Asian Nursing Research 6 (2012) 121–127

Contents lists available at SciVerse ScienceDirect

Asian Nursing Research

journal homepage: www.asian-nursingresearch.com

Invited Article

# Review of the Psychometric Evidence of the Perceived Stress Scale

Eun-Hyun Lee, RN, PhD *

*Keywords:*
stress
psychometrics
questionnaire

S U M M A R Y

*Purpose:* The purpose of this study was to review articles related to the psychometric properties of the Perceived Stress Scale (PSS).
*Methods:* Systematic literature searches of computerized databases were performed to identify articles on psychometric evaluation of the PSS.
*Results:* The search finally identified 19 articles. Internal consistency reliability, factorial validity, and hypothesis validity of the PSS were well reported. However, the test-retest reliability and criterion validity were relatively rarely evaluated. In general, the psychometric properties of the 10-item PSS were found to be superior to those of the 14-item PSS, while those of the 4-item scale fared the worst. The psychometric properties of the PSS have been evaluated empirically mostly using populations of college students or workers.
*Conclusion:* Overall, the PSS is an easy-to-use questionnaire with established acceptable psychometric properties. However, future studies should evaluate these psychometric properties in greater depth, and validate the scale using diverse populations.


## Introduction

Stress has long been a major research concept in health science since it is linked to various health outcomes and illnesses, including cancer, diabetes, cardiovascular disease, asthma, and rheumatoid arthritis (Cohen, Janicki-Deverts, & Miller, 2007; Johnson, Perry, & Rozensky, 2002). The ways in which the concept of stress has been assessed in research can be classified broadly into three perspectives: (a) environmental, focusing on stressors or life events; (b) psychological, assessing subjective stress appraisal and affective reactions; and (c) biological, assessing the activation of the physiological systems involved in the stress response (Cohen & Kessler, 1997; Kopp et al., 2010).

The Perceived Stress Scale (PSS; Cohen, Kamarch, & Mermelstein, 1983) is one of the more popular tools for measuring psychological stress. It is a self-reported questionnaire that was designed to measure "the degree to which individuals appraise situations in their lives as stressful" (Cohen et al., 1983, p. 385). The PPS items evaluate the degree to which individuals believe their life has been unpredictable, uncontrollable, and overloaded during the previous month. The assessed items are general in nature rather than focusing on specific events or experiences.

There are three versions of the PSS. The original instrument is a 14-item scale (PSS-14) that was developed in English (Cohen et al., 1983), with 7 positive items and 7 negative items rated on a 5-point Likert scale. Five years after the introduction of the PSS-14, it was shortened to 10 items (PSS-10) using factor analysis based on data from 2,387 U.S. residents. A four-item PSS (PSS-4) was also introduced as a brief version for situations requiring a very short scale or telephone interviews (Cohen & Williamson, 1988). According to Cohen's Laboratory for the Study of Stress, Immunity, and Disease (2012), the PSS is currently translated into 25 languages other than English.

While the psychometric properties of the PSS have been evaluated in various cultures and countries, its psychometric properties have never been reviewed across studies. The purpose of this paper was therefore to review the psychometric properties of the PSS.

## Methods

### *Searching and study selection*

Three computerized databases (PubMed, SCOPUS, and CINAHL) were searched (up to June 2012) to find relevant articles for this study. The index terms used were "stress", "measurement", "questionnaire", "psychometrics", "reliability", and "validity". The references provided in the included studies were also screened manually for additional relevant articles. A study was included if it was reported in a full-text, original article, and measured stress in humans using the PSS, with a focus on psychometric evaluation. Only English language articles were included.

* Correspondence to: Eun-Hyun Lee, RN, PhD, Graduate School of Public Health, Ajou University, San 5, Wonchon-Dong, Yeongtong-Gu, Suwon 443-721, South Korea
*E-mail address:* ehlee@ajou.ac.kr

1976-1317/$ – see front matter 
http://dx.doi.org/10.1016/j.anr.2012.08.004

*Psychometric properties evaluated*

Reliability which is the degree to which the measurement is free from measurement error. Internal consistency reliability (homogeneity of multi-item scales) and test-retest reliability (score consistency between two time points) were extracted for reliability of the PSS in this review paper. Validity, which is the degree to which the scale measures the constructs it purports to measure. Four types of validity were extracted: factorial (or structural) validity, criterion validity (relation to a gold standard), hypothesis testing (relation to other measures in a way an investigator would expect), and known-groups validity (anticipation of differences in scores between a certain specific known group) (Mokkink et al., 2010). In addition, general characteristics of the PSS including version used, language used, and method of translation into another language were examined. Population studied including the type of subjects and the sample size were examined.

## Results

*Study selection*

The initial search of the three computerized databases and the manual search yielded 654 records, with 36 of the studies being potentially suitable for inclusion. The full texts of all 36 of these potential studies were screened; 14 studies were found to be duplicated, and 3 were in languages other than English (i.e., Portuguese, French, and Spanish). The inclusion criteria were ultimately met by 19 studies (Figure 1).

*General characteristics of the PSS and the study sample*

Table 1 presents the characteristics of the PSS and samples of the 19 selected studies. PSS-14, PSS-10, and PSS-4 were psychometrically evaluated in 12, 12, and 6 studies, respectively. They were evaluated simultaneously in five studies, while PSS-14 and PSS-10 were evaluated simultaneously in one study.

In 13 studies, the PSS was evaluated psychometrically after translation from English into nine other languages (Japanese, Spanish, Turkish, Portuguese, Chinese, Thai, Arabic, Greek, and French). A translation and back-translation technique was implemented in nine of these studies to change the PSS from English into their target languages, while in three studies a previously translated French or Spanish version was used. The method of translation was not stated for the remaining study.

The most common cohort in these psychometric studies of the PSS comprised college students (Table 2), and the sample sizes in the studies ranged from 60 to 2,387 (Table 1).

*Psychometric properties*

The psychometric properties of each study are summarized in Table 3. Cronbach's alpha is a measure of internal consistency reliability, with a value >.70 considered a minimum measure of internal consistency (Nunnally & Bernstein, 1994). Cronbach's alpha of the PSS-14 was >.70 in 11 of the 12 studies in which this version was evaluated, and was not evaluated in the 12th study. Cronbach's alpha of the PSS-10 was evaluated at >.70 in all 12 studies in which it was used. However, the reported Cronbach's alpha was <.70 in half of the six studies in which the PSS-4 was evaluated.

Test-retest reliability was evaluated using a correlation coefficient, such as Pearson's, Spearman's, or the intraclass correlation coefficient (ICC); coefficient values >.70 are usually recommended (Terwee et al., 2007). The test-retest reliability for the PSS-14 was assessed in three studies, all except one of which met the criterion of a coefficient value of >.70. In that study the test-retest reliability of the PSS-14 was evaluated only after a 6-week interval, while in the other two it was evaluated between 2 days and 4 weeks (Cohen et al., 1983). The test-retest reliability of the PSS-10 was assessed in four studies, and met the criterion of >.70 in all cases. None of the studies in which the PSS-4 was evaluated assessed its reliability. The interval between the first and second administrations of the PSS for assessment of the test-retest reliability ranged from 2 days to 6 weeks.

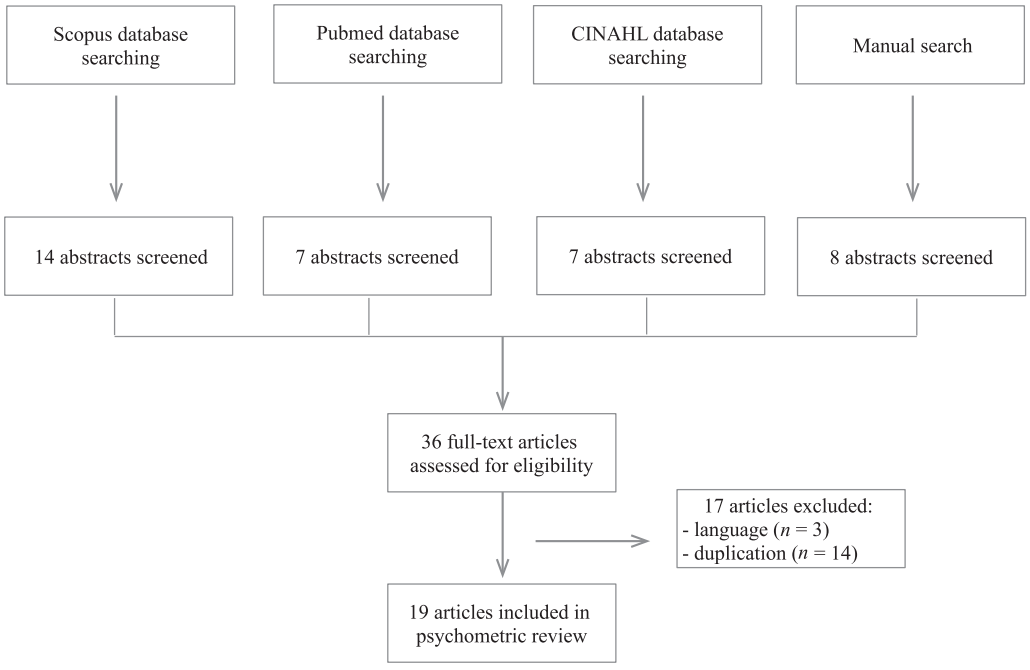


*Figure 1.* Flow chart of the search protocol.

**Table 1** *General Characteristic of Selected Studies*

| Reference | PSS version | Language/country | Translation | Population (sample size) |
|---|---|---|---|---|
| Cohen, Kamarck, & Mermelstein (1983) | PSS-14 | English/USA | N/A | Two samples of college students ($n = 332$ & $n = 114$)<br>Enrolled in a smoking-cessation program ($n = 64$) |
| Cohen & Williamson (1988) | PSS-14<br>PSS-10<br>PSS-4 | English/USA | N/A | Stratified random sampling ($N = 2,387$) |
| Pbert, Doerfler, & DeCosimo (1992) | PSS-14 | English/USA | N/A | Enrolled in health-promotion program ($n = 59$)<br>Enrolled in cardiac-rehabilitation program ($n = 41$) |
| Hewitt, Flett, & Mosher (1992) | PSS-14 | English/Canada | N/A | Psychiatric patients ($N = 96$) |
| Mimura & Griffiths (2004) | PSS-14 | Japanese & English/Japan & UK | Translation and back-translation for Japanese version | Students undertaking postgraduate programs:<br>Japanese students ($n = 23$)<br>UK students ($n = 38$) |
| Remor (2006) | PSS-14<br>PSS-10 | European Spanish/ Spain | Translation and back-translation | Adults ($N = 440$): Parents of chronically ill children, substance abusers, healthy undergraduate students, and HIV-positive patients |
| Roberti, Harrington, & Storch (2006) | PSS-10 | English/USA | N/A | Undergraduate students ($N = 285$) |
| Ramírez & Hernández (2007) | PSS-14 | Spanish/Mexico | Previously translated Spanish version | Psychology students ($N = 365$) |
| Mitchell, Crane, & Kim (2008) | PSS-14<br>PSS-10<br>PSS-4 | English/USA | N/A | Survivors within 1 month of the death by suicide of a family member or significant other ($N = 60$) |
| Örücü & Demir (2009) | PSS-10 | Turkish/Turkey | Translation and back-translation | University students ($N = 508$) |
| Reis, Hino, & Rodriguez-Añez (2010) | PSS-10 | Brazilian Portuguese/Brazil | Translation and back-translation | Full-time teachers ($N = 793$) |
| Leung, Lam, & Chan (2010) | PSS-14<br>PSS-10<br>PSS-4 | Chinese/China (Hong Kong) | Unclear | Cardiac patients who smoked ($N = 1,800$) |
| Wongpakaran & Wongpakaran (2010) | PSS-10 | Thai/Thailand | Translation and back-translation | Adults ($n = 479$): medical students ($n = 368$), patients ($n = 111$) |
| Chaaya, Osman, Naassan, & Mahfoud (2010) | PSS-10 | Arabic/Qatar | Translation and back-translation | Women ($N = 268$): pregnant ($n = 113$), postpartum ($n = 97$), university students ($n = 58$) |
| Andreou et al. (2011) | PSS-14<br>PSS-10<br>PSS-4 | Greek/Greece | Translation and back-translation | Adults recruited from hospitals, financial/tax offices, or universities ($N = 941$) |
| Wang et al. (2011) | PSS-10 | Chinese/China | Translation and back-translation | Policewomen ($N = 240$) |
| Karam et al. (2012) | PSS-4 | French and English/ USA and Canada | Previously translated French version | Pregnant women ($N = 217$) |
| Almadi, Cathers, Mansour, & Chow (2012) | PSS-14 | Arabic/Jordan | Translation and back-translation | Teachers and technical workers ($N = 90$) |
| Lesage, Berjot, & Deschamps (2012) | PSS-14<br>PSS-10<br>PSS-4 | French/France | Previously translated French version | Workers ($N = 501$) |

PSS = Perceived Stress Scale.

The factorial (structural) validity of a construct to be measured is usually performed using factor analysis. Exploratory factor analysis for the PSS-14 and PSS-10 indicated that a two-factor structure was more dominant than a one-factor structure (Tables 3 and 4). This was confirmed by the findings of confirmatory factor analysis. However, in many of the studies, the two-factor structure for the PSS-14 accounted for less than 50% of the total variance, which is the minimum percentage of cumulative variance extracted by successive factors (Pett, Lackey, & Sullivan, 2003).

**Table 2** *Populations of All Studies Reviewed*

| Population | No. of studies (%) |
|---|---|
| General population | 1(5.26) |
| Adults (recruited from hospital words, finance or tax offices, universities) | 1(5.26) |
| College or postgraduate students | 4(21.05) |
| Psychiatry patients or survivors of death by suicide of significant persons | 2(10.52) |
| Cardiac patients | 1(5.26) |
| Women (pregnant, postpartum, students, policewomen) | 3(15.78) |
| Adults (patients, students) | 2(10.52) |
| Workers or teachers | 3(15.78) |
| Enrolled in health-promotion program (college students) | 2(10.52) |

To establish criterion validity, the scores of an instrument should be strongly correlated (i.e., $r > .70$) with the scores of its gold-standard instrument (Terwee et al., 2007). The criterion validity of PSS was evaluated in a few studies, of which the PSS was strongly correlated with only the mental component of health status as measured by the Medical Outcomes Study–Short Form 36 (Ware, Snow, Kosinski, & Grandek, 1993).

Hypothesis testing revealed that the PSS was either moderately or strongly correlated with the hypothesized emotional variables, such as depression or anxiety, as measured using the Center for Epidemiologic Studies Depression Scale (Radloff, 1977), Inventory to Diagnose Depression (Zimmerman & Coryell, 1987), Beck Depression Inventory (Beck, Steer, & Garbin, 1988), Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), State-Trait Anxiety Inventory (Spielberger, 1983), Escala de Cansancio Emocional (Scale of Emotional Exhaustion; Ramos, Manga, & Moran, 2005), General Health Questionnaire (Goldberg & Williams, 1991), Edinburgh Postnatal Depression Scale (Cox, Holden, & Sagovsky, 1987), Thai Depression Inventory (Lotrakul & Sukanich, 1999), and Depression Anxiety Stress Scale-21 (Lyrakos, Arvaniti, Smyrnioti, & Kostopanahiotou, 2011).

The known-groups validity of the PSS was assessed using general characteristics related groups of participants. As might be

**Table 3** *Summary of PSS Psychometric Properties*

| Reference | PSS version | Cronbach's alpha | Test-retest reliability | Factorial validity | Criterion validity | Hypothesis testing | Known-groups validity |
|---|---|---|---|---|---|---|---|
| Cohen, Kamarck, & Mermelstein (1983) | PSS-14 | .84–.86 for three samples | $r = .85$ for 2-day interval<br>$r = .55$ for 6-week interval | — | CSLES ($r = .35$–$.49$) | CES-D ($r = .76$–$.65$)<br>CHIPS ($r = .52$–$.70$)<br>Health center utilization ($r = .11$–$.20$)<br>SADS ($r = .37$–$.48$) | Gender: Not significant<br>Age: Not significant |
| Cohen & Williamson (1988) | PSS-14 | .75 | — | EFA with varimax rotation: Two-factor structure accounted for41.6% of variance | Other measures of appraisal of stress: Moderately to weakly correlated | PSS-14: Age ($r = -.13$)<br>PSS-10: Age ($r = -.13$)<br>PSS-4; Age ($r = -.11$)<br>PSS scores were moderately or weakly correlated with self-reported physical illness<br>PSS scores were weakly correlated with self-reported health behaviors.<br>Life satisfaction ($r = .47$) | Income: Perceived higher stress for lower income<br>Education: Lower scores on PSS for more education<br>Race: Lower scores on PSS for whites<br>Marital: Lower scores on PSS for married and cohabitating subjects<br>Employment: Lower scores on PSS for the employed |
| | PSS-10 | .78 | — | EFA with varimax rotation: Two-factor structure accounted for48.9% of variance) | | | |
| | PSS-4 | .60 | — | EFA with varimax rotation: One-factor structure accounted for 45.6% of variance | | | |
| Pbert, Doerfler, & DeCosimo (1992) | PSS-14 | — | — | — | LES ($r = .25$–$.36$) | IDD ($r = .63$–$.67$)<br>SAI ($r = .80$)<br>CHIPS ($r = .42$–$.54$) | Gender: Not significant |
| Hewitt, Flett, & Mosher (1992) | PSS-14 | .80 | — | EFA with varimax rotation: Two-factor structure (accounted for 46.6% of variance) | — | BDI ($r = .57$) | Gender: Higher scores on PSS for women |
| Mimura & Griffiths (2004) | PSS-14 | .81–.88 | — | EFA with varimax rotation: Two-factor structure (accounted for 50–53.2% of variance) | — | — | — |
| Remor (2006) | PSS-14 | .81 | $r = .73$ for 2-week interval | — | — | HADS ($r = .71$–$.64$) | Gender: Higher scores on PSS for women |
| | PSS-10 | .82 | $r = .77$ for 2-week interval | — | — | HADS ($r = .72$–$.66$) | Age ($r = -.18$, for both PSS-14 and PSS-10)<br>Population: Higher scores on PSS for parents of chronically ill children |
| Roberti, Harrington, & Storch (2006) | PSS-10 | .89 | — | EFA with oblique rotation: Two-factor structure (accounted for 61.9% of variance). The two factors were strongly correlated ($r = .65$)<br>CFA: Indicated an adequate fit | — | MHLC Form A ($r = .18$–$.20$)<br>STAI-T ($r = .73$)<br>Weak or no correlations with SSS-V, SCSRFQ-SF, OA, and RA (divergent validity) | — |
| Ramírez & Hernández (2007) | PSS-14 | .83 | — | EFA with oblimin rotation: Two-factor structure (accounted for 48.02% of variance)<br>CFA indicated an adequate fit | — | BDI ($r_s = .553$)<br>ECE ($r_s = .521$) | Gender: No difference |
| Mitchell, Crane, & Kim (2008) | PSS-14 | .89 | — | EFA: One-factor structure accounted for 51% of variance | MOS-SF36: Mental component ($r = -.65$), physical component ($r = -.27$) | IES ($r = .51$)<br>PTS-AS ($r = .68$) | — |
| | PSS-10 | .91 | — | EFA: One-factor structure accounted for 56.6% of variance | MOS-SF36: Mental component ($r = -.70$), physical component ($r = -.21$) | IES ($r = .54$)<br>PTS-AS ($r = .69$) | — |
| | PSS-4 | .82 | — | EFA: One-factor structure accounted for 65.2% of variance | MOS-SF36: Mental component ($r = -.70$), physical component ($r = -.23$) | IES ($r = .58$)<br>PTS-AS ($r = .70$) | — |
| Örücü & Demir (2009) | PSS-10 | .84 | — | EFA with a varimax rotation: Two-factor structure accounted for 56.24% of variance. CFA indicated almost a good fit. | — | GHQ ($r = .61$) | — |
| Reis, Hino, & Rodriguez-Añez (2010) | PSS-10 | .87 | ICC with 24 teachers (.86 for 7-day interval) | EFA with a varimax rotation: Two-factor structure accounted for 56.87% of variance. CFA indicated an adequate fit | — | HEALTH ($r = -.37$)<br>MENT ($r = -.32$)<br>PHYS ( $= -.24$) | — |
| Chaaya, Osman, Naassan, & Mahfoud (2010) | PSS-10 | .74 | $r_s = .74$ for 1-week interval | EFA with varimax rotation: Two-factor structure accounted for 47.3% of variance | — | GHQ-12 ($r_s = .59$)<br>EPDS ($r_s = .49$)<br>Life events ($r_s = .30$) | — |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Leung, Lam, & Chan (2010) | PSS-14 | .85 | — | CFA indicated an adequate fit | — | FTND ($r = .11$)<br>Confidence in not smoking again ($r = -.19$)<br>Perceived health status ($r = -.16$)<br>Anxiety ($r = .18$)<br>Depression ($r = .22$) | Gender: Higher scores on PSS-14 for women |
| | PSS-10 | .83 | — | CFA indicated an adequate fit | — | FTND ($r = .11$)<br>Confidence in not smoking again ($r = -.19$)<br>Perceived health status ($r = -.17$)<br>Anxiety ($r = .19$)<br>Depression ($r = .24$) | Gender: Higher scores on PSS-10 for women |
| | PSS-4 | .67 | — | CFA indicated an adequate fit | — | FTND ($r = .10$)<br>Confidence in not smoking again ($r = -.18$)<br>Perceived health status ($r = -.19$)<br>Anxiety ($r = .19$)<br>Depression ($r = .24$) | Gender: Higher scores on PSS-4 for women |
| Wongpakaran & Wongpakaran (2010) | PSS-10 | .80–.84 | ICC (.72–.88) for 4-week interval | EFA with a maximum likelihood method: Two-factor structure accounted for 66.47% of variance.<br>CFA indicated an adequate fit | — | STAI ($r = .60$)<br>TDI ($r = .55$)<br>RSES ($r = -.46$) | — |
| Andreou et al. (2011) | PSS-14 | .82 | — | CFA: Two-factor models fit well for PSS-10 and PSS-4 and marginally for PSS-14. One-factor models did not fit significantly for PSS-14 and PSS-10 | — | Moderate correlation with subscales of DASS-21 | Gender: Higher scores on PSS-14 and PSS-10 for women.<br>Marital status: Higher scores on PSS-14 and PSS-10 for divorcees and widows than for married and single women |
| | PSS-10 | .82 | — | | — | | |
| | PSS-4 | .68 | — | | — | | |
| Wang et al. (2011) | PSS-10 | .86 | — | EFA with varimax rotation: Two-factor structure accounted for 62.41% of variance.<br>CFA indicated an adequate fit | — | BDI-II ($r = .67$)<br>BAI ($r = .58$) | — |
| Karam et al. (2012) | PSS-4 | .79 (English and French versions pooled) | — | — | PSS-10 ($r = .63$) | EPDS ($r = .67$, English and French pooled)<br>MCS SF-12 ($r = -.62$, English and French pooled) | — |
| Almadi, Cathers, Mansour, & Chow (2012) | PSS-14 | .80 | ICC = .90 for 2-week interval | EFA with varimax rotation: Two-factor structure accounted for 45.0% of variance | — | — | — |
| Lesage, Berjot, & Deschamps (2012) | PSS-14 | .84 | — | EFA with oblimin rotation: Two-factor structure accounted for 49% of variance | | | Age: Higher scores for older subjects<br>Parental status: Higher scores for workers with children |
| | PSS-10 | .83 | — | EFA with oblimin rotation: Two-factor structure accounted for 55% of variance | | | Age: Higher scores for older subjects<br>Parental status: Higher scores for workers with children Gender: Higher scores for women |
| | PSS-4 | .73 | — | EFA with oblimin rotation: One-factorstructure accounted for 55% of variance | | | Age: Higher scores for older subjects<br>Gender: Higher scores for women |

PSS = Perceived Stress Scale; CSLES = College Student Life-Event Scale; CES-D = Center for Epidemiologic Studies Depression Scale; CHIPS = Cohen-Hoberman Inventory of Physical Symptoms; SADS = Social Avoidance and Distress Scale; LES = Life Experience Scale; IDD = Inventory to Diagnose Depression; SAI = State Anxiety Inventory; BDI = Beck Depression Inventory; EFA = exploratory factor analysis; CFA = confirmatory factor analysis; HADS = Hospital Anxiety and Depression Scale; SSS-V = Sensation-Seeking Scale; MHLC Form A = Multidimensional Health Locus of Control, Form A; SCSRFQ-SF = Santa Clara Strength of Religious Faith Questionnaire–Short Form; OA = Adult Overt Aggregation Scale from the Adult Aggression Scale; RA = Relational Aggression Scale from the Adult Aggression Scale; ECE = Escala de Cansancio Emocional (Scale of Emotional Exhaustion); STAI-T = State-Trait Anxiety Inventory–Trait; ICC = intraclass correlation coefficient; MOS-SF36 = Medical Outcomes Study–Short Form 36; IES = Impact of Event Scale; PTS-AS = Post Traumatic Stress-Arousal Scale; GHQ = General Health Questionnaire; HEALTH = perceived health; MENT = perceived effect of the work on mental health; PHYS = perceived effect of the work on physical health; EPDS = Edinburgh Postnatal Depression Scale; FTND = Fagerström Test of Nicotine Dependency; STAI = State-Trait Anxiety Inventory; TDI = Thai Depression Inventory; RSES = Rosenberg Self-Esteem Scale; DASS-21 = Depression Anxiety Stress Scale-21; BDI-II = Beck Depression Inventory Revised; BAI = Beck Anxiety Inventory; MCS SF-12 = Mental Health Component Summary, Short Form.

**Table 4** *Factorial Structure of the PSS*

| PSS version | EFA | | CFA | | $n$[a] | $n$[a] |
|---|---|---|---|---|---|---|
| | Two-factor structure | One-factor structure | Two-factor structure | One-factor structure | | |
| | $n$[a] (>50%[b]) | $n$[a] (50%[b]) | $n$[a] (>50%[b]) | $n$[a] (50%[b]) | | |
| PSS-14 | 1 | 5 | 1 | 0 | 4[c] | 1[d] |
| PSS-10 | 6 | 2 | 1 | 0 | 7 | 1[d] |
| PSS-4 | 0 | 0 | 2 | 1 | 2 | 0 |

PSS = Perceived Stress Scale. EFA = exploratory factor analysis; CFA = confirmatory factor analysis.
[a] Number of studies.
[b] Percentage of total variance explained by the factor structure.
[c] Marginal fit.
[d] Does not fit one-factor structure.

expected, the PSS scores were significantly lower for groups of participants who were young, white, married, employed, earning a high income, and with parents with a smaller number of children or not having chronically ill children. However, there was an inconsistent finding with regard to gender: some studies found no gender difference (Cohen et al., 1983; Pbert, Doerfler, & DeCosimo, 1992; Ramírez & Hernández, 2007), while others found that PSS scores were higher in women than in men (Andreou et al., 2011; Hewitt, Flett, & Mosher, 1992; Lesage, Berjot, & Deschamps, 2012; Leung, Lam, & Chan, 2010; Remor, 2006).

## Discussion

This paper has reviewed studies of the psychometric properties of the PSS. It was found that the internal consistency reliability of this tool has been established, although Cronbach's alpha values obtained for the PSS-4 were only marginally acceptable. This may be attributable to the PSS-4 including fewer items than the PSS-14 and PSS-10, since Cronbach's alpha tends to increase with the number of items in an instrument (Pedhazur & Schmelkin, 1991).

The test-retest reliability of the PSS was evaluated in only six studies. Moreover, in three of these studies the Pearson's or Spearman's correlation coefficient was implemented for the test, which is a measure of association. Calculation of the ICC is a more sophisticated approach for assessing test-retest reliability when the score of an instrument is continuous, like in the PSS (Fayers & Machin, 2007). Therefore, the test-retest reliability of the PSS needs to be evaluated further using ICCs. With respect to the administration interval, the PSS demonstrated satisfactory test-retest reliability when its first and second administrations were separated by between 2 days and 4 weeks. However, when this interval was 6 weeks, the test-retest reliability was not satisfactory ($r = .55$). This may imply that the duration of the stability of PSS might be less than 6 weeks. If this is the case, clinicians or researchers using the PSS may consider reassessing the PSS score every 6 weeks. A systematic, longitudinal study of changes in PSS scores is required to further clarify this.

Examination of the factorial validities of the PSS-14 and PSS-10 demonstrated that the two-factor structure predominated, rather than one dimensionality. Mitchell, Crane, and Kim (2008) reported on the one-dimensional PSS construct based on 60 adults who had survived the death of a family member or significant other by suicide. However, the sample of that study was relatively small for a factor analysis, which may have resulted in an incorrect estimation of both the number of factors and their structure (Fayers & Machin, 2007). Even though the two-factor structure of the PSS-14 was predominant, it should be considered that most studies have shown that the two-factor structure accounts for less than 50% of the total variance in the 14 items. On the other hand, the PSS-4 structure was not consistent.

The criterion validity of the PSS was evaluated in only a few studies; the criteria used were all questionnaires. The correlation coefficients used with criterion questionnaires showed a weak to moderate association, demonstrating unsatisfactory criterion validity. In addition, it is questionable whether or not the used criteria were gold standards for the PSS. Future studies may use a biomarker of stress, such as cortisol, as a criterion variable (van Eck & Nicolson, 1994).

Hypothesis tests of the PSS consistently demonstrated a satisfactory correlation with depression or anxiety. This finding is consistent with the report of Cohen et al. (1983, p. 391) that "there is some overlap between what is measured by depressive symptomatology scales and measured by the PSS, since the perception of stress may be a symptom of depression."

For the known-groups validity test of the PSS, demographic categorical variables (e.g., marital status, educational status, gender, and having children) were used mostly without prior determined expectations or evidence. It is recommended that the known-groups validity for groups that have been previously well determined be implemented in future studies. With respect to gender, five of the studies in this review found that the PSS scores were significantly higher in women than in men. The gender-related difference in PSS scores remains a matter of debate. Some believe that it is an artifact of measurement bias, given that the women are more likely to score on the negatively worded items of the PSS (Gitchel, Roessler, & Turner, 2012), while others believe that there is a true gender difference arising from social, biological, or psychological influences (Lavoie & Douglas, 2012). Therefore, gender should be considered carefully when evaluating known-groups validity in the PSS.

According to Cohen et al. (1983), the PSS measures general stress and is thus relatively free of content that is specific to any particular population. However, the PSS has been empirically validated with populations of mainly college students or workers. It is necessary to validate the PSS with more diverse populations (e.g., specific or mixed clinical populations) and in various cultures. It has been translated into 25 languages, but some of the translated (i.e., non-English-language) forms have yet to be empirically validated. Furthermore, a multicultural psychometric evaluation of the PSS is recommended.

## Conclusion

In summary, the PSS is a short and easy to use questionnaire established with acceptable psychometric properties. However, the test-retest reliability, criterion validity, and known-groups validity of the PSS need to be evaluated further. In general, the psychometric properties of the PSS-10 are superior to those of the PSS-14. Therefore, it is recommended that the PSS-10 be used to measure perceived stress, both in practice and research. The PSS-4 is the

least effective of these tools, although as proposed by Cohen et al. (1983), it may be useful and feasible in situations where a short questionnaire is required, such as telephone interviews.

## Conflict of interest

The author declares no conflict of interest.

## Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1B5000978).

## References

Almadi, T., Cathers, I., Mansour, A. M. H., & Chow, C. M. (2012). An Arabic version of the Perceived Stress Scale: Translation and validation study. *International Journal of Nursing Studies, 49*, 84–89.

Andreou, E., Alexopoulos, E. C., Lionis, C., Varvogli, L., Gnardellis, C., Chrousos, G. P., et al. (2011). Perceived stress scale: Reliability and validity study in Greece. *International Journal of Environmental Research and Public Health, 8*, 3287–3298.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77–100.

Chaaya, M., Osman, H., Naassan, G., & Mahfoud, Z. (2010). Validation of the Arabic version of the Cohen perceived stress scale (PSS-10) among pregnant and postpartum women. *BioMed Central Psychiatry, 10*, 111.

Cohen, S., Janicki-Deverts, D., & Miller, G. E. (2007). Psychological stress and disease. *Journal of the American Medical Association, 298*, 1684–1687.

Cohen, S., Kamarch, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385–396.

Cohen, S., & Kessler, R. C. (1997). *Measuring stress: A guide for health and social scientists*. NY: Oxford University Press.

Cohen's Laboratory for the Study of Stress, Immunity and Disease. (2012). Dr. Cohen's Scales. Retrived June 11, 2012, from www.psy.cmu.edu/~scohen

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan, & S. Oskamp (Eds.), *The social psychology of health: Claremont symposium on applied social psychology*. Newbury Park, CA: Sage.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 19-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry, 150*, 782–786.

Fayers, P. M., & Machin, D. (2007). *Quality of life* (2nd ed.). West Sussex, UK: John Wiley & Sons.

Gitchel, W. D., Roessler, R. T., & Turner, R. (2012). Gender effect according to item directionality on the perceived stress scale for adults with multiple sclerosis. *Rehabilitation Counseling and Bulletin, 55*, 20–28.

Goldberg, D., & Williams, P. (1991). *A user's guide to the general health questionnaire*. Windsor, UK: NFER-Nelson.

Hewitt, P. L., Flett, G. L., & Mosher, S. W. (1992). The perceived stress scale: Factor structure and relation to depression symptoms in a psychiatric sample. *Journal of Psychological and Behavioral Assessment, 14*, 247–257.

Johnson, S. B., Perry, N. W., & Rozensky, R. H. (2002). *Handbook of clinical health psychology: Medical disorder and behavioral applications*. Washington, DC: American Psychological Association.

Karam, F., Bérard, A., Sheehy, O., Huneau, M., Briggs, G., Chambers, C., et al. (2012). Reliability and validity of the 4-item perceived stress scale among pregnant women: Results from the OTIS antidepressants study. *Research in Nursing and Health, 35*, 363–375.

Kopp, M. S., Thega, B. K., Balog, P., Stauder, A., Salaveez, G., Rózsa, S., et al. (2010). Measures of stress in epidemiological research. *Journal of Psychosomatic Research, 69*, 211–225.

Lavoie, J. A. A., & Douglas, K. S. (2012). The perceived stress scale: Evaluating configural, metric and scalar invariance across mental health status and gender. *Journalof Psychopathology and Behavioral Assessment, 34*, 48–57.

Lesage, F., Berjot, S., & Deschamps, F. (2012). Psychometric properties of the French versions of the perceived stress scale. *International journal of Occupational Medicine, 25*, 178–184.

Leung, D. Y., Lam, T., & Chan, S. S. (2010). Three versions of perceived stress scale: Validation in a sample of Chinese cardiac patients who smoke. *BioMed Central Public Health, 10*, 513.

Lotrakul, M., & Sukanich, P. (1999). Development of the Thai Depression Inventory. *Journal of the Medical Association of Thailand, 82*, 1200–1207.

Lyrakos, G. N., Arvaniti, C., Smyrnioti, M., & Kostopanahiotou, G. (2011). Translation and validation study of the depression anxiety stress scale in the Greek general population and in a psychiatric patient's sample. *European Psychiatry, 26*(Suppl. 1), 1731.

Mimura, C., & Griffiths, P. (2004). A Japanese version of the perceived stress scale: Translation and preliminary test. *International Journal of Nursing Studies, 41*, 379–385.

Mitchell, A. M., Crane, P., & Kim, Y. (2008). Perceived stress in survivors of suicide: Psychometric properties of the perceived stress scale. *Research in Nursing and Health, 31*, 576–585.

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Strarford, P., Knol, D., et al. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epistemology, 63*, 737–745.

Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Örücü, M. C., & Demir, A. (2009). Psychometric evaluation of perceived stress scale for Turkish university students. *Stress and Health, 25*, 103–109.

Pbert, L., Doerfler, L. A., & DeCosimo, D. (1992). An evaluation of the perceived stress scale in two clinical populations. *Journal of Psychopathology and Behavioral Assessment, 14*, 363–375.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.

Radloff, L. (1977). The center for epidemiological studies-depression (CES-D) scale: A self-reported depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385–401.

Ramírez, M. T., & Hernández, R. L. (2007). Factor structure of the perceived stress scale (PSS) in a sample from Mexico. *The Spanish Journal of Psychology, 10*, 199–206.

Ramos, F., Manga, D., & Moran, C. (2005). Escala de Cansancio Emocional (ECE) para estudiantes universitarios: Propiedades psicométricas y asociación.

Reis, R. S., Hino, A. A. F., & Rodriguez-Añez, C. R. (2010). Perceived stress scale: Reliability and validity study in Brazil. *Journal of Health Psychology, 15*, 107–114.

Remor, E. (2006). Psychometric properties of a European Spain version of the Perceived Stress Scale (PSS). *The Spanish Journal of Psychology, 9*, 86–93.

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further psychometric support for the 10-item version of the perceived stress scale. *Journal of Counseling Association, 9*, 135–147.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory: STAI*. Palo Alto, CA: Mind Garden.

Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*, 34–42.

van Eck, M. M., & Nicolson, N. A. (1994). Perceived stress and salivary cortisol in daily life. *Annals of Behavioral Medicine, 16*, 221–227.

Wang, Z., Chen, J., Boyd, J. E., Zhang, H., Jia, X., Qiu, J., et al. (2011). Psychometric properties of the Chinese version of the perceived stress scale in policewomen. *Public Library of Science ONE, 6*, e28610.

Ware, J. E., Snow, K. K., Kosinski, M., & Grandek, B. (1993). *The SF-36 health survey manual and interpretation guide*. Boston, MA: Nimrod press.

Wongpakaran, N., & Wongpakaran, T. (2010). The Thai version of the perceived stress scale (PSS-10): An investigation of its psychometric properties. *Biopsychosocial Medicine, 4*, 1–6.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital and anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361–370.

Zimmerman, M., & Coryell, W. (1987). The inventory to diagnose depression (IDD): A self-report scale to diagnose depression. *Journal of Consulting and Clinical Psychology, 55*, 55–59.