare the CES-D on a factorial level, at a minimum, the validity and feasibility of the four-factor model must be performed and supported across groups. Findings in the current study support the validity of the four-factor structure upon which the CES-D was developed and therefore suggest similar interpretations across the three groups. Alternatively, if the four-factor structure had varied between whites and African Americans, then group comparisons on these four symptom domains would not be meaningful as results could be due to measurement artifacts. Under such circumstance, the four-factor model would no longer be analyzed as a confirmatory model and subsequently would need to be respecified and reestimated using the exploratory route to identity an alternative structural model that would best describe the CES-D for groups under study (Byrne, 1998; Jöreskog and Sorböm, 2001).

In line with previous studies, our results seem to indicate a lack of conceptual distinctions between somatic complaints and depressive affect. When analyzed as a three-factor model with somatic complaints and depressive affect factors combined, however, the model was rejected due to the significant reduction in fit. There are a couple of reasons why the two factors are related. From a socio-cultural perspective, research has suggested that some cultures may somaticize and describe depression as a physiologic phenomenon with somatic symptoms rather than affective symptoms (Angel and Thoits, 1987; Dick et al., 1994; Miller et al., 1997). The higher correlation between the two factors could also indicate a stronger co-dependent relationship compared with other factors. Individuals with high levels of somatic complaints are much more likely to experience higher levels of depressive affect than other domains such as interpersonal problems or positive affect. However, from a conceptual perspective of factor invariance, correlations among factors should be minor in order to satisfy evidence of latent factor unique attributes (as indicated by large factor loadings). Such findings are highly

unlikely with respect to psychological data in particular. Yet large factor correlations are possibly a result of models that either specify too many latent variables or not enough observed variables (Maitland, 1997; Byrne, 1998). This hypothesis remains to be tested in our study.

Results obtained from the analyses provided support for at least invariance of factor structures for all groups. That is, all groups demonstrate equivalence in the pattern of factor structure, indicating that there is a reliable relationship between the factors and the variables by which the factors are measured. Our analyses, however, generally did not support invariance of factor loadings across groups. This finding may imply that all variables constituting the four selected factors are not measuring the same CES-D facets in exactly the same way for the three groups. Notably, a higher number of loading similarities were found between the two African-American groups. This finding was expected given the same ethnicity and cultural backgrounds. The lack of loading invariance across groups was also expected given what has been suggested in the factor invariance literature. That is, the same number of factors is often supported across groups while other levels of invariance including loading invariance are more difficult to establish in the social science domain. Rank order of salient loadings and factor variances/covariances and uniqueness are generally expected to vary from one group to another. In most instances, particularly those with significant sample differentiation, the same number of factors is regarded to be more important than factor loading invariance. This is sufficient in most cases (Thurstone, 1947; Horn et al., 1983; Cunningham, 1991; Maitland, 1997).

In considering the implications of the findings in the current study, it is essential to take into account limitations of the data and the analysis. Though the study groups differed by race and socioeconomic status, our analyses were not framed within an exploratory mode. This route may provide potential factor structures that are uniquely associated with each specific group. Because the CES-D has been fully developed and its factor structures extensively validated in the general population, we elected to follow the confirmatory factor analysis route to further test for its validity given the unique characteristics of the HANDLS group. Based on findings of adequate fit and no indication of model misspecification, in addition to consistent results obtained from cross-validation, the hypothesized four-factor structure of the CES-D in low SES African Americans was not rejected. Given that the estimation of the targeted parameters was substantively meaningful, we did not continue to explore other alternative factorial structures.

The current study also did not seek to determine whether our findings held equally across gender. For example, in a study of the CES-D factor structure among races and gender, Callahan and Wolinsky (1994) found gender differences regarding item loadings in African Americans. Two items including 'mind' and 'failure' were found to have nonsignificant loadings in African-American women. Given the sample size in the HANDLS, such sample compositions would result in small cell sizes. Yet sample size is a crucial factor in determining the degree to which existing model evaluation procedures can be viewed with confidence (Byrne, 1998). Model testings across gender in this study may not have enough power to detect the differences between several competing models using the $\Delta\chi^2$ value statistic for model selection or evaluation.

In addition, our study lacks longitudinal data that would allow the examination of the CES-D factor structure across time in African Americans. Data in the current study were collected using different time periods, and results were based on cross-sectional data. Longitudinal data would provide information on the stability of the relationship between the factors and the variables by which the factors were measured over time. The demonstration of factorial invariance is important in representing valid within-group changes and reliable change processes over time. This area of research is especially limited in minority populations including African Americans. However, a full-scale population study of approximately 3000 participants is currently being collected, and follow-up data of the original 426 HANDLS participants could be used eventually to validate the results discussed in the present study. These limitations

notwithstanding, there is much to be learned from these analyses that will contribute significantly to the literature on the measurement validity of the CES-D factor structure in African Americans, particularly among individuals from low SES areas.

References

Angel, R., Thoits, P., 1987. The impact of culture on the cognitive structures of illness. Culture, Medicine and Psychiatry 11, 465–494.

Baker, F.M., Okwumabua, J., Philipose, V., Wong, S., 1996. Screening African-American elderly for the presence of depressive symptoms: a preliminary investigation. Journal of Geriatric Psychiatry and Neurology 9 (3), 127–132.

Bentler, P.M., 1992. On the fit of models to covariances and methodology to the Bulletin. Psychological Bulletin 112, 400–404.

Berkman, L.F., Berkman, C.S., Kasl, S., Freeman Jr, D.H., Leo, L., Ostfeld, A.M., Cornoni-Huntley, J., Brody, J.A., 1986. Depressive symptoms in relation to physical health and functioning in the elderly. American Journal of Epidemiology 124 (3), 372–388.

Blazer, D., Burchett, B., Service, C., George, L.K., 1991. The association of age and depression among the elderly: an epidemiologic exploration. Journal of Gerontology: Medical Sciences 46 (6), M210–M215.

Blazer, D.G., Landerman, L.R., Hays, J.C., Simonsick, E.M., Saunders, W.B., 1998. Symptoms of depression among community-dwelling elderly African–American and white older adults. Psychological Medicine 28 (6), M210–M215.

Bollen, K.A., 1989. Structural Equations with Latent Variables. Wiley, New York.

Brown, C., Schulberg, H.C., Madonia, M.J., 1996. Clinical presentations of major depression by African Americans and whites in primary medical care practice. Journal of Affective Disorders 41 (3), 181–191.

Browne, M.W., Cudeck, R., 1993. Alternative ways of assessing model fit. In: Bollen, K.A., Long, J.S. (Eds.), Testing Structural Equation Models. Sage, Newbury Park, CA, pp. 445–455.

Byrne, B.M., 1998. Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS. Lawrence Erlbaum Associates, Mahwah, NJ.

Callahan, C.M., Wolinsky, F.D., 1994. The effect of gender and race on the measurement properties of the CES-D in older adults. Medical Care 32 (4), 341–356.

Chapleski, E.E., Lamphere, J.K., Kaczynski, R., Lichtenberg, P.A., Dwyer, J.W., 1997. Structure of a depression measure among American Indian elders: confirmatory factor analysis of the CES-D Scale. Research on Aging 19, 462–485.

Cornoni-Huntley, J., Barbano, H.E., Brody, J.A., Cohen, B., Feldman, J.J., Kleinman, J.C., Madans, J., 1983. National health and nutrition examination I—epidemiologic follow-up survey. Public Health Reports 98 (3), 245–251.

Cunningham, W., 1991. Issues in factorial invariance. In: Collins, L., Horn, J.L. (Eds.), Best Methods for the Analysis of Change: Recent Advances, Unanswered Questions, Future Directions. American Psychological Association, Washington, DC, pp. 106–113.

Dick, R.W., Beals, J., Keane, K.M., Manson, S.M., 1994. Factorial structure of CES-D among American Indian Adolescents. Journal of Adolescence 17, 73–79.

Eaton, W.W., Kessler, L.G., 1981. Rates of symptoms of depression in a national sample. American Journal of Epidemiology 114 (4), 528–538.

Foley, K.L., Reed, P.S., Mutran, E.J., DeVellis, R.F., 2002. Measurement adequacy of the CES-D among a sample of older African-Americans. Psychiatry Research 109 (1), 61–69.

Guarnaccia, P.J., Angel, R., Worobey, J.L., 1989. The factor structure of the CES-D in the Hispanic Health and Nutrition Examination Survey: the influences of ethnicity, gender, and language. Social Science and Medicine 29 (1), 85–94.

Hertzog, C., Van Alstine, J., Usala, P., Hultsch, D.F., Dixon, R., 1990. Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. Psychological Assessment 2, 64–72.

Horn, J.L., McArdle, J.J., Mason, R., 1983. When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. Southern Psychologist 1, 179–188.

Jones-Webb, R.J., Snowden, L.R., 1993. Symptoms of depression among blacks and whites. American Journal of Public Health 83 (2), 240–244.

Jöreskog, K.G., Sorböm, D., 2001. LISREL 8: User's Reference Guide. Scientific International, Chicago.

Knight, R.G., Williams, S., McGee, R., Olaman, S., 1997. Psychometric properties of the Center for Epidemiological Studies Depression Scale (CES-D) in a sample of women in middle life. Behavior Research and Therapy 35, 373–380.

Lewinsohn, P.M., Hoberman, H.M., Rosenbaum, M., 1988. A prospective study of risk factors for unipolar depression. Journal of Abnormal Psychology 97 (3), 251–264.

Liang, J., Tran, T.V., Krause, N., Markides, K.S., 1989. Generational differences in the structure of the CES-D scale in Mexican Americans. Journal of Gerontology 44, 110–120.

Maitland, S. 1997. Factorial invariance and concordance of health behaviors and health status: a study of individual differences in familial context. Unpublished Dissertation, Pennsylvania State University, University Park, PA.

McBarnette, L.S., 1996. African American women. In: Bayne-Smith, M. (Ed.), Race, Gender, and Health. Sage Publications, Newbury Park, CA, pp. 43–67.

Miller, T.Q., Markides, K.S., Black, S.A., 1997. The factor structure of the CES-D in two surveys of elderly Mexican Americans. Journal of Gerontology: Social Sciences 52B (5), S259–S269.

Murrell, S.A., Himmelfarb, S., Wright, K., 1983. Prevalence of depression and its correlates in older adults. American Journal of Epidemiology 117 (2), 173–185.

Pearlin, L.I., 1989. The sociological study of stress. Journal of Health and Social Behavior 30, 241–256.

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

Rhee, S.H., Petroski, G.F., Parker, J.C., Smarr, K.L., Wright, G.E., Multon, K.D., Buchholz, J.L., Komatireddy, G.R., 1999. A confirmatory factor analysis of the Center for Epidemiologic Studies Depression Scale in rheumatoid arthritis patients: additional evidence for a four-factor model. Arthritis Care Research 12 (6), 392–400.

Roberts, R.E., 1980. Reliability of the CES-D Scale in different ethnic contexts. Psychiatry Research 2 (2), 125–134.

Shock, N.W., Greulich, R.C., Andres, R., Arenberg, D., Costa, P.T., Lakatta, E.G., Tobin, J.D., 1984. Normal Human Aging: The Baltimore Longitudinal Study of Aging, NIH Publication No. 84-2450. U.S. Government Printing Office, Washington, DC.

Thurstone, L.L., 1947. Multiple Factor Analysis. University of Chicago Press, Chicago.

Weissman, M.M., Sholamskas, D., Pottenger, M., Prusoff, B.A., Locke, B.Z., 1977. Assessing depressive symptoms in five psychiatric populations: a validation study. American Journal of Epidemiology 106, 203–214.