

Factor structure and longitudinal invariance of the Center for Epidemiological Studies Depression Scale (CES-D) in adult women: application in a population-based sample of mothers of children with epilepsy

Mark A. Ferro · Kathy N. Speechley

Received: 18 August 2012 / Accepted: 4 February 2013 / Published online: 19 February 2013
© Springer-Verlag Wien 2013

Abstract The objective of this study was to examine the factor structure and longitudinal measurement invariance of the Center for Epidemiological Studies Depression Scale (CES-D). The population-based sample included 347 adult women who had children participating in the Health-related Quality of Life in Children with Epilepsy Study. Longitudinal confirmatory factor analysis was used to confirm the factor structure and test for measurement invariance. The original CES-D four-factor model (depressed affect, positive affect, somatic activity, and interpersonal relations) provided the best

fit to the data compared to alternate models: [$\chi^2=362.95$ ($df=164$); Comparative Fit Index (CFI)=965; Tucker-Lewis Index (TLI)=0.960; Weighted Root Mean Square Residual (WRMR)=0.966; Root Mean Square Error of Approximation (RMSEA)=0.059, 90 % confidence interval (CI), 0.051, 0.068] and was used for invariance testing. Results suggested full measurement invariance over time with the final model demonstrating adequate fit [$\chi^2=2303.92$ ($df=1,776$); CFI=0.957; TLI=0.952; WRMR=1.149; RMSEA=0.033, 90 % CI, 0.030, 0.038]. We conclude that the four-factor structure of the CES-D is supported for adult women and that the measurement of the scale is longitudinally invariant. Clinicians and researchers can be confident that longitudinal changes in CES-D scores reflect true changes in depressive symptomatology and are not an artifact of changes in the interpretation of items in the scale over time.

M. A. Ferro
Department of Psychiatry and Behavioural Neurosciences,
McMaster University, Hamilton, Canada

M. A. Ferro
Offord Centre for Child Studies, McMaster University,
Hamilton, Canada

K. N. Speechley
Department of Paediatrics, Western University,
Commissioners Road East, Room E6-103,
London, ON N6A 5W9, Canada

K. N. Speechley
Department of Epidemiology and Biostatistics,
Western University, London, Canada

K. N. Speechley
Children's Health Research Institute, Western University,
London, Canada

M. A. Ferro (✉)
Department of Psychiatry and Behavioural Neurosciences,
Offord Centre for Child Studies, McMaster University,
Chedoke Site, Central Building, Room 310,
1280 Main Street West,
Hamilton, ON L8S 4K1, Canada
e-mail: ferroma@mcmaster.ca

Keywords Age · Depression · Measurement · Psychometrics

Introduction

Caring for a child recently diagnosed with a chronic health condition can be the source of tremendous stress for parents, particularly mothers, who most often assume the primary caregiving role in families (Gillespie and Primavera 2000). As a result, mothers may have an elevated risk for psychological distress, including depression, in response to their child's diagnosis (Dolgin et al. 2007; Ferro et al. 2011b). Understanding maternal depression has important clinical implications—not only are symptoms of depression unhealthy for mothers, research has provided evidence showing that depression in mothers has a significant negative effect on children's health (Beardslee et al. 1998; Kim-Cohen et al. 2005; Pilowsky et al.

2006; Goodman et al. 2011). In pediatric epilepsy, maternal depression has been linked to internalizing and externalizing behavior problems in children, particularly symptoms of anxiety and depression, and poor health-related quality of life (Ferro and Speechley 2009).

In epidemiological and clinical research, symptoms of depression are routinely measured using screening instruments that aim to capture various related constructs underlying the multidimensionality of depression. One of the most commonly used screening instruments is the Center for Epidemiological Studies Depression Scale (CES-D; Radloff 1977), which has received over 20,000 citations since its publication. The CES-D is a 20-item self-report instrument that measures depressive symptoms across four domains: depressed affect, positive affect, somatic activity, and interpersonal relations. This four-factor structure of the CES-D has been replicated in several different study populations (Hertzog et al. 1990; Knight et al. 1997; Williams et al. 2007; Edwards et al. 2010; Morin et al. 2011). Furthermore, a recent meta-analysis of studies confirmed the four-factor structure as robust, despite slight variations in the items comprising each factor and magnitude of factor loadings (Shafer 2006). While the factor structure of the CES-D has been well established, concerns regarding the validity and reliability of this instrument in different groups and over time have been expressed. Despite evidence of response heterogeneity within and between populations, the majority of studies that have examined differences in depression scores using the CES-D, either between groups or over time, have taken measurement invariance for granted (Muthén 1989). This is problematic since violation of the assumption of invariance may result in measuring depression using indicators that are not conceptualized similarly by respondents over time. Thus, comparisons over time can be biased.

Measurement invariance concerns the extent to which the psychometric properties of the observed indicators are generalizable across groups or over time (Vandenberg and Lance 2000; Brown 2006); it tells researchers that the same construct is being measured in the same way in different groups (e.g., males and females) or over time (i.e., repeated measures on the same individuals). Demonstrating measurement invariance makes it possible to interpret differences between groups or over time as real and meaningful. The test for these differences is called structural invariance, and it involves comparing model fit before and after constraining the unobserved factor means to be equal. Without the evaluation of invariance over time, researchers cannot be certain whether observed changes over time reflect true change or if it is the result of changes in the interpretation of items or structure of the construct (Brown 2006).

Previous cross-sectional studies examining measurement invariance between subgroups of women have been mixed. Recently, Van Lieshout et al. (2011) reported partial

measurement invariance of the CES-D comparing immigrant and non-immigrant women in the post-partum period. In contrast, Williams et al. (2007) confirmed the four-factor structure of the CES-D, but found that the measure was non-invariant among younger and older African American women. Finally, Knight et al. (1997) replicated the original factor structure of the CES-D, but did not assess measurement invariance.

We are aware of only two studies that have examined the longitudinal invariance of the CES-D; one sampled children and the other included mostly adult male cardiac patients (Motl et al. 2005; Contrada et al. 2006). In the study by Motl et al. (2005), support was found for the original four-factor structure of the CES-D. In addition, full longitudinal measurement invariance was observed using 2-year follow-up data of children during their transition between grade seven and eight. In contrast, Contrada et al. (2006) did not replicate the original four-factor structure of the CES-D, but instead demonstrated longitudinal invariance of a three-factor solution with modifications among a predominately male sample of patients undergoing cardiac surgery. However, it is unknown if these findings can be replicated in a sample of adult women. In particular, there are no studies which have examined the factor structure or measurement invariance of depression scales in mothers of children with a chronic illness. Assessment of longitudinal invariance of the CES-D is important to ensure that changes in symptoms of depression are true changes and not simply an artifact of changes in the interpretation of scale items over time. This is particularly salient given that previous evidence has shown that depressive symptoms in mothers have a negative impact on the health-related quality of life of children with epilepsy (Ferro et al. 2011a). Thus, given the paucity of research surrounding longitudinal invariance of the CES-D, the objectives of this research were to (1) confirm the factor structure and (2) assess the longitudinal measurement invariance of the CES-D over a 24-month period in a sample of adult women.

Methods

Sample

Data for this study came from the Health-related Quality of Life of Children with Epilepsy Study (HERQULES), a prospective cohort study designed to examine the determinants of health-related quality of life in children with epilepsy during the first 24 months post-diagnosis. Families were recruited from pediatric neurology practices and consisted of a sample of English-speaking families with a child diagnosed with epilepsy between 4 and 12 years of age. Participants were recruited over a 36-month period by pediatric neurologists ($n=52$) and provided written consent prior to being enrolled in the study. Primary

caregivers were contacted by telephone to determine participation status and mailed self-administered questionnaires after diagnosis (baseline), and at 6, 12, and 24 months. For this analysis, data at baseline, 12, and 24 months were utilized due to the equal spacing of measurement occasions. Responses from mothers who completed at least one measurement occasion were retained, resulting in a sample size of $N=347$ at baseline. A total of $N=258$ (76 %) completed all measurements' occasions. Approval for HERQULES was obtained from all relevant research ethics boards across the country.

Measurement

The CES-D was included in HERQULES. The CES-D is a 20-item questionnaire designed to assess depressive symptoms in the general adult population over the past week (Radloff 1977). The scale includes 20 items that assesses depressed affect (7 items), positive affect (4 items), somatic activity (7 items), and interpersonal relations (2 items). A four-point adjectival response scale (0–3) is used with anchor points for frequency of symptoms experienced ranging from “rarely or none of the time (less than 1 day)” to “most or all of the time (5–7 days).” The total score ranges from 0 to 60, with a higher score indicating greater impairment. Individuals who score <8 are considered to have clinically irrelevant symptoms, whereas individuals in the 8–15 range are defined as having subclinical depressive symptoms, and those scoring ≥ 16 are typically identified as having clinically relevant levels of depressive symptoms (Radloff 1977; Cohen et al. 2010).

Procedure and statistical analysis

Confirmatory factor analysis was used to examine longitudinal measurement invariance of the CES-D. In this approach, a hierarchical set of procedures specifying increasingly stringent equality constraints was employed to assess longitudinal invariance (Millsap and Yun-Tein 2004). Tests of longitudinal measurement invariance were as follows: (1) configural invariance imposes no equality constraints on parameters and is used as the basis for more complex models to be tested (Byrne 2012); (2) metric invariance examines the extent to which the magnitude of the factor loadings (λ_i) for particular items are the same over time (Vandenberg and Lance 2000; Brown 2006) and is a prerequisite for making valid comparisons (Bollen 1989); (3) scalar invariance tests for evidence that item thresholds (τ_i) are invariant over time (Brown 2006); and (4) strict invariance is performed to determine whether the residuals (θ_i) of the regression equations for each indicator are equivalent over time (where subscript i is the item) (Steinmetz et al. 2009).

Measurement invariance over time was considered to be present when, after imposing a constraint, there was no appreciable worsening of model fit. If this condition was met,

invariance testing proceeded to the application of the next equality constraint. If there was significant worsening of fit, modification indices were reviewed and constraints on non-invariance were removed and allowed to vary freely. This respecified model was then tested against the less constrained model to determine if model fit significantly improved. This strategy, known as partial invariance, requires that some, but not all parameters be constrained over time (Byrne et al. 1989).

Determination of model fit was based on three goodness-of-fit indices. These were the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Weighted Root Mean Square Residual (WRMR), and Root Mean Square Error of Approximation (RMSEA) and 90 % confidence interval (MacCallum et al. 1996; Cheung and Rensvold 2002; Singh 2009). Because the χ^2 test is sensitive to sample size (Singh 2009), it was examined, but not considered in decisions regarding model fit. Adequate model fit was defined using the following cutoffs: CFI >0.90 , TLI >0.90 , WRMR <0.90 , and RMSEA <0.06 (Yu 2002; Singh 2009; Byrne 2012). If at least three of the four fit indices exceeded the specified cutoff, model fit was deemed adequate (Van Lieshout et al. 2011; Ferro and Boyle 2012).

Due to the ordered categorical nature of responses for the CES-D, the confirmatory factor model was estimated with a weighted least squares estimator using a diagonal weight matrix to generate robust parameter estimates (Muthén and Muthén 2010). The χ^2 difference test (using the DIFFTEST option in *Mplus*), was used to determine the presence of invariance at each level of the analysis (Muthén and Muthén 2010). However, given that the $\Delta\chi^2$ is influenced by sample size (Chen et al. 2005), the following changes in fit indices were also used to determine invariance at each stage of testing (Chen 2007): $\Delta\text{CFI} \geq -0.010$ and $\Delta\text{RMSEA} \geq 0.015$. In the case of mixed results, closeness to threshold scores and consistency across fit statistics were used to determine invariance.

Full information maximum likelihood was employed to include participants with missing data for the dependent variable, depressive symptoms. Complete data were available for $n=219$ individuals, and the pattern of missingness was as follows: missing on one variable, $n=59$; two variables, $n=75$; three variables, $n=1$; all variables, $n=21$. Thus, the total sample used for invariance testing was $N=326$. All statistical analyses associated with the longitudinal invariance testing were performed with *Mplus* 6 (Muthén & Muthén, USA). Descriptive statistics of the sample were obtained using SAS 9.2 (SAS Inc., USA).

Results

The mean age of women in this study was 37.6 (6.1) years at the time their children were diagnosed with epilepsy. As shown in Table 1, 80 % of mothers were married, 62 % had

Table 1 Sociodemographic characteristics of the sample at baseline

Characteristic	<i>N</i>	%
Marital status		
Not married	68	19.6
Married ^a	279	80.4
Education		
Primary school	22	8.0
Secondary school	67	22.3
College/University	185	61.6
Graduate/Professional school	24	8.0
Employment status		
Not employed	33	9.6
Employed	228	66.3
Homemaker	78	22.7
Student	5	1.5
Household income		
<\$20,000	27	8.11
\$20–39,999	47	14.1
\$40–59,999	69	20.7
\$60–79,999	59	17.7
≥\$80,000	123	36.9

Frequencies and percentages may not sum to $N=347$ or 100 % due to missing data and rounding error

^a Includes mothers in married and common-law relationships

graduated from college or university, 66 % were employed either full- or part-time, and 37 % had annual household incomes \geq \$80,000. The mean CES-D score at baseline, 12, and 24 months was 20.7 (7.4), 19.1 (6.5), and 18.9 (7.2), respectively. At total of 81 women were lost during follow-up. Compared to those who completed all measurement occasions, non-completers were more likely to be younger ($p=0.0002$), not married ($p=0.0040$), had lower education ($p=0.0122$), and had lower household income ($p=0.0191$).

Prior to testing measurement invariance of the original factor structure of the CES-D, alternative models were examined to determine which provided the best fit to the data. The alternate models tested were based on previously reported CES-D models and included (Perreira et al. 2005; Shafer 2006): (1) three-factor model with depressed and positive affect

domains combined, (2) three-factor model with depressed affect and somatic activity domains combined, (3) two-factor model with depressed affect and somatic activity and positive affect and interpersonal relations combined, and (4) one-factor model. Results from fitting these alternate models are shown in Table 2 and suggested that the original four-factor structure of CES-D demonstrated the best fit, and thus, tests for longitudinal invariance continued with this model. Parameter estimates for the four-factor model are shown in Fig. 1.

The first step in assessing longitudinal measurement invariance was to establish a well-fitting baseline model (configural invariance) whereby no constraints are placed on the model. Invariance was modeled to test whether an equal number of factors and the pattern of loadings in this longitudinal sample represented those observed in the original construction of the scale. This configural model demonstrated adequate fit, $\chi^2(1,644)=2,213.74$; CFI=0.965; TLI=0.963; WRMR=1.050; RMSEA=0.029 [0.026, 0.033] and thus was used as the basis for testing more constrained models. A summary of model fit indices associated with the longitudinal invariance testing of measurement and structure of the CES-D is shown in Table 3.

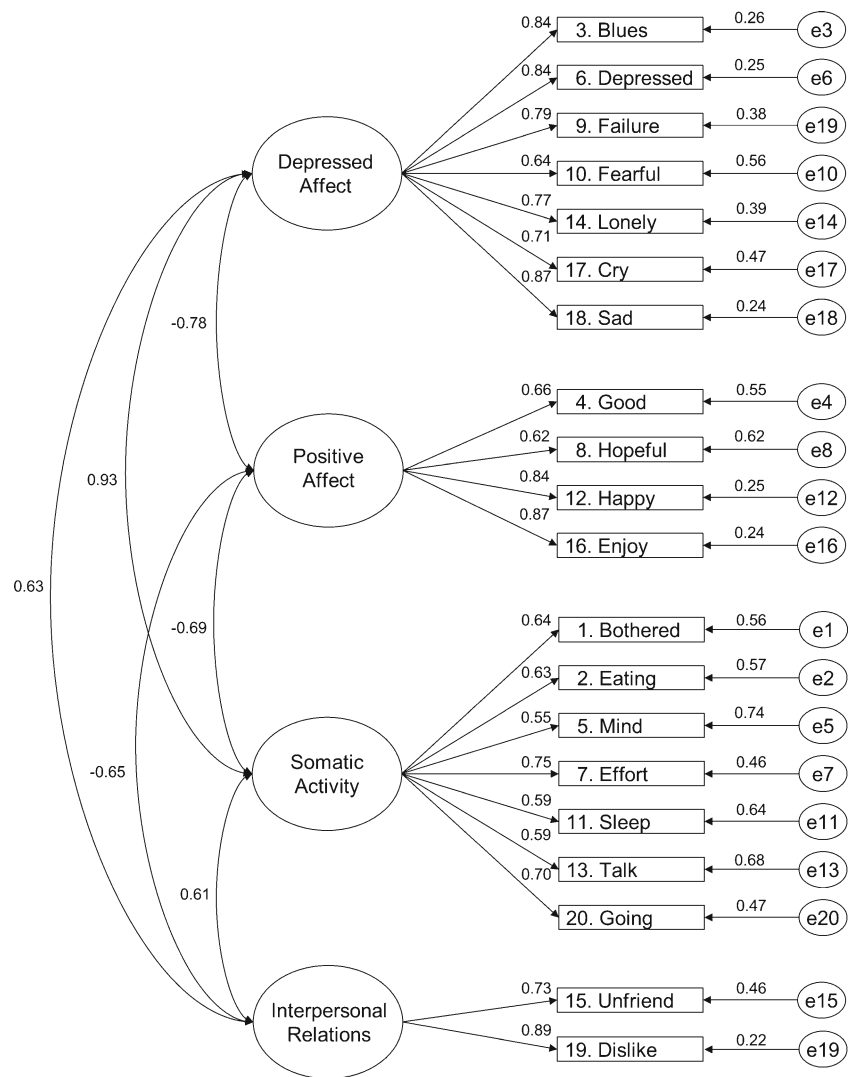
In the first stage of testing, which focuses on longitudinal measurement invariance, equality constraints were set on the factor loadings to test whether the magnitudes of factor loadings were equal over time (metric invariance). In comparison to the configural model, there was no substantial worsening of model fit, Δ CFI=-0.001; Δ RMSEA=0.001, indicating metric invariance was observed. Next, thresholds were constrained to be equal over time (scalar invariance). Results suggested the model was invariant over time, Δ CFI=-0.003; Δ RMSEA=0.001. In the final step of measurement invariance testing, residual variances were constrained to be equal over time (strict invariance). Compared to the scalar model, there was no substantial difference in model fit, Δ CFI=-0.004; Δ RMSEA=0.002. Thus, the CES-D demonstrated full longitudinal measurement invariance in this sample of mothers. Parameter estimates and the covariance matrix of final invariant model are shown in Tables 4 and 5, respectively. Internal consistency estimates (using Cronbach's α) were good at each measurement occasion, ranging from 0.76 to 0.81.

Table 2 Comparison of alternate CES-D factor structures

	χ^2 (<i>df</i>)	CFI	TLI	WRMR	RMSEA (90 % CI)
Four-factor	362.95 (164)	0.965	0.960	0.966	0.059 (0.051, 0.068)
Three-factor (DA and PA combined)	525.18 (167)	0.938	0.929	1.222	0.079 (0.071, 0.087)
Three-factor (DA and SA combined)	384.13 (167)	0.962	0.957	1.002	0.061 (0.053, 0.069)
Two-factor (DA and SA, PA and IR combined)	419.52 (169)	0.956	0.951	1.076	0.066 (0.058, 0.073)
One-factor	611.42 (170)	0.923	0.914	1.339	0.087 (0.079, 0.094)

DA depressed affect, PA positive affect, SA somatic activity, IR interpersonal relations

Fig. 1 Standardized parameter estimates for the four-factor CES-D model. All parameter estimates are statistically significant at $p < 0.001$



Discussion

To our knowledge, this is the first study to examine the factor structure and test measurement invariance of the CES-D using a longitudinal framework in a large sample of women. Results from the current study support the original model proposed by Radloff (1977); the CES-D can be explained by four latent factors: depressed affect, positive affect, somatic activities, and interpersonal difficulties. This finding is congruent with

a recent meta-analysis which demonstrated that the four-factor model of the CES-D is the most commonly reported structure of the measure (Shafer 2006). These findings also align with the results from Knight et al. (1997), who reported an adequate fit for the four-factor model in a large sample of women 36–67 years of age.

The extent to which the CES-D exhibited longitudinal invariance over time across three measurement occasions taken at 1-year intervals was also examined in this prospective

Table 3 Model fit summary of measurement and structural invariance testing

	χ^2 (df)	TLI	CFI	WRMR	RMSEA (90 % CI)	$\Delta\chi^2$ (df)	Δ CFI	Δ RMSEA
Configural	2,185.96 (1,644)	0.963	0.965	1.050	0.029 (0.026, 0.033)	–	–	–
Metric	2,213.74 (1,676)	0.962	0.964	1.102	0.030 (0.026, 0.033)	43.70 (32)	-0.001	0.001
Scalar	2,287.36 (1,716)	0.958	0.961	1.126	0.031 (0.029, 0.035)	110.25 (40)	-0.003	0.001
Strict	2,303.92 (1,776)	0.952	0.957	1.149	0.033 (0.030, 0.038)	166.34 (60)	-0.004	0.002

CI confidence interval

Table 4 Parameter estimates and standard errors for the final invariant model

	Factor loading (λ)	Residual variance (θ)
Depressed affect		
03. Blues	0.92 (0.04)	0.43 (0.03)
06. Depressed	0.95 (0.04)	0.38 (0.02)
09. Failure	0.89 (0.05)	0.33 (0.03)
10. Fearful	0.73 (0.05)	0.67 (0.03)
14. Lonely	0.92 (0.05)	0.43 (0.03)
17. Cry	0.83 (0.05)	0.56 (0.03)
18. Sad	0.99 (0.04)	0.35 (0.02)
Positive affect		
04. Good	0.68 (0.05)	0.73 (0.06)
08. Hopeful	0.75 (0.05)	0.66 (0.05)
12. Happy	0.95 (0.04)	0.27 (0.03)
16. Enjoy	0.94 (0.04)	0.30 (0.03)
Somatic activity		
01. Bothered	0.69 (0.03)	0.79 (0.03)
02. Eating	0.61 (0.05)	0.80 (0.04)
05. Mind	0.67 (0.05)	0.87 (0.04)
07. Effort	0.88 (0.04)	0.64 (0.05)
11. Sleep	0.70 (0.05)	0.90 (0.05)
13. Talk	0.74 (0.04)	0.73 (0.03)
20. Going	0.86 (0.04)	0.57 (0.03)
Interpersonal relations		
15. Unfriendly	0.90 (0.04)	0.54 (0.03)
19. Dislike	1.12 (0.06)	0.25 (0.03)

All parameter estimates are statistically significant at $p < 0.001$

cohort study of Canadian mothers. Results provided support for full measurement invariance; an important prerequisite for longitudinal comparison of depressive symptoms (Steinmetz

et al. 2009). The presence of metric invariance suggested that the 20 observed indicators were perceived similarly over time, and scalar invariance provided evidence that the magnitude of any biases in responses to an item was equal over time. Strict invariance was also demonstrated which suggested that indicators had the same accuracy and were reliable over time.

While full measurement invariance was established in the current study, researchers have argued that full invariance is not required for substantive analyses if at least a subset of parameters are invariant (Byrne et al. 1989). Indeed, previous research has demonstrated that two metric and scalar invariant parameters (i.e., factor loadings and intercepts) are sufficient for meaningful comparisons between groups or over time (Baumgartner and Steenkamp 1998). Consequently, these results provide the evidence to suggest that the latent structure of the CES-D exhibited stationarity and stability (Tisak and Meredith 1990; Pitts et al. 1996). In other words, the same underlying construct is being measured over time (stationarity) and that the rank order of individuals on the construct remains constant over time (stability).

Given that longitudinal invariance was established, results have practical implications for clinicians and researchers. Since the CES-D demonstrated no difference in structure, scaling, or reliability, changes in mean levels of depressive symptoms can be attributable to real differences in depressive symptoms over time, as opposed to differences arising as an artifact associated with changes in the interpretation of items as individuals mature. This finding is particularly relevant for investigators interested in maternal-child health as depression is understood to be the leading cause of disease burden among women in their reproductive years (Lopez et al. 2006). In addition, it is a fundamental construct underlying the mental health of mothers and how they adapt and cope with their

Table 5 Covariance matrix (ϕ) for the final invariant model

	$F_{(1,1)}$	$F_{(2,1)}$	$F_{(3,1)}$	$F_{(4,1)}$	$F_{(1,2)}$	$F_{(2,2)}$	$F_{(3,2)}$	$F_{(4,2)}$	$F_{(1,3)}$	$F_{(2,3)}$	$F_{(3,3)}$	$F_{(4,3)}$
$F_{(1,1)}$												
$F_{(2,1)}$	-0.76											
$F_{(3,1)}$	0.91	-0.66										
$F_{(4,1)}$	0.59	-0.61	0.58									
$F_{(1,2)}$	0.69	-0.64	0.57	0.48								
$F_{(2,2)}$	-0.55	0.58	-0.53	-0.48	-0.89							
$F_{(3,2)}$	0.58	-0.52	0.67	0.49	0.92	-0.74						
$F_{(4,2)}$	0.45	-0.41	0.39	0.64	0.64	-0.55	0.61					
$F_{(1,3)}$	0.70	-0.50	0.59	0.47	0.66	-0.65	0.53	0.44				
$F_{(2,3)}$	-0.49	0.58	-0.42	-0.35	-0.56	0.71	-0.37	-0.35	-0.88			
$F_{(3,3)}$	0.58	-0.45	0.70	0.45	0.57	-0.60	0.74	0.49	0.94	-0.80		
$F_{(4,3)}$	0.49	-0.42	0.49	0.58	0.47	-0.45	0.41	0.52	0.80	-0.51	0.79	

The first subscript denotes the factor, and the second subscript denotes the measurement occasion, where factor 1 is depressed affect; factor 2 is positive affect; factor 3 is somatic activity; and, factor 4 is interpersonal relations. All covariance estimates are statistically significant at $p < 0.001$

children's illness (Ferro et al. 2011b; Ferro and Speechley 2012) and is a salient risk factor predicting health outcomes in children (Ferro and Speechley 2009; Ferro et al. 2011a). For example, we have previously shown that mothers of children with epilepsy are not homogeneous, but instead consist of distinct heterogeneous subgroups with unique trajectories of depressive symptoms (Ferro et al. 2011b). Furthermore, we have documented that maternal depression has a negative impact on the health-related quality of life of children with epilepsy, such that children of mothers with depression have lower health-related quality of life scores and less favorable trajectories during the first 24 months after their diagnosis (Ferro et al. 2011a). Importantly, since the sociodemographic characteristics of the study sample were similar to those of Canadian women in the general population (Canada 2006), results are likely externally valid. Thus, when using the CES-D in other adult woman populations, researchers should be confident that changes in CES-D scores over time are reflective of true changes in depressive symptomatology and not an artifact of changes in the interpretation of items in the scale.

Conclusions

The present study provided evidence to support the use of the original, four-factor model of the CES-D in a large sample of adult women. Additionally, results from the measurement invariance testing allow for meaningful comparisons of mean CES-D factor scores over time. Research should continue examining longitudinal invariance of the CES-D in other mother-child populations so that if invariance is established, clinicians and researchers can be confident in using this measure to document depressive symptoms. Future work building upon the results of this study and incorporating tests of invariance between subgroups (e.g., women with vs. without children) over time would make an important contribution to the understanding of depression among mothers.

Acknowledgments Dr. Ferro is supported by a Banting Postdoctoral Fellowship from the Canadian Institutes of Health Research. The HERQULES project was supported by a grant from the Canadian Institutes of Health Research (MOP-64311).

References

- Baumgartner H, Steenkamp JBEM (1998) Multi-group latent variable models varying numbers of items and factors with cross-national and longitudinal applications. *Mark Lett* 9:21–35
- Bearslee WR, Versage EM, Gladstone TR (1998) Children of affectively ill parents: a review of the past 10 years. *J Am Acad Child Psychiatry* 37:1134–1141
- Bollen KA (1989) *Structural equations with latent variables*. Wiley, New York
- Brown TA (2006) *Confirmatory factor analysis for applied research*. The Guilford Press, New York
- Byrne BM (2012) *Structural equation modeling with Mplus: basic concepts, applications, and programming*. Taylor & Francis Group, LLC, New York
- Byrne BM, Shavelson RJ, Muthen B (1989) Testing for the equivalence of factor covariance and mean structures: the issue of partial measurement invariance. *Psychol Bull* 105:456–466
- Canada S (2006) *Census of Canada*. Statistics Canada, Ottawa
- Chen FF (2007) Sensitivity of goodness of fit indices to lack of measurement invariance. *Struct Equ Model* 14:464–504
- Chen FF, Sousa KH, West SG (2005) Testing measurement invariance of second-order factor models. *Struct Equ Model* 12:471–492
- Cheung GW, Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Model* 9:233–255
- Cohen CI, Goh KH, Gustave M (2010) A prospective study of outcome and predictors of subclinical and clinical depression in an older biracial sample of psychiatric outpatients. *J Affect Disord* 121:204–211
- Conrada RJ, Boulifard DA, Idler EL, Krause TJ, Labouvie EW (2006) Course of depressive symptoms in patients undergoing heart surgery: confirmatory analysis of the factor pattern and latent mean structure of the Center for Epidemiologic Studies Depression Scale. *Psychosom Med* 68:922–930
- Dolgin MJ, Phipps S, Fairclough DL, Sahler OJ, Askins M, Noll RB, Butler RW, Varni JW, Katz ER (2007) Trajectories of adjustment in mothers of children with newly diagnosed cancer: a natural history investigation. *J Pediatr Psychol* 32:771–782
- Edwards MC, Cheavens JS, Heiy JE, Cukrowicz KC (2010) A reexamination of the factor structure of the Center for Epidemiologic Studies Depression Scale: is a one-factor model plausible? *Psychol Assess* 22:711–715
- Ferro MA, Boyle MH (2012) Longitudinal invariance of measurement and structure of global self-concept: a population-based study examining trajectories among adolescents with and without chronic illness. *J Pediatr Psychol*. doi:10.1093/jpepsy/jss112
- Ferro MA, Speechley KN (2009) Depressive symptoms among mothers of children with epilepsy: a review of prevalence, associated factors, and impact on children. *Epilepsia* 50:2344–2354
- Ferro MA, Speechley KN (2012) Examining clinically relevant levels of depressive symptoms in mothers following a diagnosis of epilepsy in their children: a prospective analysis. *Soc Psychiatry Psychiatr Epidemiol* 47:1419–1428
- Ferro MA, Avison WR, Campbell MK, Speechley KN (2011a) The impact of maternal depressive symptoms on health-related quality of life in children with epilepsy: a prospective study of family environment as mediators and moderators. *Epilepsia* 52:316–325
- Ferro MA, Avison WR, Campbell MK, Speechley KN (2011b) Prevalence and trajectories of depressive symptoms in mothers of children with newly diagnosed epilepsy. *Epilepsia* 52:326–336
- Gillespie JF, Primavera J (eds) (2000) *Diverse families, competent families: innovations in research and preventive intervention practice*. Haworth Press, New York
- Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D (2011) Maternal depression and child psychopathology: a meta-analytic review. *Clin Child Fam Psychol Rev* 14:1–27
- Hertzog C, Van Alstine J, Usala P, Hultsch DF, Dixon R (1990) Measurement properties of the Center for Epidemiological Studies Depression Scale (CES-D) in older populations. *Psychol Assess* 2:64–72
- Kim-Cohen J, Moffitt TE, Taylor A, Pawlby SJ, Caspi A (2005) Maternal depression and children's antisocial behavior: nature and nurture effects. *Arch Gen Psychiatry* 62:173–181
- Knight RG, Williams S, McGee R, Olaman S (1997) Psychometric properties of the Centre for Epidemiologic Studies Depression

- Scale (CES-D) in a sample of women in middle life. *Behav Res Ther* 35:373–380
- Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJL (eds) (2006) *Global burden of disease and risk factors*. The World Bank and Oxford University Press, New York
- MacCallum RC, Browne MW, Sugawara HM (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods* 1:130–149
- Millsap RE, Yun-Tein J (2004) Assessing factorial invariance in ordered-categorical measures. *Multivar Behav Res* 39:479–515
- Morin AJ, Moullec G, Maiano C, Layet L, Just JL, Ninot G (2011) Psychometric properties of the Center for Epidemiologic Studies Depression Scale (CES-D) in French clinical and nonclinical adults. *Rev Epidemiol Sante Publique* 59:327–340
- Motl RW, Dishman RK, Birnbaum AS, Lytle LA (2005) Longitudinal invariance of the Center for Epidemiological Studies-Depression Scale among girls and boys in middle school. *Educ Psychol Meas* 65:90–108
- Muthén BO (1989) Latent variable modeling in heterogeneous populations. *Psychometrika* 54:557–585
- Muthén LK, Muthén BO (2010) *Mplus user's guide*, 6th edn. Muthén & Muthén, Los Angeles
- Perreira KM, Deeb-Sossa N, Harris KM, Bollen K (2005) What are we measuring? An evaluation of the CES-D across race/ethnicity and immigrant generation. *Soc Forces* 83:1567–1602
- Pilowsky DJ, Wickramaratne P, Nomura Y, Weissman MM (2006) Family discord, parental depression, and psychopathology in offspring: 20-year follow-up. *J Am Acad Child Adolesc Psychiatry* 45:452–460
- Pitts SC, West SG, Tein J (1996) Longitudinal measurement models in evaluation research: examining stability and change. *Eval Prog Plan* 19:333–350
- Radloff LS (1977) The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1:385–401
- Shafer AB (2006) Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *J Clin Psychol* 62:123–146
- Singh R (2009) Does my structural model represent the real phenomenon? A review of the appropriate use of structural equation modeling (SEM) model fit indices. *Mark Rev* 9:199–212
- Steinmetz H, Schmidt P, Tina-Booh A, Wieczorek S, Schwartz SH (2009) Testing measurement invariance using multiple group CFA: differences between educational groups in human values measurement. *Qual Quant* 43:599–616
- Tisak J, Meredith W (1990) Longitudinal factor analysis. In: von Eye A (ed) *Statistical methods in longitudinal research*, volume 1: principles and structuring change. Academic, San Diego
- Van Lieshout RJ, Cleverley K, Jenkins JM, Georgiades K (2011) Assessing the measurement invariance of the Center for Epidemiologic Studies Depression Scale across immigrant and non-immigrant women in the postpartum period. *Arch Womens Ment Health* 14:413–423
- Vandenberg RJ, Lance CE (2000) A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Methods* 3:4–70
- Williams CD, Taylor TR, Makambi K, Harrell J, Palmer JR, Rosenberg L, Adams-Campbell LL (2007) CES-D four-factor structure is confirmed, but not invariant, in a large cohort of African American women. *Psychiatry Res* 150:173–180
- Yu CY (2002) *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. University of California, Los Angeles