HHS Public Access

Author manuscript

Psychol Assess. Author manuscript; available in PMC 2018 March 01.

Published in final edited form as:

Psychol Assess. 2017 March ; 29(3): 320–328. doi:10.1037/pas0000336.

Factor Structure of the Perceived Stress Scale-10 (PSS) Across English and Spanish Language Responders in the HCHS/SOL Sociocultural Ancillary Study

Marisa J. Perera, M.S.[1], **Carrie E. Brintz, M.S.**[1], **Orit Birnbaum-Weitzman, Ph.D.**[1], **Frank J. Penedo, Ph.D.**[2], **Linda C. Gallo, Ph.D.**[3], **Patricia Gonzalez, Ph.D.**[3], **Natalia Gouskova, M.S.**[4], **Carmen R. Isasi, MD, Ph.D.**[5], **Elena L. Navas-Nacher, M.S.**[6], **Krista M. Perreira, Ph.D.**[4], **Scott C. Roesch, Ph.D.**[3], **Neil Schneiderman, Ph.D.**[1], and **Maria M. Llabre, Ph.D.**[1]

[1] University of Miami, Coral Gables FL

[2] Northwestern University Feinberg School of Medicine, Chicago, IL

[3] San Diego State University, San Diego, CA

[4] University of North Carolina at Chapel Hill, Chapel Hill, NC

[5] Einstein College of Medicine, Bronx, NY

[6] University of Illinois, Chicago, IL

Abstract

Despite widespread use, psychometric investigation of both original English and translated Spanish versions of the 10-item Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983) has been limited among the U.S. Hispanic/Latino population. The present study examined the factor structure, factorial invariance, and reliability and validity of PSS scores from English and Spanish versions using data from 5,176 Hispanics/Latinos who participated in the Hispanic Community Health Study/Study of Latinos Sociocultural Ancillary Study. Total sample and language multigroup confirmatory factor analyses supported a bi-factor model with all 10 PSS items loading on a general perceived stress factor, and the four reverse worded items also loading on a reverse worded factor. Internal consistency ranged from .68-.78 and it was indicated that reliable variance exists beyond the general perceived stress factor. The model displayed configural, metric, scalar, and residual invariance across language groups. Convergent validity analyses indicated that both the general perceived stress factor and the reverse worded factor were related to scores of depression, anxiety, and anger in the expected directions. The reverse worded factor added to the validity of the PSS beyond the general perceived stress factor. The total computed score of the PSS can be recommended for use with Hispanics/Latinos in the U.S. that complete the measure in English or Spanish and the reverse worded factor can enhance prediction.

Keywords

perceived stress; Hispanic/Latino; psychometrics; factor structure; factorial invariance

Corresponding author: Marisa J. Perera, Department of Psychology, University of Miami, 5665 Ponce de Leon, Miami, FL 33146, Tel: (305) 284-2687, mperera@miami.edu.

The Perceived Stress Scale-10 (PSS; Cohen, Kamarack, & Mermelstein, 1983) is the most widely used measure of global perceived stress, and is a robust predictor of health and disease (Cohen, Janicki-Deverts, & Miller, 2007; Cohen, Tyrrell, & Smith, 1993). The original English version of the PSS has been translated into several languages to increase its use with non-English speaking populations (Almadi, Cathers, Mansour, & Chow, 2012; Mimura & Griffiths, 2004), including Spanish translations for use with Hispanics/Latinos (Ramírez & Hernández, 2007). Hispanics/Latinos in the U.S. are the largest and fastest growing minority group (Passel, Cohn, & Lopez, 2011; Suro & Singer, 2002), and there is increased interest in the health status of this group (Franzini, Ribble, & Keddie, 2000). Many Hispanics/Latinos residing in the U.S. have limited English proficiency (Derose & Baker, 2000). To enable the inclusion of monolingual Spanish speakers in research and clinical contexts alongside English speakers, both English and Spanish versions of the PSS are used to assess perceived stress in the U.S. Hispanic/Latino population (Alegria et al., 2004; Chiriboga, Black, Aranda, & Markides, 2002). Despite use of the PSS in Spanish language responders in the U.S., psychometric investigation of the Spanish translated PSS has been limited to populations in Spain and Mexico (Ramírez & Hernández, 2007; Remor, 2006; Remor & Carrobles, 2001; Trujillo & González-Cabrera, 2007). To date, the psychometric properties of the PSS have not been investigated in U.S. Hispanic/Latino communities.

There has been limited investigation of the factorial invariance of the original English and the translated Spanish PSS. Factorial invariance is the notion that the same construct of perceived stress is measured across both English and Spanish versions of the PSS (Gregorich, 2006). Invariance is often assumed rather than established for translated measures (Borsboom, 2006). A measure can vary, or perform differently across groups, as a result of true group differences in the construct and/or differences in the ways the groups experience and perceive the questionnaire content (Corral & Landrine, 2010). Violations to factorial invariance can result in misleading comparison of PSS scores between English and Spanish language responders. For instance, variance by language of response can reflect variance by acculturation, nativity (Schumann, 1986), and/or level of response bias (van Herk, Poortinga, & Verhallen, 2004). With fewer than one-in-four Hispanic/Latino U.S. immigrants reporting that they are fluent in English (Hakimzadeh & Cohn, 2007), both English and Spanish versions of the PSS are required to study perceived stress and related health outcomes in a representative sample of the U.S. Hispanic/Latino population. It is important to determine whether English and Spanish versions of the PSS are invariant, or equivalent to, one another prior to use of both English and Spanish PSS scores in practice. To the best of our knowledge, no study has investigated factorial invariance across English and Spanish versions of the PSS.

PSS Factor Structure

Past examination of the factor structure of the PSS has supported a two-factor structure for both the original English (Cohen, 1988; Lee, 2012; Taylor, 2015) and the translated Spanish versions (Remor, 2006; Ramírez & Hernández, 2007). The two-factor structure of the PSS includes one-factor that is grouped by negative wording of items (non reverse-worded scoring), and a second factor that is grouped by positive wording of items (reverse-worded

scoring). Some researchers have posited that the negative-worded factor measures perceived helplessness or negative stress, and that the positive-worded factor measures perceived self-efficacy or positive stress (Reis, Hino, & Anez, 2010; Roberti, Harrington, & Storch, 2006). Despite robust empirical support for a two-dimensional structure, the factor distinction has been regarded as superficial and not meaningful (Carretero-Dios & Pérez, 2007; Cohen, 1988; González-Ramírez, Rodríguez-Ayán, & Hernández, 2013; Lee, 2012; Ramírez & Hernández, 2007), and a single summed PSS score is most often used in research and clinical contexts. Overall, there is little research on the validity or reliability for each of the two factors. Notably, studies have examined the full 14-item PSS and have noted appropriate reliability (Cronbach alpha's ≥ .78) for each factor (Ramírez & Hernández, 2007), and that the the reverse-worded factor significantly predicted depressive symptoms (Hewitt, Flett, & Mosher, 1992).

Alternative to a two-factor model, a bi-factor model can determine whether the PSS is sufficiently unidimensional to support the use of an overall score, while allowing for additional specific variance resulting from reverse worded items. A bi-factor model of the PSS can examine a general factor that represents commonality of all scale items (in this case, perceived stress). A bi-factor model can also examine orthogonal, or unrelated, factors that represent the unique influence of a specific factor on subsets of scale items above and beyond the general factor (Reise, Morizot, & Hays, 2007). In a bi-factor model, PSS items can load on both a general perceived stress factor and on additional factors, such as a reverse worded factor. Inclusion of reverse worded items to control for response bias assumes that the reverse worded items measure the same construct as the non-reverse worded items (Barnette, 2000). Yet, the orthogonal (or nuisance) reverse worded factor can arise because of the reverse wording format that interferes with the measurement of the main construct of perceived stress. It is important to compare two-factor and bi-factor models to determine whether scale items are driven by two distinct factors or a general perceived stress factor. It may be especially important to model the reverse worded, or nuisance, variance when working with Spanish language responders. Reverse worded items can hold up poorly cross-culturally (Wong, Rindfleisch, & Burroughs, 2003), and particularly with non-English language responder populations for whom English measures have been translated (Llorente, Warren, de Eulate, & Gleaves, 2013).

In addition to factor structure and factorial invariance, the present study sought to examine whether both English and Spanish PSS versions reliably and validly measure perceived stress in a heterogeneous and representative sample of U.S. Hispanics/Latinos. The PSS was designed for use in community samples with at least a junior or high school education, however its use with diverse populations has increased. The populations that have been sampled for psychometric studies on the PSS have been university students and working populations (Lee, 2012). Some studies have examined the reliability and validity of the Spanish translated PSS, however samples have been specific to Mexico (González-Ramírez et al., 2013; Ramírez & Hernández, 2007) and Spain (Remor, 2006). Cronbach's alpha values for the PSS have ranged from .78 to .91 for English and Spanish versions (Lee, 2012; Remor, 2006). Evidence for the generalizability of the PSS for use with the U.S. Hispanic/Latino population is limited, at best. With regard to validity, studies have supported the convergent validity of the PSS with several emotion-based measures, such as depression and

anxiety. Given the limited generalizability to U.S. Hispanics/Latinos, it would be important to examine convergent validity to provide evidence of the construct validity of the PSS with U.S. Hispanics/Latinos for both English and Spanish versions.

The present study sought to evaluate the factor structure, factorial invariance, and reliability and validity of English and Spanish versions of the PSS in a population-based sample of 5,176 Hispanics/Latinos from the four urban U.S. communities that participated in the Hispanic Community Health Study/Study of Latinos (HCHS/SOL) and HCHS/SOL Sociocultural Ancillary Study. The sample, design, and procedures of the HCHS/SOL and the HCHS/SOL Sociocultural Ancillary study have been described elsewhere (Gallo et al., 2014; Sorlie et al., 2010). The aims of the present study were to: (1) examine the factor structure of the PSS by comparing one-factor, two-factor, and bi-factor models in the full sample and language groups; (2) test for configural, metric, scalar, and residual invariance in language groups; (3) examine the internal consistency of all PSS items in the full sample and in the English and Spanish language groups; (4) establish the convergent validity of the PSS with measures of depressive symptoms, trait anxiety, and trait anger. We hypothesized that a bi-factor structure would indicate better model fit relative to one- and two-factor models because of the underlying unidimensionality of the PSS (Carretero-Dios & Pérez, 2007; Cohen, 1988; González-Ramírez, Rodríguez-Ayán, & Hernández, 2013; Lee, 2012; Ramírez & Hernández, 2007). Although reliability and convergent validity have not been tested with a representative sample of Hispanics/Latinos in the US, we expected acceptable reliability and convergent validity given that they have been demonstrated with select Hispanic/Latino samples in the past (Lee, 2012, Remor, 2006). Based on previous research, we expected that the best fitting model of the PSS would demonstrate moderate to large positive associations with depressive symptoms, trait anxiety, and trait anger.

Method

Participants and Procedure

The present sample was derived from the HCHS/SOL Sociocultural Ancillary Study. The HCHS/SOL is a national cohort study that seeks to establish prevalence and risk factors for major chronic diseases, including cardiovascular disease, among 16,415 self-identified Hispanic/Latinos recruited in four field centers in the U.S. (Bronx, NY; Chicago, IL; Miami, FL; San Diego, CA). The HCHS/SOL sampling two-stage area household probability design strategy (LaVange et al., 2010) and approach have been previously detailed (Sorlie et al., 2010). The HCHS/SOL Sociocultural Ancillary Study comprised a separate assessment of socioeconomic, social, psychological, and cultural factors with hypothesized cardiovascular-metabolic health relevance among a subset of 5,313 participants from the original cohort (see Gallo et al., 2014). HCHS/SOL original cohort participants able and willing to complete a second visit within 3-9 months of the parent study baseline clinic exam were eligible for the Sociocultural Ancillary Study. Recruitment began during the second wave of parent study enrollment. A total of 5,313 (72.6%) of 7321 parent study participants for whom recruitment was attempted participated. The sample is considered a random subsample of HCHS/SOL participants, and there was the exception that participation was lower in some higher socioeconomic strata as discussed in LaVange et al., 2010. All self-report assessments

were administered via interview using a standardized approach to accommodate the wide range of education and literacy levels among participants. Interviews lasted one to two hours and participants were compensated $60 for their participation. Participants elected to complete the interview in Spanish (n = 4,169) or English (n = 1,143). Participants with missing data on any of the ten PSS items were excluded from the present analyses because the estimation method used in the factor analyses required complete data. A total of 137 participants (.025% of the total sample) were considered missing, resulting in an analytic sample of 5,176.

Measures

Demographic variables—Demographic information was collected during the HCHS/SOL baseline clinic exam. Age, sex, self-identified Hispanic/Latino background, marital status, income, education, number of years living in the United States, and language in which the interview was completed (either English or Spanish) were collected.

Measure translations—Although there are existing Spanish translations for some of the measures used in HCHS/SOL, not all translations were previously validated for use with all Hispanic/Latino ancestry groups, but rather with specific populations in Mexico or Spain, for example. Given the presence of multiple Hispanic/Latino ancestry groups in HCHS/SOL (i.e., Dominican, Cuban, Mexican, etc.) and in an effort to improve generalizability to the US Hispanic/Latino population, the Spanish versions of the PSS and other HCHS/SOL measures were translated by the HCHS/SOL team following recommended translation guidelines. The translation steps were as follows: (1) creation of two independent translations; (2) comparison and review of translations by a committee comprised of bilingual/bicultural members from each of the primary Hispanic/Latino background groups; (3) pilot-testing of the approved version via focus groups comprised of bilingual and monolingual representatives of each ethnic group; and (4) validation of translated instruments with a group of bilingual representatives. In addition, for the PSS the translation followed recommendations from Acquadro, Conway, Giroudet, and Mear (2006), Acquadro, Jambon, Ellis, and Marquis (1996) and Gallo et al. (2014) translation guidelines since they are specific to the PSS. Salman (1998) and Butcher, Cabiya, Lucio, and Garrido (2007) were referenced in translating the Spielberger inventories since they address Spanish translation issues for assessment of personality and traits.

Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983). The 10-item PSS measures global perceived stress experienced across the past 30 days on a 5-point scale (0 – *never*, 1 = *almost never*, 2 = *once in a while*, 3 = *often*, 4 = *very often*). Sample items include: “in the last month, how often have you felt that things were going your way” and “in the last month, how often have you found that you could not cope with all the things that you had to do”. Six of the 10 items were worded and scored in the non-reversed direction (i.e., “how often have you felt that you were unable to control the important things in your life”). Four of the 10 items were worded and scored in the reversed direction (i.e., “how often have you felt that things were going your way”). Total scores range from 0 to 40. Internal consistency reliability was α = .84 for the total sample, α = .86 for English responders, and α = .84 for Spanish responders.

Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977)— The CES-D measures frequency of depressive symptoms experienced during the past week from *rarely or none of the time* (*< 1 day*) to *all the time* (*5-7 days*). An abbreviated 10-item version was used in the present study, with total scores ranging from 0 to 30 (Kohout, Berkman, Evans, & Cornoni-Huntley, 1993). Internal consistency reliability for the total sample was $\alpha = .83$. The validity, reliability, and invariance of the English and Spanish translated CES-D versions that were used by HCHS/SOL have been supported (Gonzalez et al., 2015).

Spielberger Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970)—The STAI measures trait anxiety, or a general tendency to experience anxious emotion-cognition. It is a 10-item measure and respondents rate how they generally feel on a 4-point scale ranging from 1 (*almost never*) to 4 (*almost always*). Total scores range from 10 to 40. Internal consistency reliability for the total sample was $\alpha = .80$.

Speilberger Trait Anger Scale (STANG; Spielberger, Jacobs, Russell, & Crane, 1983)—The STANG measures trait anger, or a general tendency to experience feelings of anger. It is a 10-item measures and respondents rate how they generally feel on a 4-point scale ranging from 1 (*almost never*) to 4 (*almost always*). Total scores range from 10 to 40. Internal consistency reliability for the total sample was $\alpha = .85$.

Statistical Analyses

Analyses were performed in IBM Statistics Version 22.0 (SPSS, Inc., Chicago IL) and Mplus Version 7.0 (Muthén & Muthén, 2006). Means and standard deviations for continuous variables, percentages for categorical variables, and psychometric and regression analyses were computed without accounting for the weights, stratification or clustering of the sampling design. Therefore, the descriptive statistics are unweighted and represent the present sample rather than the population from which the sample was drawn.

Confirmatory factor analysis (CFA) was used to compare one-factor, two-factor, and bi-factor models at the item level in the full Sociocultural Ancillary Study sample and in the English and Spanish language samples. At the item level, skewness ranged from −.666 to .662 and kurtosis ranged from −.563 to .606. The maximum likelihood mean-adjusted estimator (MLM) was used in all analyses. In the one-factor model, all 10 PSS items loaded onto a single factor. In the oblique two-factor model, the six non-reverse worded items loaded onto one factor and the four reverse worded items loaded onto a separate factor. Whereas the factors in the two-factor model were oblique, in the bi-factor model, all 10 items loaded onto a general perceived stress factor and the four reverse worded items loaded onto an additional nuisance reverse worded factor (see **Figure 1** for two-factor and bi-factor models). Several fit indices were evaluated to determine model fit. Closer model fit was indicated by lower values on the Satorra-Bentler χ^2 likelihood ratio (S-Bχ^2; Satorra & Bentler, 2001), a standardized root-mean-square residual (SRMR) ≤ .08 (Hu & Bentler, 1999), a root-mean-square error of approximation (RMSEA) ≤ .06 (Steiger, 1990), and a comparative fit index value (CFI) ≥ .95. One-factor, two-factor, and bi-factor models are nested, and the difference in relative model fit between models was examined using ΔS-Bχ^2,

where a non-significant $\Delta\chi^2$ ($p > .05$) indicated no difference between nested models. $\Delta S\text{-}B\chi^2$ can be biased against invariance with large sample sizes (Dimitrov, 2010; MacCallum, Browne, & Cai, 2006), and ΔSRMR and ΔRMSEA were also used for relative model fit. where values < .015 indicated no difference between nested models (Chen, 2007; Cheung & Rensvold, 2002). The best-fitting model was carried forward throughout the remaining analyses.

Factorial invariance across English and Spanish versions of the PSS was examined with a series of nested models. A step-wise approach was applied where configural, metric, scalar, and residual invariance models were subsequently tested by sequentially imposing equality restrictions on parameters. Separate models for English and Spanish language responders were simultaneously estimated. For the configural model, the factor-item correspondence was tested across language responder groups to determine whether the same number of factors and the pattern of factor-item relationships were equivalent across language responders. For the metric model, the model from configural testing was used and all of the item factor loadings were constrained equal between groups to test whether items loaded similarly onto the factors across groups. For the scalar model, the model from metric testing was used and all item intercepts were constrained equal between groups to test whether item intercepts on the PSS variables are similar across groups. For the residual invariance model, the model from scalar testing was used and all item residuals were constrained equal between groups to test whether residuals on the PSS items are similar across language groups. Relative model fit was examined between each of the invariance levels.

To examine the internal consistency reliability of PSS scores we initially estimated Cronbach's alpha (α) for the full sample, and for the language respondent groups separately. A coefficient of ≥ .70 was considered to represent adequate reliability (Cortina, 1993). Convergent validity was examined by regressing each depressive symptom, trait anxiety, and trait anger scores on the PSS factors in the path model context. We expected a positive relationship with each of the validity measures with PSS scores, and negative relationships for with the reverse-worded PSS factors.

Results

Descriptive Analyses

The percentages of participants identifying with each of the Hispanic/Latino background groups (Dominican, Central American, Cuban, Mexican, Puerto Rican, South American) are reported in **Table 1**, along with other sample characteristics. A majority of the sample was female (61.9%). The majority of the present sample indicated living in the U.S. for ten years or more (76.3%) and was born outside of the U.S. (72.6%). Spanish was the most common language of response (80.9%). The average PSS score for the sample ($M = 14.82$, $SD = 6.85$) was similar to average PSS scores reported in previous studies of Spanish-speakers (González-Ramírez et al., 2013). Among English responders ($M = 16.14$, $SD = 7.06$) the total PSS score was significantly higher than among Spanish responders ($M = 14.49$, $SD = 6.76$), $t(5174) = -6.88$, $p < .001$, Cohen's $d = .22$ for a small effect size. Means and standard deviations of each PSS item for the full sample are presented in **Table 2**.

CFA: One-Factor, Two-Factor, and Bi-factor Models

Fit indices for the one-factor, two-factor, and bi-factor models for the full sample are presented in **Table 3**. The S-Bχ^2 indicated closest fit for the bi-factor model. The SRMR, RMSEA, and CFI further supported the bi-factor model as the closest fit. The ΔS-Bχ^2 indicated that the bi-factor model had better fit as compared to the one-factor model, $\chi^2(3) = 2500.74$, $p < .001$, and as compared to the two-factor model, $\chi^2(3) = 216.47$, $p < .001$. For the bi-factor model, all standardized factor loadings were statistically significant for both the non-reverse worded and reverse worded factors (all p's < .001). The bi-factor model was retained for multigroup invariance and validity analysis.

Multigroup CFA: Language Groups—A series of multigroup CFAs testing configural, metric, scalar, and residual invariance fitting the bi-factor solution to the data for both language responder groups showed good model fit at every level of invariance. The configural model indicated a bi-factor structure with all ten items loading onto the general factor and four reverse-worded items loading onto the reverse-worded factor for both English and Spanish responders. The metric model indicated that each of the item loadings onto the factors is similar across language groups, and the scalar model indicated each of the item intercepts is similar across groups. The residual model indicated the item residuals for each of the ten PSS items are also similar across groups. The descriptive fit indices (*ΔRMSEA* and *ΔSRMR*) indicated no significant differences in relative model fit between more restrictive and less restrictive models, indicating that the criteria for configural, metric, scalar, and residual invariance across languages were met. **Table 3** presents fit indices for configural, metric, scalar, and residual invariance models across languages for the bi-factor solution. **Table 4** presents factor loadings for the full sample from the bi-factor model.

Internal Consistency

Cronbach's alpha indicated good internal consistency in the full sample ($\alpha = .84$), and in English ($\alpha = .86$) and Spanish responders ($\alpha = .84$). The omega hierarchical coefficient (ω_H) is a model-based reliability index for item response data that are consistent with a bi-factor structure, and is analogous to coefficient alpha (Reise, 2012; Zinbarg, Revelle, Yovel, & Li, 2006). The degree to which the composite PSS scale score was interpretable as a measure of a single common factor for the full sample of U.S. Hispanics/Latinos was $\omega_H = .70$. Omega subscale (ω_S) is a reliability estimate for a residualized subscale that controls for the reliability due to the general factor of perceived stress. Reliability for the reverse-worded subscale was $\omega = .78$ when the effect of the general factor was included, and was $\omega_S = .68$ when the general factor was controlled. Of note, the coefficient (.78) of the reverse-worded factor was greater relative to the general factor (.70), indicating reliable variance may exist beyond the general factor.

Convergent Validity

All external correlations were significant at $p < .05$ and in the same direction with the general PSS factor (PSS) , as well as with the reverse worded PSS factor (PSS reverse). Scores on measures of each depressive symptoms, trait anxiety, and trait anger were regressed on the general perceived stress latent factor (PSS) and the nuisance reverse worded

latent factor (PSS reverse) from the bi-factor model to establish a degree of convergent validity. For depressive symptoms (N = 5,163), both PSS (β = .677, *s.e.* = .009, $p < .001$) and PSS reverse (β = −.206, *s.e.* = .013, $p < .001$) were significant predictors. For trait anxiety (N = 5,166), both PSS (β = .648, *s.e.* = .010, $p < .001$) and PSS reverse (β = −.301, *s.e.* = .013, $p < .001$) were significant predictors. For trait anger (N = 5,166), both PSS (β = .541, *s.e.* = .012, $p < .001$) and PSS reverse (β = −.057, *s.e.* = .015, $p < .001$) were significant predictors. In all cases, PSS reverse was a significant predictor of the outcome beyond PSS, with a negative coefficient. We also tested invariance in the regression weights when depressive symptoms, trait anxiety, and trait anger was each regressed on PSS and PSS reverse. Invariance of loadings was indicated in all cases ($\Delta\chi^2$, ΔRMSEA, ΔCFI, and ΔSRMR were *n.s.* in all analyses).

Discussion

The present study evaluated the factor structure, factorial invariance, internal consistency, and convergent validity of the 10-item Perceived Stress Scale in a large-scale sample of U.S. Hispanics/Latinos. Results indicated that a bi-factor model was a closer fit to the data than the two-factor or the one-factor models. The bi-factor model included a general perceived stress factor for all items and a separate factor for positively worded, reversed items only. Configural invariance indicated a bi-factor structure with all ten items loading onto the general factor and four reverse-worded items loading onto the reverse-worded factor for both English and Spanish responders. Metric invariance indicated that a unit difference in the construct of perceived stress holds the same meaning across indicators, and that relations between factors on the PSS and external variables are meaningful across groups. Scalar invariance indicated that the origin, or zero point, of the indicators is the same across language groups, and that factor mean comparisons between the groups are meaningful. The residual model indicated that item and scale reliabilities are similar across groups. The multiple-group analyses overall supported that the same general construct, and the same systematic response bias, exist in the English and Spanish language groups, and that group differences in PSS scores are attributable only to group differences in the variances of the general and reverse-worded perceived stress factors (Vandenberg & Lance, 2000). A confirmation of similarities in factor structure provides some indication of similar psychometric properties across different Hispanic/Latino groups. This work is a necessary test of the assumptions required to compare PSS scores between groups, as well as to combine English and Spanish PSS scores to obtain a more representative sample of U.S. Hispanics/Latinos.

Additional analyses indicated adequate internal consistency of the general factor, as well as of the reverse worded factor. In addition, scores from the general factor and from the reverse worded factor were robustly and positively related to each depressive symptoms, trait anxiety, and trait anger in support of convergent validity. Further, invariance testing with depressive symptoms, trait anxiety, and trait anger indicated convergent validity for the general perceived stress factor and the reverse-worded factor in the same directions. These findings indicated that reliable variance may exist beyond the general perceived stress (PSS) factor. It may be informative to interpret both the general factor and reverse-worded factor scores.

As aforementioned, past studies have supported a two-factor structure and posited that the distinction between factors is not meaningful due to the superficial nature of reversed wording (Cohen, 1988; González-Ramírez et al., 2013; Lee, 2012). However, the grouping of items onto a factor by reverse wording can violate the unidimensionality of the measurement of perceived stress and the use of a total summed PSS score. Our results directly tested the unidimensionality of the PSS and supported a bi-factor structure with the existence of two distinct latent factors that both contribute to the validity of the PSS: 1) a general perceived stress latent factor, and 2) a reverse worded latent factor. The two-factors were uncorrelated, yet the reverse worded factor added significantly to the validity of the PSS suggesting that the PSS measures perceived stress and an additional unknown construct (possibly the nuisance construct of reverse wording). Whereas past research has supported a two-factor model with no general perceived stress factor but rather with two separate factors, the present analyses support a general perceived stress factor. However, the present analyses indicate and emphasize the contribution of a reverse wording factor in measuring perceived stress.

Notably, the reverse worded PSS latent factor was related to trait anger ($\beta = -.057$) at a significantly lesser degree compared to either depressive symptoms ($\beta = -.206$) or trait anxiety ($\beta = -.301$). The CES-D and STAI include reverse worded items and the STANG contains no reverse worded items, which may underlie trait anger's reduced degree of convergent validity with the Perceived Stress Scale.

Future studies should apply appropriate modeling of PSS items to account for the variation introduced by reverse wording when studying perceived stress with the Perceived Stress Scale. Although reverse wording is intended to reduce careless and/or biased responding, it is possible that some undesired source of variation is introduced with the reverse worded format. Future measurement development should consider whether the benefit of content distraction outweighs the potential cost of undesired variance that can interfere with the measurement of a target construct.

For appropriate use of the PSS in clinical contexts, it may be important to obtain both a score based on all ten items and a score based on the four reverse worded items. Another option is to use additional measures to control for the contribution of the reverse worded factor in order to accurately measure perceived stress using the Perceived Stress Scale. Regarding the computation of a total summed PSS score, it is important to note that reverse worded item 4 and item 7 (see **Table 3** for the specific item questions) performed poorly (both items with loadings < .3) in the one-factor model for Hispanics/Latinos completing the PSS in English or Spanish. The low loadings further indicate a need to use caution in using and interpreting summed PSS scores, and a need to consider the variance associated with the reverse worded latent factor when computing and interpreting composite PSS scores.

It is also worthwhile to note that the item-level response format of a Likert-type scale of the PSS measure can be challenging when making group comparisons due to the subjective nature of the response. It has been reported that individuals of different cultural groups adopt different standards when evaluating themselves on subjective Likert scales (Heine, Lehman, Peng, & Greenholtz, 2002). For instance, although all participants in the present study

identified as Hispanic/Latino, cultural differences between those opting to complete measures in English compared to those completing measures in Spanish might lead to differences in the construct being measured. The finding that the PSS items were invariant both in terms of loadings and intercepts indicate that, with respect to U.S. Hispanics/Latinos, their language of choice did not alter the meaning of the latent construct. This important finding supports the practice of grouping English and Spanish PSS scores with U.S. Hispanic/Latino samples, thereby allowing for a more representative sample of U.S. Hispanics/Latinos in the study of perceived stress that is not restricted to Hispanics/Latinos with English proficiency.

There are some limitations to the current study that should be considered in interpreting the findings. Our examination was limited to the PSS-10 and not the original 14-item version of the measure. However, the PSS-10 has been reported to have psychometric properties that are superior to those of the 14-item version, and its use has increased internationally in recent decades (Cohen, 1988; Lee, 2012; Remor, 2006). In addition, the external criterion measures (i.e., depression, anxiety, anger) and the PSS measure were translated using novel HCHS/SOL-approved translations as described in the Method section. The CES-D has been validated (Gonzalez et al., 2015) and all measures are generalizable to the U.S. Hispanic/Latino population, however it is overall important to note that the PSS was validated against measures that are themselves novel. An additional limitation of the present study is the lack of attention to discriminant validity. It would be important to examine both the general and the reverse-scored factors of the PSS with constructs such as positive affect (Watson, Clark, & Tellegen, 1988) and/or life satisfaction (Diener, Emmons, Larsen, & Griffin, 1985) in future studies. Finally, although this is a population-based study, the sample is only representative of the four geographic regions from which it was obtained and may not be representative of Hispanics/Latinos in the U.S. as a whole.

Despite these limitations, the present study is the first to examine a bi-factor structure, factorial equivalence, and psychometric properties of the commonly used PSS in a heterogeneous, large-scale, population-based sample of Hispanics/Latinos in the United States. Researchers and clinicians working with U.S. Hispanic/Latino populations should exercise caution when using and interpreting a total summed score from English and/or Spanish-translated PSS measures. Overall, the present findings indicate that the PSS presents appropriate psychometric properties for English- and Spanish-speaking Hispanics/Latinos residing in the U.S. The use of a total summed PSS score is supported only when variance due to the reverse worded factor is taken into account when measuring perceived stress. Researchers and clinicians must be cautious when using measures that include reverse worded items.

Acknowledgement

The authors thank the staff and participants of HCHS/SOL for their important contributions. A complete list of staff and investigators has been provided by Sorlie P., et al. in Ann Epidemiol. 2010 Aug;20: 642-649 and is also available on the study website http://www.cscc.unc.edu/hchs/. The Hispanic Community Health Study/Study of Latinos was carried out as a collaborative study supported by contracts from the National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (N01-HC65233), University of Miami (N01-HC65234), Albert Einstein College of Medicine (N01-HC65235), Northwestern University (N01-HC65236), and San Diego State University (N01-HC65237). The following Institutes/Centers/Offices contribute to the HCHS/SOL through a

transfer of funds to the NHLBI: National Center on Minority Health and Health Disparities, the National Institute of Deafness and Other Communications Disorders, the National Institute of Dental and Craniofacial Research, the National Institute of Diabetes and Digestive and Kidney Diseases, the National Institute of Neurological Disorders and Stroke, and the Office of Dietary Supplements.

References

Acquadro C, Conway K, Giroudet C, Mear I. Linguistic Validation Manual for Patient-Reported Outcomes (PR0) Instruments. Hispanic Journal of Behavioral Sciences. 2006; 28:498–515.

Acquadro, C., Jambon, B., Ellis, D., Marquis, P. Language and translation issues.. In: Spilker, B., editor. Quality of life and pharamcoeconomics in clinical trials. Lippincott-Raven Publishers; Philadelphia, PA: 1996. p. 575-585.

Alegria M, Takeuchi D, Canino G, Duan N, Shrout P, Meng X, Gong F. Considering context, place and culture: The National Latino and Asian American Study. International Journal of Methods in Psychiatric Research. 2004; 13:208–220. doi:10.1002/mpr.178. [PubMed: 15719529]

Almadi T, Cathers I, Hamdan Mansour AM, Chow CM. An Arabic version of the Perceived Stress Scale: Translation and validation study. International Journal of Nursing Studies. 2012; 49:84–89. doi:10.1016/j.ijnurstu.2011.07.012. [PubMed: 21851941]

Barnette JJ. Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. Educational And Psychological Measurement. 2000; 60:361–370. doi:10.1177/00131640021970592.

Borsboom D. When Does Measurement Invariance Matter?. Medical Care. 2006; 44:S176–S181. doi: 10.1097/01.mlr.0000245143.08679.cc. [PubMed: 17060825]

Butcher, JN., Cabiya, J., Lucio, E., Garrido, M. Assessing Hispanic clients using the MMPI-2 and MMPI-A. American Psychological Association; 2007.

Carretero-Dios H, Pérez C. Normas para el desarrollo y revisión de estudios instrumentales. = Norms to development and to review instrumental studies. International Journal of Clinical And Health Psychology. 2005; 5:521–551.

Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Structural equation modeling. 2007; 14:464–504.

Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural equation modeling. 2002; 9:233–255.

Chiriboga DA, Black SA, Aranda M, Markides K. Stress and depressive symptoms among Mexican American elders. The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences. 2002; 57B:P559–P568. doi:10.1093/geronb/57.6.P559.

Cohen, S. Perceived stress in a probability sample of the United States.. In: Spacapan, S.Oskamp, S.Spacapan, S., Oskamp, S., editors. The social psychology of health. Sage Publications, Inc; Thousand Oaks, CA, US: 1988. p. 31-67.

Cohen S, Janicki-Deverts D, Miller GE. Psychological stress and disease. JAMA: Journal of The American Medical Association. 2007; 298:1685–1687. doi:10.1001/jama.298.14.1685. [PubMed: 17925521]

Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. Journal Of Health And Social Behavior. 1983; 24:385–396. doi:10.2307/2136404. [PubMed: 6668417]

Cohen S, Tyrrell DA, Smith AP. Negative life events, perceived stress, negative affect, and susceptibility to the common cold. Journal of Personality and Social Psychology. 1993; 64:131–140. doi:10.1037/0022-3514.64.1.131. [PubMed: 8421249]

Corral, I., Landrine, H. Methodological and statistical issues in research with diverse samples: The problem of measurement equivalence.. In: Landrine, H.Russo, NF.Landrine, H., Russo, NF., editors. Handbook of diversity in feminist psychology. Springer Publishing Co; New York, NY, US: 2010. p. 83-134.

Cortina JM. What is coefficient alpha? An examination of theory and applications. Journal of applied psychology. 1993; 78:98.

Derose KP, Baker DW. Limited English proficiency and Latinos' use of physician services. Medical Care Research and Review. 2000; 57:76–91. doi:10.1177/107755870005700105. [PubMed: 10705703]

Diener ED, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. Journal of personality assessment. 1985; 49:71–75. [PubMed: 16367493]

Dimitrov DM. Testing for factorial invariance in the context of construct validation. Measurement and Evaluation in Counseling and Development. 2010; 43:121–149.

Franzini L, Ribble JC, Keddie AM. Understanding the Hispanic paradox. Ethnicity & Disease. 2000; 11:496–518.

Gallo LC, Penedo FJ, Carnethon M, Isasi CR, Sotres-Alvarez D, Malcarne VL, Talavera GT. The Hispanic community health study/study of Latinos sociocultural ancillary study: sample, design, and procedures. Ethnicity & Disease. 2014; 24:77–83. [PubMed: 24620452]

Gonzalez P, Nunez A, Merz E, Navas E, Weitzman O, et al. Gallo LC. Measurement Properties of the Center of Epidemiologic Studies-Depression Scale (CESD-10): Findings from the Hispanic Community Health Study/Study of Latinos (HCHS/SOL). Under Review in Psychological Assessment. 2015

González-Ramírez MT, Rodríguez-Ayán MN, Hernández RL. The Perceived Stress Scale (PSS): Normative data and factor structure for a large-scale sample in Mexico. The Spanish Journal of Psychology. 2013; 16:E47. [PubMed: 23866243]

Gregorich SE. Do Self-Report Instruments Allow Meaningful Comparisons Across Diverse Population Groups? Testing Measurement Invariance Using the Confirmatory Factor Analysis Framework. Medical Care. 2006; 44:S78–S94. doi:10.1097/01.mlr.0000245454.12228.8f. [PubMed: 17060839]

Hakimzadeh, S., Cohn, DV. English Usage among Hispanics in the United States. Pew Hispanic Center; 2007.

Harmon-Jones E. Anger and the behavioral approach system. Personality and Individual Differences. 2003; 35:995–1005. doi:10.1016/S0191-8869(02)00313-6.

Heine SJ, Lehman DR, Peng K, Greenholtz J. What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. Journal of Personality And Social Psychology. 2002; 82:903–918. doi:10.1037/0022-3514.82.6.903. [PubMed: 12051579]

Hewitt PL, Flett GL, Mosher SW. The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample. Journal of Psychopathology and Behavioral Assessment. 1992; 14:247–257.

Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55. doi: 10.1080/10705519909540118.

Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association. 1995; 90:773–795.

Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D Depression Symptoms Index. Journal of Aging and Health. 1993; 5:179–193. [PubMed: 10125443]

LaVange LM, Kalsbeek WD, Sorlie PD, Avilés-Santa LM, Kaplan RC, Barnhart J, Elder JP. Sample design and cohort selection in the Hispanic Community Health Study/Study of Latinos. Annals of Epidemiology. 2010; 20:642–649. [PubMed: 20609344]

Lee EH. Review of the psychometric evidence of the perceived stress scale. Asian Nursing Research. 2012; 6:121–127. [PubMed: 25031113]

Llorente E, Warren CS, de Eulate LP, Gleaves DH. A Spanish version of the Sociocultural Attitudes Towards Appearance Questionnaire- 3 (SATAQ- 3): Translation and psychometric evaluation. Journal of Clinical Psychology. 2013; 69:240–251. [PubMed: 23307143]

MacCallum RC, Browne MW, Cai L. Testing differences between nested covariance structure models: Power analysis and null hypotheses. Psychological Methods. 2006; 11:19. [PubMed: 16594765]

Mimura C, Griffiths P. A Japanese version of the Perceived Stress Scale: Translation and preliminary test. International Journal of Nursing Studies. 2004; 41:379–385. doi:10.1016/j.ijnurstu.2003.10.009. [PubMed: 15050849]

Muthén, LK., Muthén, BO. Mplus [computer program]. Muthén & Muthén; Los Angeles, CA: 2006.

Passel, JS., Cohn, D., Lopez, MH. Hispanics account for more than half of nation's growth in past decade. Pew Hispanic Center; Washington, DC: 2011.

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

Raftery AE. Bayesian model selection in social research. Sociological Methodology. 1995; 25:111–164.

Ramírez MG, Hernández RL. Factor structure of the Perceived Stress Scale (PSS) in a sample from Mexico. The Spanish Journal of Psychology. 2007; 10:199–206. [PubMed: 17549893]

Reis RS, Hino AF, Rodriguez-Añez CR. Perceived Stress Scale: Reliability and validity study in Brazil. Journal of Health Psychology. 2010; 15:107–114. [PubMed: 20064889]

Reise SP, Morizot J, Hays RD. The role of the bifactor model in resolving dimensionality issues in health outcomes measures. Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation. 2007; 16:19–31. doi:10.1007/s11136-007-9183-7.

Remor E. Psychometric Properties of a European Spanish Version of the Perceived Stress Scale (PSS). The Spanish Journal of Psychology. 2006; 9:86–93. [PubMed: 16673626]

Remor E, Carrobles JA. Versión Española de la Escala de Estrés Percibido (PSS- 14): Estudio psicométrico en una muestra VIH+. = Spanish version of the Perceived Stress Scale (PSS-14): Psychometric study in a HIV+ sample. Ansiedad Y Estrés. 2001; 7:195–201.

Roberti JW, Harrington LN, Storch EA. Further Psychometric Support for the 10-Item Version of the Perceived Stress Scale. Journal of College Counseling. 2006; 9:135–147. doi:10.1002/j.2161-1882.2006.tb00100.x.

Salman, E. Research Edition Spanish Translation: State-Trait Anxiety Inventory (Form Y). Mind Garden, Inc; Redwood City, CA: 1998.

Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. Psychometrika. 2001; 66:507–514. doi:10.1007/BF02296192.

Schumann JH. Research on the acculturation model for second language acquisition. Journal of Multilingual & Multicultural Development. 1986; 7:379–392.

Sorlie PD, Avilés-Santa LM, Wassertheil-Smoller S, Kaplan RC, Daviglus ML, Giachello AL, Heiss G. Design and implementation of the Hispanic Community Health Study/Study of Latinos. Annals of Epidemiology. 2010; 20:629–641. [PubMed: 20609343]

Spielberger, CD., Gorsuch, RL., Lushene, RE. The State-Trait Anxiety Inventory: Test manual. Consulting Psychologist Press; Palo Alto, CA: 1970.

Spielberger CD, Jacobs G, Russell S, Crane RS. Assessment of anger: The state-trait anger scale. Advances in Personality Assessment. 1983; 2:159–187.

Steiger JH. Structural model evaluation and modification: An interval estimation approach. Multivariate Behavioral Research. 1990; 25:173–180. [PubMed: 26794479]

Suro, R., Singer, A. Latino growth in metropolitan America: Changing patterns, new locations. Brookings Institution, Center on Urban and Metropolitan Policy in collaboration with the Pew Hispanic Center; 2002.

Taylor JM. Psychometric analysis of the Ten-Item Perceived Stress Scale. Psychological Assessment. 2015; 27:90–101. doi:10.1037/a0038100. [PubMed: 25346996]

Trujillo HM, González-Cabrera JM. Propiedades psicométricas de la versión española de la ‘Escala de Estrés Percibido’ (EEP). = Psychometric properties of the Spanish version of the Perceived Stress Scale (PSS). Psicología Conductual Revista Internacional De Psicología Clínica De La Salud. 2007; 15:457–477.

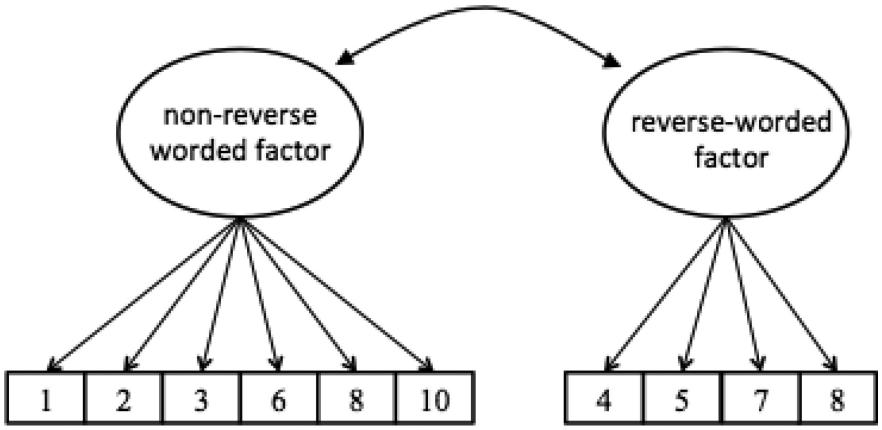
van Herk H, Poortinga YH, Verhallen TM. Response Styles in Rating Scales: Evidence of Method Bias in Data From Six EU Countries. Journal of Cross-Cultural Psychology. 2004; 35:346–360. doi:10.1177/0022022104264126.

Vandenberg RJ, Lance CE. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. Organizational Research Methods. 2000; 3:4–69. doi:10.1177/109442810031002.

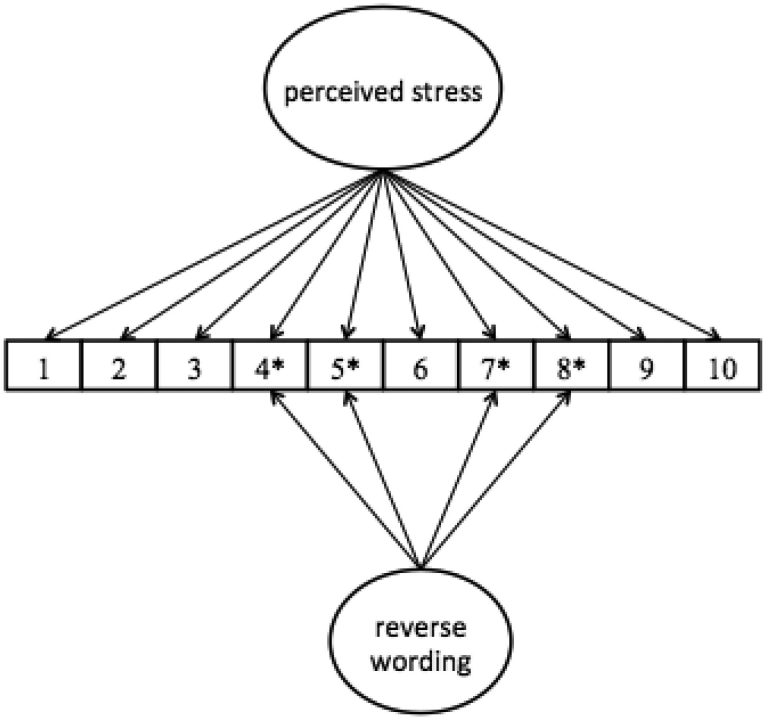
Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. Journal of personality and social psychology. 1988; 54:1063. [PubMed: 3397865]

Wong N, Rindfleisch A, Burroughs JE. Do Reverse worded Items Confound Measures in Cross-Cultural Consumer Research? The Case of the Material Values Scale. Journal of Consumer Research. 2003; 30:72–91. doi:10.1086/374697.

Zinbarg RE, Yovel I, Revelle W, McDonald RP. Estimating Generalizability to a Latent Variable Common to All of a Scale's Indicators: A Comparison of Estimators for ωh. Applied Psychological Measurement. 2006; 30:121–144. doi:10.1177/0146621605278814.



(I)



(II)

Figure 1.

I is a diagram of the two-factor model. II is a diagram of the bi-factor model. PSS items load onto either the non-reverse worded or reverse worded factor in the two-factor model. All items load into the general perceived stress factor, and the four reverse worded items also load onto the reverse worded nuisance factor in the bi-factor model. * indicates reverse worded item.

Table 1

Unweighted Sample Characteristics for the Total Sample (N = 5,176)

Variable	*n*	%
Women	3206	61.9
Hispanic/Latino Background (*n* = 5,172)		
Dominican	527	10.2
Central American	536	10.4
Cuban	742	14.3
Mexican	2034	39.3
Puerto Rican	857	16.6
South American	341	6.6
More than one heritage	122	2.4
Other	13	.3
Married or cohabitating (*n* = 5,167)	2580	49.8
Household Income (*n* = 4,753)		
<$10,000	857	16.6
$10,001-20,000	1620	31.3
$20,001-40,000	1549	29.9
$40,001-75,000	553	10.7
>$75,000	174	3.4
Education (*n* = 5,167)		
Less than high school	1859	35.9
High school/GED	1350	26.1
More than high school/GED	1958	37.8
Years in U.S. (*n* = 5,165)		
<10	1215	23.5
≥10	3950	76.3
U.S. born (*n* = 5,173)	1416	27.4
Spanish language (*n* = 5,176)	4170	80.6

	M	SD
Age (*n* = 5,176)	46.49	13.66
PSS total (*n* = 5,176)	14.82	6.85
Depression (*n* = 5,110)	7.79	6.00
Anxiety (*n* = 5,074)	17.78	5.32
Anger (*n* = 5,120)	16.83	5.36

Note. PSS = Perceived Stress Scale.

Table 2

Perceived Stress Scale (PSS) Item-Level Descriptive Statistics for Full Sample

PSS Item		M	SD
In the last month, how often have you...			
PSS 1	been upset because of something that happened unexpectedly?	1.43	1.10
PSS 2	felt that you were unable to control the important things in your life?	1.32	1.12
PSS 3	felt nervous and “stressed”?	1.87	1.13
PSS 4	felt confident about your ability to handle your personal problems?	2.72	1.05
PSS 5	felt that things were going your way?	2.56	1.01
PSS 6	found that you could not cope with all the things you had to do?	1.61	1.07
PSS 7	been able to control irritations in your life?	2.58	1.02
PSS 8	felt that you were on top of things?	2.56	1.01
PSS 9	been angered because of things that were outside of your control?	1.60	1.04
PSS 10	felt difficulties were piling up so high that you could not overcome them?	1.39	1.12

Note. N = 5,176. PSS = Perceived Stress Scale.

Table 3

Goodness-of-Fit Statistics for CFA Models of the Perceived Stress Scale

Model	S-Bχ2	*df*	*p*	RMSEA	SRMR	CFI	Ref. model no.	ΔRMSEA	ΔSRMR
1. One-factor	3453.56	35	<.001	.137	.105	.780			
2. Two-factor	721.09	34	<.001	.062	.044	.956	1	.075	.061
3. Bi-factor	486.72	31	<.001	.053	.024	.971	2	.009	.020
English	75.852	31	<.001	.038	.022	.986			
Spanish	439.182	31	<.001	.056	.025	.967			
4. Configural	514.002	62	<.001	.053	.024	.971			
5. Metric	557.321	74	<.001	.050	.030	.969	4	−.003	.006
6. Scalar	742.065	82	<.001	.056	.036	.958	5	.006	.006
7. Residual	738.224	92	<.001	.052	.036	.958	6	−.004	--

Note. N = 5,176. CFA = confirmatory factor analysis; S-Bχ2 = Satorra-Bentler χ2 statistic; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CFI = comparative fit index. Invariance results (4-7) based on the bi-factor model (model 3).

Table 4

Factor Loadings from the Bi-factor Model

Bifactor-general loading			Bifactor-RW loading	
Item	**Unstandardized**	**Standardized**	**Unstandardized**	**Standardized**
1	1.00	.72	--	--
2	1.09	.77	--	--
3	1.07	.75	--	--
4*	−.26	−.19	1.00	.62
5*	−.52	−.39	.93	.61
6	.79	.58	--	--
7*	−.23	−.17	.92	.59
8*	−.55	−.41	.95	.62
9	.92	.69	--	--
10	1.07	.76	--	--

Note. N = 5,176, Spanish n = 4,045, English n = 1,131. The unstandardized factor loading for the first item was fixed to one to set the metric for the latent variable, all *ps* < .001.

RW = reverse worded.

* indicates reverse worded item.