Reevaluating Positive Affect in the Center for Epidemiologic Studies-Depression scale

Trevor Randell Olson [a],*, Michelle Dawn Presniak [a], Michael William MacGregor [b]

[a] Institute of Community and Family Psychiatry at SMBD Jewish General Hospital, Department of Psychiatry, McGill University, Montreal, Canada
[b] Department of Psychology, University of Saskatchewan, Saskatoon, Canada

ABSTRACT

In 2000, Schroevers and colleagues examined the reliability and validity of a two-factor structure for the Dutch version of the Center for Epidemiologic Studies-Depression (CES-D) scale in cancer patients and in a community sample. The authors concluded that a two-factor structure assessing Positive Affect (PA) and Depressed Affect was a better fit to the data than the standard four-factor structure. They argued further that the four reverse-scored items composing the PA scale should be dropped. Using similar measures and analyses as Schroevers and colleagues, we examined the factor structure and concurrent validity of the English version of the CES-D in university student and community samples. Across both samples the factor structure was more similar to Radloff's (1977) original four-factor structure than to the proposed two-factor structure. Additionally, our validity analyses indicated no problems with the PA items and suggested that PA might be more specifically related to depression than to other forms of psychopathology. We recommend that clinicians and researchers using the English version of the CES-D continue to use the full 20-item version.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The Center for Epidemiologic Studies-Depression (CES-D) scale is a self-report measure of depressive symptomatology for use in the general population (Radloff, 1977). The reliability and validity of the CES-D have been examined extensively (e.g., Hertzog et al., 1990; McCallum et al., 1995; MacKinnon et al., 1998). In Radloff's original study, the CES-D demonstrated high internal consistency and acceptable test–retest reliability. Strong evidence for the validity of the CES-D scores was demonstrated by relations to observer and self-ratings of depression.

To evaluate the factor structure of the CES-D, Radloff (1977) conducted a principal components analysis (PCA). A four-factor pattern (Depressed Affect, Positive Affect, Somatic and Retarded Activity, and Interpersonal) emerged consistently across three samples. Subsequent evaluations have demonstrated results similar to Radloff's findings (e.g., Fava, 1983; Ross and Mirowsky, 1984) and strong evidence for the same factor structure across several age, cultural, and socioeconomic groups (Hertzog et al., 1990; McCallum et al., 1995; MacKinnon et al., 1998; Wong, 2000; Boisvert et al., 2003). In 2006, Shafer conducted a meta-analysis of four depression questionnaires (i.e., Beck, CES-D, Hamilton, and Zung). Across the 21 articles (28 studies) and 22,340 participants included by Shafer, every CES-D item replicated into Radloff's original four-factor structure. In

summarizing the results, Shafer wrote "Overall, the CES-D had relatively little variability across the factor analysis studies included in this meta-analysis compared to the other three depression tests. The results were clear and highly consistent with the initial factor analyses conducted by Radloff during the development of the CES-D" (p. 133–134).

Not every study, however, has replicated the four-factor structure. For example, Thomas and Brantley (2004) found that three factors (Depressed Affect/Somatic, Positive Affect, and Interpersonal) provided the best fit to their CES-D data using a confirmatory factor analysis (CFA) in a sample of women with low-income. The authors hypothesized that minority groups might express depression more somatically than the general population. They further noted that their results were similar to the three-factor structure found by Guarnaccia et al. (1989) in a sample of Mexican-American women.

In 2000, Schroevers and colleagues argued that a two-factor solution provided the best fit for their data using the Dutch version of the CES-D. The authors conducted a forced two-factor PCA and across a group of cancer patients and a nonclinical reference group, they found one factor consisting of the 16 negatively formulated items (Depressed Affect; DA) and a second factor consisting of the 4 positively formulated items (Positive Affect; PA). Next, they investigated the validity of the two factors by correlating each factor with several measures of psychosocial functioning. The authors determined that DA was strongly related ($r > 0.40$) to measures of anxiety, depression, and general psychological distress, while PA was more weakly correlated ($r \leq 0.40$) with the same measures. Finally, they examined the utility of DA and PA in discriminating between the cancer patient and community reference

* Corresponding author. FIT for Active Living C/O Saskatoon City Hospital 8th Floor, 701 Queen Street Saskatoon, SK S7K 0M7, Canada. Tel.: +1 306 655 8950.
E-mail address: trevor.olson@usask.ca (T.R. Olson).

groups on depressive symptoms. They found that the patients had significantly higher DA mean scores compared to the reference group, while there was no significant difference between groups on PA mean scores. Schroevers and colleagues concluded that there was weak support for the validity of the PA items as a measure of depressive symptomatology and argued that a summed score of the 16 DA items was a "more valid measure of depressive symptomatology, both in cancer patients and in a matched reference group of healthy individuals from the general population" (p. 1026).

The first goal of the current study was to attempt to replicate some of Schroevers et al.'s (2000) empirical findings in two samples. Their results appear to contrast previous findings that have replicated the original four-factor structure across nations, age groups, and cultural groups. Moreover, previous work (e.g., Watson et al., 1988a; Clark and Watson, 1991) has indicated that low positive affect is an important factor in the assessment of depression. As such, we examined the factor structure of the English CES-D and the concurrent validity of the proposed DA and PA factors in university student and community participant samples. Both the student and community samples were previously collected for unrelated studies. From the measures available in these samples, we selected validity comparisons as similar as possible to those used by Schroevers and colleagues as well as those available across both of the previously collected samples to increase their comparability. We attempted to answer three questions:

1) Do the factor analyses in each sample replicate the findings of either Schroevers et al. (2000) or Radloff (1977)?
2) Are psychosocial variables related weakly to PA items but strongly to DA items?
3) In the student sample, we also investigated a) are there mean differences in DA and PA summed scores between participants with low, moderate, and high self-reported depression symptomatology? and b) do the DA and PA scores account for unique variance in several psychopathology scales?

2. Method

2.1. Participants

2.1.1. Student sample
The sample consisted of 1513 undergraduate students from a Western Canadian university. The mean age of the sample was 19.7 years (range = 17–40; S.D. = 2.6), and the majority of the participants were female (70.9%). Self-reported ethnicity was predominantly Caucasian (87.7%), followed by Asian (4.5%) and First Nations (2.0%).

2.1.2. Community sample
The sample consisted of 3227 participants from the 1995 Nova Scotia Health Survey, which was representative of the adult provincial population by participant age and sex (Maclean et al., 1996). The mean age was 48.1 years (range 18–99; S.D. = 19.8), and 50.1% of the participants were female. Self-reported ethnicity was predominantly Caucasian (98.6%), followed by African-Canadian (1.1%) and Asian (0.2%).

2.2. Measures

2.2.1. CES-D (Radloff, 1977)
The CES-D is a 20-item self-report measure of current depressive symptomatology. Using a 4-point Likert scale, participants rate how often they experienced each symptom during the previous week. A summed score represents a participant's overall depressive level, with a score of 16 or greater indicating possible clinical depression.

2.2.2. Personality Assessment Inventory (PAI; Morey, 1991)
The PAI is a 344-item self-report measure of personality that assesses 22 scales including 4 validity, 2 interpersonal, 5 treatment, and 11 clinical scales. Participants rate their agreement with each statement on a 4-point Likert scale. To more closely match the constructs used by Schroevers et al. (2000), only the Somatic Complaints, Anxiety, Anxiety Related Disorders, Depression, Stress, and Nonsupport scales were used for correlation analyses. All 11 clinical scales were used in the regression analyses.

2.2.3. The State Trait Anxiety Inventory (STAI; Spielberger, 1983)
The STAI is a 40-item measure assessing both State and Trait Anxiety. Participants rate statements on a 4-point Likert scale indicating their experience of anxiety and tendency to perceive stressful events as threatening. For the current study, only the 20-item Trait Anxiety score was used.

2.2.4. The Anger Expression Scale (AES; Spielberger et al., 1985)
The AES is a 20-item measure of aspects of anger. Participants rate items on a 4-point Likert scale to indicate the extent to which they experience various aspects of anger.

2.2.5. Cook–Medley Hostility Scale (Ho; Cook and Medley, 1954)
The Ho is a 50-item true/false measure of various aspects of hostility that was developed from the MMPI.

2.2.6. Positive and Negative Affect Scale (PANAS; Watson et al., 1988b)
The PANAS is a 20-item measure of positive and negative affect. Participants rate each item on a 5-point Likert scale to indicate the extent to which they experience various positive and negative symptoms.

2.2.7. Big Five Inventory-54 (BFI-54; John et al., 1992)
The BFI-54 is a 54-item self-report measure of the five factors of personality (i.e., Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness). Participants respond on a 5-point Likert scale to short phrases describing trait adjectives known to be prototypical of the five factors.

2.2.8. Defense-Q (MacGregor and Olson, 2005)
The Defense-Q is an observer report measure of defense mechanisms. The Adaptive Defense Profile Similarity Score (ADPSS) is a measure of the healthiness or adaptiveness of a person's overall defense mechanism use.

2.3. Statistical analysis

2.3.1. Factor analyses
To replicate Schroevers et al.'s (2000) study, a PCA with varimax rotation forcing a two-factor solution was conducted. In addition, we conducted a non-forced principal axis factoring (PAF) analysis with oblique rotation, based on previous findings that the CES-D factors are interrelated and recommendations that when examining scales for underlying constructs, particularly when there are a small number of variables (i.e., 20), PAF with oblique rotation is the preferred analysis (Hertzog et al., 1990; Tabachnick and Fidell, 2001; Field, 2005). Of note, although many studies have examined the factor structure using confirmatory factor analyses, CFA and exploratory factor analysis (EFA) procedures share few statistical similarities and as such, their results should not be expected to be the same (Kline, 2004). Given that Schroevers et al. (2000), as well as Radloff (1977) used EFA, we chose to use exploratory analyses as well.

2.3.2. Validity
Pearson correlations were used to examine the relation of the above psychosocial variables to DA, PA and the full 20-item CES-D score. In the student sample, ANOVAs were conducted to examine mean differences in both DA and PA scores across participants with low, moderate, and high depression scores on the PAI. Finally, regression analyses were conducted in the student sample to determine the degree to which DA and PA (entered together as predictors) accounted for unique variance in the 11 PAI clinical scales.

3. Results

3.1. Factor analysis

3.1.1. Forced two-factor PCA with varimax rotation
The total variance accounted for in the student sample was 44.2%. The 4 PA items loaded onto the second factor along with 6 of 16 DA items (see Table 1). Additionally, there were a large number of residuals (51.0%) over 0.05 and four of the eigenvalues were greater than 1 before rotation. The total variance accounted for in the community sample was 38.5%. The 4 PA items loaded onto the second factor along with 2 of 16 DA items. One DA item did not load onto either factor (<0.32).[1] There were a large number of residuals (40%) over 0.05 and four eigenvalues were greater than 1 before rotation.

3.1.2. Non-forced PAF with oblique rotation
This student sample solution produced four factors (eigenvalues>1) in which the variance accounted for was 55.2% (total) and 44.8% (common). Only 3.0% of the residual values were over 0.05. Three factor loadings were below 0.32 (see Table 2). In the community sample the solution produced four factors (eigenvalues>1) in which the variance accounted for was 49.9% (total) and 38.7% (common). Only 3% of the residual values were over 0.05. Three factor loadings were below 0.32. Taking into account the variance accounted for, number of residuals

---

[1] Based on Tabachnick and Fidell (2001), a significance level of 0.32 was chosen for the factor loadings.