

Perceived Stress Scale: confirmatory factor analysis of the PSS14 and PSS10 versions in two samples of pregnant women from the BRISA cohort

Escala de Estresse Percebido: análise fatorial confirmatória das versões EEP14 e EEP10 em duas amostras de gestantes da coorte BRISA

Escala de Estrés Percibido: análisis factorial confirmatorio de las versiones EEP14 y EEP10 en dos muestras de gestantes de la cohorte BRISA

Ana Valéria Carvalho Pires Yokokura ¹
Antônio Augusto Moura da Silva ¹
Juliana de Kássia Braga Fernandes ¹
Cristina Marta Del-Ben ²
Felipe Pinheiro de Figueiredo ²
Marco Antonio Barbieri ²
Heloisa Bettiol ²

doi: 10.1590/0102-311X00184615

Abstract

This study aimed to assess the dimensional structure, reliability, convergent validity, discriminant validity, and scalability of the Perceived Stress Scale (PSS). The sample consisted of 1,447 pregnant women in São Luís (Maranhão State) and 1,400 in Ribeirão Preto (São Paulo State), Brazil. The 14 and 10-item versions of the scale were assessed using confirmatory factor analysis, using weighted least squares means and variance (WLSMV). In both cities, the two-factor models (positive factors, measuring resilience to stressful situations, and negative factors, measuring stressful situations) showed better fit than the single-factor models. The two-factor models for the complete (PSS14) and reduced scale (PSS10) showed good internal consistency (Cronbach's alpha ≥ 0.70). All the factor loadings were ≥ 0.50 , except for items 8 and 12 of the negative dimension and item 13 of the positive dimension. The correlations between both dimensions of stress and psychological violence showed the expected magnitude (0.46-0.59), providing evidence of an adequate convergent construct validity. The correlations between the scales' positive and negative dimensions were around 0.74-0.78, less than 0.85, which suggests adequate discriminant validity. Extracted mean variance and scalability were slightly higher for PSS10 than for PSS14. The results were consistent in both cities. In conclusion, the single-factor solution is not recommended for assessing stress in pregnant women. The reduced, 10-item two-factor scale appears to be more appropriate for measuring perceived stress in pregnant women.

Pregnant Women; Psychological Stress; Statistical Factor Analysis

Correspondence

A. V. C. P. Yokokura
Departamento de Saúde Pública, Universidade Federal do Maranhão.
Rua Barão de Itapary 155, São Luís, MA 65020-070, Brasil.
valeria.yokokura@hotmail.com

¹ Departamento de Saúde Pública, Universidade Federal do Maranhão, São Luís, Brasil.

² Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brasil.



Introduction

Stress is a process of body reactions to everyday situations – called stressors – that can threaten a person's well being¹. The pregnancy period is considered a phase of emotional stress due to numerous and intense feelings associated with physiological and physical changes^{2,3}. Stress may be present in more than 75% of pregnant women, and frequent exposure to long-term stressors can involve health risks^{4,5,6,7}.

Maternal exposure to prenatal stress is associated with hypertension, spontaneous abortion, and preterm delivery⁸. The fetus may show increased heart rate and impaired development, in addition to predisposition to mental diseases, allergies and asthma⁹. Stress during pregnancy affects the behavior and emotional control of the children, making them more irritable, restless and prone to attention problems^{8,10}. Because of the high prevalence of stress during pregnancy and its association with adverse effects on the health of both mother and child, it is important to detect stress early in pregnancy using valid and reliable tools.

In 1983, Cohen et al.¹¹ elaborated the *Perceived Stress Scale* (PSS), which was applied to a sample of university students in the United States. The scale is based on the observation of the extent to which an individual perceives as stressful different situations faced throughout his/her life¹². There are three versions of the scale, one with 14 items (PSS14), corresponding to its complete version, and two reduced versions, with ten (PSS10) and four items (PSS4) – the latter is more used in telephone surveys. PSS has been widely used throughout the world^{13,14,15,16,17,18,19,20} and was translated into several languages besides English: Arabic¹⁰, Chinese²¹, Thai¹⁶, Mexican Spanish¹⁸, Spanish²², Greek¹⁹, Japanese²⁰, French²³, and Brazilian Portuguese²⁴. PSS can be used for stress detection in a given population because it does not contain specific context questions^{11,25}.

In Brazil, the first study about the psychometric properties of PSS was conducted by Luft et al.²⁴ in Santa Catarina State. The scale was translated into Portuguese, and semantic, idiomatic, experimental, cultural and conceptual equivalences were evaluated. Through the principal component analysis (PCA), PSS showed good fit as a two-dimensional construct for the complete version (PSS14), and as a one-dimensional construct for the short version (PSS10).

Machado et al.¹² also found a one-factor solution similar to that of Luft et al.²⁴ for the PSS10 using PCA in a sample of school teachers in northeastern Brazil. However, Reis et al.²⁶, using PCA and confirmatory factor analysis (CFA) with the maximum likelihood (ML) estimator, found the two-factor solution to have the best fit in a sample of university teachers in southern Brazil. Recently, Faro²⁷ tested the three versions of PSS in the general population of Aracaju (Sergipe State), by CFA with the ML estimator, and found the two-factor model to be the best solution. Dias et al.²⁸ tested the one-factor model for the three versions of PSS in female university students using CFA with the ML estimator. They concluded that PSS10 presented the better fit, validity, internal consistency and stability compared to the two other versions. However, they did not test the two-factor solution.

Several validation studies of PSS have been carried out. However, there is no consensus in the literature about its dimensionality. A recent review pointed out that most of the studies highlight the two-factor model as having the best fit when compared to the one-factor model²⁹. In contrast, in Brazil, most studies found the best fit for the one-factor solution, especially for PSS10^{12,24,28}.

The lack of consensus about the PSS dimensionality may have been due to inadequate analytical strategies used in most previous studies. Most studies have examined the psychometric characteristics of the scale using PCA^{12,17,24,25,30}, which is a technique of data reduction that provides limited psychometric information and is not recommended to identify common factors and latent variables^{14,30,31}. In addition, we did not find investigations on PSS validity using CFA with the weighted least squares mean and variance adjusted (WLSMV) estimator³², which is more appropriate for categorical variables. All studies that applied CFA used the ML estimator^{14,16,18,19,21,26,27,28}.

Moreover, validation studies of PSS have been conducted predominantly on samples of university professors, students and workers²⁹, whereas there is a lack of studies examining the PSS validity among pregnant women. Only one Arabic study analyzed the psychometric properties of the scale in a small sample (n = 113) of pregnant women¹⁰.

Thus, this study aimed to evaluate the dimensional structure of two PSS versions (PSS10 and PSS14) using CFA, with the WLSMV estimator among pregnant women in two cities from two

different Brazilian regions with contrasting socioeconomic conditions. In addition, reliability, convergent validity, discriminant validity and scalability were assessed. Finally, we believe that this broader analysis may provide further evidence of PSS validity in the Brazilian context.

Methods

Study design

This is a cross-sectional study and the data were obtained from the cohort study *Etiological Factors of Preterm Birth and Consequences of Perinatal Factors for Child Health: Brazilian Birth Cohort Study of Ribeirão Preto and São Luís – BRISA* ²³. The study was conducted in two Brazilian cities located in distinct regions. Details of the methods have been published elsewhere by Silva et al. ³³.

Ribeirão Preto is located in the State of São Paulo, Southeastern Brazil, in a wealthy and industrialized region. Its Human Development Index (HDI) is 0.800, occupying the 40th place in Brazil. It had 604,682 inhabitants in 2010 and the per capita income was BRL 1,314.04 (USD 573.82). It is one of the most developed cities in Brazil, with 99% of residences receiving piped water and equipped with sanitary sewage (<http://www.atlasbrasil.org.br/2013>, accessed on 20/Jan/2015).

São Luís, the capital of the State of Maranhão, is located in the Northeastern region of the country. Its HDI is 0.768, occupying the 249th place in Brazil. It had 1,014,837 inhabitants in 2010 and the per capita income was BRL 805.36 (equivalent to USD 351.69). The city is in one of the poorest regions of the country, where only 50% of households are equipped with sanitary sewage, and 75% receive piped water (<http://www.atlasbrasil.org.br/2013>, accessed on 20/Jan/2015).

Participants

A total of 1,447 pregnant women from São Luís participated in the study, from February 2010 to June 2011. In Ribeirão Preto, the sample consisted of 1,400 pregnant women who were interviewed from February 2010 to February 2011. Both groups were convenience samples due to the difficulty of obtaining a random sample of pregnant women from the two cities.

Data collection

Previously trained undergraduate students contacted the women during the prenatal period in ultrasound clinics, prenatal outpatient clinics and public and private maternity hospitals. Those who had performed obstetric ultrasound before the 20th gestational week, when gestational age is determined more reliably, were invited to participate. Women with multiple fetuses were not included.

The interviews were held at the Clinical Research Center of the Mother-Child Unit of the University Hospital in São Luís, and at the Clinical Research Unit of the University Hospital, University of São Paulo, in Ribeirão Preto. Women completed a self-administered questionnaire about PSS during this visit when their gestational age was 22 to 25 weeks.

Measuring instrument

The PSS was the instrument used to measure stress during pregnancy, which evaluates the perception of stressful experiences in the preceding month using a Likert-type five-point scale ¹¹. PSS contains 14 items, seven of them positive and seven negative, and responses range from 0 to 4 (0 = never; 1 = almost never; 2 = sometimes; 3 = fairly often; 4 = very often).

Most questions reflect negative feelings and the inability to deal with stress, although some questions address positive emotions and the ability to act in stressful situations ¹⁰. All items were designed to identify to what extent the respondents evaluate their life as unpredictable, uncontrollable and overloaded, and included central components of the experience of stress. The version used in this study was translated and adapted to Brazilian Portuguese by Luft et al. ²⁴ and was self-administered.

Data analysis

Data analysis was conducted following the procedures here described:

(a) We first tested PSS with 14 categorical questions (PSS14) as a one-factor latent variable. Then we tested PSS14 as a two-factor latent variable, composed of two dimensions, one positive and the other negative. After this, we tested PSS10 as a one-factor solution or a two-factor solution. We used the WLSMV estimator, suitable for the analysis of categorical variables³².

(b) Global model fit was evaluated by the following fit indices: root mean square error of approximation (RMSEA) – values close to 0.06 or below indicate good fit^{34,35}; values lower than 0.08 suggest adequate model fit^{35,36}, and values ≥ 0.10 indicate poor model fit^{35,36}; comparative fit index (CFI) and Tucker-Lewis index (TLI) ≥ 0.95 suggest good model fit^{34,35}, and values from 0.90 to < 0.95 indicate acceptable model fit³⁷. The chi-square test was also calculated but not used to assess the model fit because of its sensitivity to a large sample size.

(c) Scale reliability was estimated using internal consistency indicators: Cronbach's alpha (α) and composite reliability (CR), which were considered acceptable when ≥ 0.70 ³⁸. These two measurements of internal consistency were calculated for each of the two PSS dimensions (positive and negative) to determine whether the questions corresponding to each dimension were consistently measured in the two studied samples.

(d) Convergent validity was evaluated based on the factor loadings and the average variance extracted (AVE). The latter was considered to be appropriate when ≥ 0.50 ³⁹. Factor loadings ≥ 0.50 and statistically significant were considered to be supportive of convergent validity^{35,38}. High correlations between each PSS dimension and psychological violence were used to indicate adequate convergent construct validity³⁹. Psychological violence was measured using the World Health Organization (WHO) violence against women questionnaire. Women were asked four questions⁴⁰: since you became pregnant has someone insulted, belittled or humiliated, intimidated or scared, or threatened to hurt you? Answers on a Likert scale were coded 0 = never; 1 = once; 2 = a few times; 3 = many times. A one-factor latent variable to measure psychological violence has been validated in a previous study⁴¹.

(e) Factor correlations < 0.85 between the PSS³⁹ positive and the negative factors and $\sqrt{\text{AVE}} \geq$ those correlations were considered to be supportive of adequate discriminant validity³⁸.

(f) We used the Mokken criteria to identify if some items could be excluded from the scale. Therefore, the Loevinger's H scalability index (HS) was calculated for each dimension of the scale and for each item. HS evaluates if the items measure the same latent trait and if there is item redundancy and validates the score as an ordinal measure of the latent trait. HS is considered acceptable for each item if > 0.30 . Scalability was considered weak if $0.30 \leq \text{HS} < 0.40$, medium if $0.40 \leq \text{HS} < 0.50$ and strong if ≥ 0.50 ⁴².

(g) We inspected the residual correlation matrix for values > 0.10 and calculated the modification indices⁴³.

The data from São Luís were analyzed first. Then the data from Ribeirão Preto were used as a validation sample to determine the stability of the factor solution in the two groups⁴⁴. The samples were analyzed separately following the same steps and procedures. The analyses were carried out using the Stata 11.0 statistical package (StataCorp LP, College Station, USA) and MPLUS 7.0 (Muthén & Muthén, Los Angeles, USA).

Ethical aspects

The Ethics Research Committees of the University Hospital, Federal University of Maranhão (protocol n. 4771/2008-30) and of the University Hospital, Faculty of Medicine of Ribeirão Preto (protocol n. 4116/2008) approved the study. All subjects gave written informed consent to participate.

Results

Table 1 presents the results of the CFA of the 14- and 10-item versions of the PSS for São Luís and Ribeirão Preto. The two-factor models showed better fit indices when compared to the one-factor

Table 1

Results of the confirmatory factor analysis of the *Perceived Stress Scale* – PSS14 and PSS10 versions. São Luís, Maranhão State, and Ribeirão Preto, São Paulo State, Brazil, 2010/2011.

Models	São Luís					Ribeirão Preto				
	χ^2	p-value	RMSEA (90%CI)	CFI	TLI	χ^2	p-value	RMSEA (90%CI)	CFI	TLI
PSS14										
One-factor	1,015.18	< 0.001	0.092 (0.091-0.103)	0.91	0.89	1,386.25	< 0.001	0.110 (0.105-0.115)	0.89	0.87
Two-factor	511.68	< 0.001	0.063 (0.058-0.068)	0.95	0.95	840.97	< 0.001	0.085 (0.080-0.090)	0.93	0.92
PSS10										
One-factor	551.05	< 0.001	0.101 (0.094-0.108)	0.92	0.90	641.48	< 0.001	0.111 (0.104-0.119)	0.92	0.90
Two-factor	240.17	< 0.001	0.065 (0.057-0.073)	0.97	0.96	322.33	< 0.001	0.078 (0.070-0.086)	0.96	0.95

90%CI: 90% confidence interval; CFI: comparative fit index; RMSEA: root mean square error of approximation; TLI: Tucker Lewis index.

models in both cities. In São Luís, the two-factor PSS14 and PSS10 models presented good fit (PSS14: RMSEA = 0.063; CFI = 0.95; TLI = 0.95 and PSS10: RMSEA = 0.065; CFI = 0.97; TLI = 0.96). In Ribeirão Preto, the 14-item model showed acceptable fit (RMSEA = 0.085; CFI = 0.93; TLI = 0.92), while the reduced 10-item model showed adequate fit (RMSEA = 0.078; CFI = 0.96; TLI = 0.95).

Table 2 lists the factor loadings, measurement errors, factor correlations, Cronbach's alpha, CR, AVE and the square root of the AVE from the two-factor PSS14 and PSS10 models for São Luís and Ribeirão Preto. Regarding PSS14, the lowest factor loadings for the São Luís sample were for items 8 – “Have you felt that you could not cope with all the things that you had to do?” (0.36), 12 – “Have you been thinking about things that you have to accomplish?” (0.09), both belonging to the negative factor, and 13 – “Have you felt you were able to control the way you spend your time?” (0.43). All other factor loadings were > 0.50. The highest measurement errors were also for items 8 (0.87), 12 (0.99) and 13 (0.81). The correlation between the negative and positive factors was negative and strong (-0.74). The correlation between psychological violence and the positive factor was negative and weak (-0.47), whereas the correlation between psychological violence and the positive factor was positive and moderate (0.59). Cronbach's alpha for each dimension was > 0.70. CR values were also above 0.70. AVE for each factor was < 0.50. PSS10 (version that does not contain items 4, 5, 12 and 13) had low factor loading only for the item 8 of the negative dimension (0.36). Factor correlations and Cronbach's alpha for the PSS10 were slightly different compared to PSS14. AVE, although a little higher compared to PSS14, remained below 0.50. For the positive and negative dimensions of the PSS14 and PSS10 models the AVE square roots were lower than their factor correlations.

For the Ribeirão Preto sample, PSS14 showed low factor loadings for items 8 (0.33) and 12 (0.09), and the residual variances of these items were high (0.89 and 0.99). Factor correlations, Cronbach's alpha and composite reliability were satisfactory. However, AVE was < 0.50. PSS10 had a low factor loading for item 8, and the correlation between the positive and negative factors was -0.78. Cronbach's alpha was > 0.70 for the two factors. AVE was 0.50 for the positive factor but still below 0.50 for the negative dimension. However, AVE was higher for the abbreviated PSS10 compared to the full PSS14 scale. The AVE square roots for the positive and negative dimensions of the PSS14 and PSS10 models were all lower than their factor correlations (Table 2).

Because low factor loadings were found for the same items in the São Luís and Ribeirão Preto models, we estimated the Loevinger's H scalability index to detect the viability of excluding these items with low factor loadings. No major violations of the monotonicity assumption occurred. As shown in Table 3, PSS14 evaluation in São Luís revealed that items 8 (HS = 0.24) and 12 (HS = 0.12)

Table 2

Factor loadings, measurement errors, factor correlations, Cronbach's alpha, composite reliability (CR), average variance extracted (AVE) and square root of the average variance extracted ($\sqrt{\text{AVE}}$) from the *Perceived Stress Scale* – PSS14 and PSS10 versions. São Luís, Maranhão State, and Ribeirão Preto, São Paulo State, Brazil, 2010/2011.

Items	São Luís						Ribeirão Preto					
	PSS14		δ	PSS10		δ	PSS14		δ	PSS10		δ
	Negative	Positive		Negative	Positive		Negative	Positive		Negative	Positive	
Negative items												
1	0.61		0.62	0.61		0.62	0.66		0.56	0.66		0.56
2	0.64		0.59	0.63		0.60	0.73		0.47	0.70		0.50
3	0.69		0.52	0.70		0.51	0.72		0.48	0.72		0.49
8	0.36		0.87	0.36		0.86	0.33		0.89	0.34		0.84
11	0.70		0.52	0.70		0.51	0.68		0.53	0.69		0.52
12	0.09		0.99				0.09		0.99			
14	0.72		0.47	0.72		0.49	0.66		0.56	0.65		0.57
Positive items												
4		0.62	0.61					0.64	0.59			
5		0.69	0.52					0.70	0.50			
6		0.67	0.55		0.62	0.62		0.73	0.47		0.66	0.56
7		0.64	0.59		0.64	0.59		0.68	0.53		0.68	0.54
9		0.67	0.55		0.69	0.52		0.69	0.52		0.70	0.50
10		0.71	0.49		0.72	0.48		0.75	0.44		0.77	0.40
13		0.43	0.81					0.48	0.77			
Factor correlations												
Positive with the negative factor	-0.74			-0.76			-0.74			-0.78		
Psychological violence with the positive factor	-0.47			-0.48			-0.46			-0.48		
Psychological violence with the negative factor	0.59			0.59			0.59			0.59		
α	0.72	0.79		0.75	0.72		0.72	0.81		0.75	0.75	
CR	0.76	0.83		0.79	0.76		0.77	0.85		0.80	0.79	
AVE	0.34	0.41		0.40	0.45		0.36	0.45		0.41	0.50	
$\sqrt{\text{AVE}}$	0.58	0.64		0.63	0.67		0.60	0.67		0.64	0.71	

δ : measurement errors.

of the negative dimension had the lowest scalability. For the positive dimension, the value for item 13 (HS = 0.26) was below the recommended cutoff point. For the PSS10 HS was < 0.30 only for item 8 of the negative factor. The HS values were higher for each dimension of PSS10 when compared to those observed for PSS14. In Ribeirão Preto, HS values were < 0.30 for items 8 and 12 of the PSS14, and for item 8 of the PSS10. Similarly to the São Luís sample, higher HS values were detected for each item and dimension of the PSS10 compared to the PSS14 version.

The residual correlation matrix showed few values slightly > 0.10. Among the modification indices suggested, none was deemed theoretically justifiable (suggestions for modification included cross loadings between the negative and positive dimensions). We considered that it is not possible for an item to measure at the same time a negative (suffering due to stress) and a positive (coping satisfactorily with stress) dimension.

Table 3

Loevinger's H for each item and for each dimension of the *Perceived Stress Scale* – PSS14 and PSS10 versions. São Luís, Maranhão State, and Ribeirão Preto, São Paulo State, Brazil, 2010/2011.

In the last month, how often have you (been/felt)	São Luís				Ribeirão Preto			
	PSS14		PSS10		PSS14		PSS10	
	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive
Negative factor								
1 – Upset by something happening unexpectedly?	0.34		0.37		0.34		0.37	
2 – Unable to control the important things in your life?	0.35		0.39		0.37		0.41	
3 – Nervous and stressed?	0.38		0.42		0.38		0.42	
8 – Could not cope with all the things that you had to do?	0.24		0.26		0.22		0.24	
11 – Angered because of things that were outside your control?	0.37		0.41		0.38		0.41	
12 – Thinking about things that you have to accomplish?	0.12				0.14			
14 – Difficulties were piling up so high that you could not overcome them?	0.38		0.42		0.36		0.40	
Positive factor								
4 – Dealt successfully with day-to-day problems and annoyances?		0.38				0.40		
5 – Effectively coping with important changes that were occurring in your life?		0.40				0.45		
6 – Confident about your ability to handle your personal problems?		0.41		0.41		0.45		0.45
7 – Things were going your way?		0.39		0.42		0.42		0.46
9 – Dealt successfully with irritating life hassles?		0.37		0.39		0.41		0.45
10 – You were on top of things?		0.42		0.45		0.45		0.51
13 – Able to control the way you spend your time?		0.28				0.33		
Loevinger's H	0.31	0.38	0.38	0.42	0.31	0.41	0.37	0.47

Discussion

CFA identified a two-factor structure for the PSS 10- and 14-item versions as showing the best fit for both São Luís and Ribeirão Preto samples. The two-factor models presented satisfactory fit indices, good reliability, and convergent and discriminant validity, in addition to acceptable scalability. AVE and scalability were slightly higher for PSS10 than for PSS14. The results were consistent for both São Luís and Ribeirão Preto samples.

In São Luís, both two-factor PSS10 and PSS14 models presented good fit indices (RMSEA close to 0.06; CFI and TLI > 0.95) ^{34,35,36}. In Ribeirão Preto, the 10-item model showed adequate fit and the 14-item version showed acceptable fit (RMSEA close to 0.08). Although this index calculates the error of approximation to the population ⁴⁵, values in the 0.08-0.10 range are admissible when other fit indices are considered good, which was the case in Ribeirão Preto, where the two-factor PSS14 model had CFI and TLI > 0.95. Furthermore, it has been suggested that the model should be rejected only when RMSEA > 0.10 ^{35,36}. Thus, in both samples two-factor PSS10 and PSS14 models with positive and negative dimensions showed good or at least acceptable fit.

PSS scale has the advantage of assessing an individual's behavior before a stressful situation. Our results indicate that PSS is a construct with two dimensions, one composed of positive items, measuring "perceived coping", and the other with negative items to measure "perceived distress". If the person reacts negatively when the body does not adapt to the new situation, and responds in a

misplaced way to the stimulus that caused this situation, this is called “perceived distress,” which is measured by the negative dimension of the scale. If the person responds positively, demonstrating resilience and coping capacity, this strategy is considered “perceived coping”, and is measured by the PSS⁴⁶ positive dimension. In this study, we called “perceived coping” a positive factor, and “perceived distress” a negative factor.

Good internal consistency indicates that the items consistently measure the same latent variable and are thus reliable³⁸. The two-factor PSS10 and PSS14 models had good internal consistency ($\alpha \geq 0.70$ and $CR \geq 0.70$). In addition, internal consistency for the negative dimension was higher for PSS10 compared to the PSS14, whereas for the positive dimension it was lower for PSS10 than for PSS14. Internal consistency is a function involving the number of items and the value of factor loadings⁴⁷. Item 12 of the negative dimension, and items 4, 5 and 13 of the positive dimension were excluded by Cohen & Williamsom²⁵ when they proposed the reduced PSS10 version because these items had low factor loadings. In our samples, items 12 and 13 also showed low factor loadings, but items 4 and 5 did not. The higher internal consistency for the PSS10 negative dimension compared to the PSS14 may be explained by the exclusion of item 12, which had a very low factor loading (0.09) in our samples. Conversely, lower internal consistency for the PSS10 positive dimension compared to the PSS14 in our samples may be explained by item reduction (from seven to four items) and also by the exclusion of two items (4 and 5) that did not have low factor loadings in our data. However, internal consistency for the reduced PSS10 remained good (≥ 0.70) in both samples, even after the exclusion of the two items (4 and 5), with factor loadings in the range 0.60-0.70 in the PSS14 model.

Construct validity indicates the extent to which a set of measured variables reflects the theoretical latent construct that those indicators are supposed to measure. According to Raykov & Marcoulides³⁹, no single test can determine construct validity and its assessment should be largely based on theory and not on single tests and fixed cutoff points. High reliability is a precondition for validity, but a construct may be reliable but lack validity³⁹. Therefore, estimating correlations between different constructs is essential to address construct validity⁴⁸. Strong correlations between measurements of closely related constructs provide evidence of convergent validity, whereas weak correlations between measurements of unrelated constructs are indicative of discriminant validity³⁹. Some authors consider AVE an indicator of convergent validity^{38,49}, but this view has been questioned by some authors – Valentini & Damásio⁴⁷ argue that AVE does not measure convergent validity but is rather a reliability indicator, closely resembling CR.

Construct validity was investigated based on factor loadings, AVE and the correlations between closely related constructs. Ideally all factor loadings should be greater than 0.70 or at least ≥ 0.50 ³⁸ and statistically significant. All but three factor loadings were ≥ 0.50 (items 8 and 12 for the negative dimension and item 13 for the positive dimension). AVE was < 0.50 for all PSS10 and PSS14 dimensions in both samples, and values were much lower for the negative dimension. This may imply problems in convergent validity, because most of PSS variance was explained by measurement error rather than by the latent trait it is supposed to measure³⁸. However, the cutoff point suggested by some authors to indicate adequate convergent validity may be too stringent. According to Valentini & Damásio⁴⁷, AVE tends to reject indiscriminately models with low factor loadings, especially when the variances of the factor loadings are low (in our samples standardized variances of the factor loadings were very low, in the 0.017-0.025 range). An AVE of 0.50 corresponds to factor loadings around 0.70. Thus, it is not consistent to recommend factor loadings of at least 0.50 or higher and an AVE ≥ 0.50 , since factor loadings of 0.50 corresponds to an AVE around 0.25. Low AVE values, which explain less than 50% of total variance, are commonly found in validation studies of other psychometric measures⁵⁰. In another PSS validation study in Brazil AVE was 0.34²⁸.

The correlations between the PSS10 and PSS14 positive and the negative subscales varied from -0.78 to -0.74 and were extremely significant. However, they were not too high (< 0.85), which suggest that, although the factors were part of a higher order latent variable, they reflected some characteristics that are unique to each factor. According to this criterion, discriminant validity exists between the subscales. On the other hand, the AVE square roots for the PSS14 and PSS10 positive and negative dimensions (0.58 to 0.71) were all lower than their factor correlations (0.72 to 0.81), suggesting problems in discriminant validity. Nevertheless, this criterion has been considered to be too stringent to provide evidence for separateness of sub-scales within a multidimensional scale⁵¹.

The correlations between the two dimensions of stress (perceived coping and perceived distress) with psychological violence were within the expected range (0.46-0.59), suggesting adequate convergent construct validity. The correlations indicate that those constructs measure different things but are moderately or weakly associated because psychological violence is expected to increase perceived stress⁵². Interestingly, the correlations were higher between psychological violence and perceived distress than between psychological violence and perceived coping.

Because some items had low factor loadings (< 0.50) and high measurement errors (items 8 and 12 for the negative dimension and item 13 for the positive dimension), we investigated the PSS14 and PSS10 scalability to determine if some items had weak scalability and could thus be dropped from the scale. Overall scalability was considered acceptable, but it was weak for the negative dimension and medium for the positive dimension. Items 8, 12 and 13 had low scalability and could thus be excluded from the scale. Item 12 was the most problematic because 99% of its residual variance was not explained by perceived distress. Perhaps “thinking about things that you have to accomplish” may not be translating into an experience of life that measures the perception of stress²⁷. Alternatively the high measurement error may be due to problems of understanding the content and meaning of the item, or even due to different interpretations about what the researcher intended to convey.

Items 12 and 13 were already removed from PSS14 in the reduced PSS10 scale. Thus, due to this fact, the reduced PSS10 version showed higher scalability values than the complete PSS14 version. However, item 8 “Could you not cope with all the things that you had to do?,” which composes the PSS10, also had low scalability in both of our samples. This has not been found in any of the comparative studies that investigated the PSS psychometric characteristics, including those whose sample consisted only of women^{10,13,28}, all displaying factor loadings ≥ 0.45 for this item. It is possible that this item has a different meaning during gestation. Perhaps thinking about the things pregnant women had to do might reflect everyday chores and responsibilities that may not always be related to stress. It has been found that women had higher perceived stress than men^{14,19,22,23,24,27}. Furthermore, Chaaya et al.¹⁰ have shown that pregnant women tend to have more perceived stress than non-pregnant women. Thus, it is possible that pregnant women may perceive stress differently than non-pregnant women, and these different perceptions of stress may influence the scale structure. Studies comparing the factorial structure of the scale between pregnant and non-pregnant women are important to address this question. The possible exclusion of item 8 should be tested in other studies.

It should be pointed out that the use of a convenience sampling limited the external generalization of our findings. A greater proportion of pregnant women who participated in our study belonged to middle economic class C both in São Luís (67%) and Ribeirão Preto (60%). Thus low and high economic classes were underrepresented in our samples.

A strong point of the study is the large sample size obtained in the two cities. An additional strong point is that it is the first Brazilian study to assess the PSS dimensional structure in pregnant women. Some studies have used PSS to measure stress in pregnant women^{5,10}, but only one of them has tested the validity of the instrument¹⁰ in a very small sample of 113 pregnant women.

In addition, our investigation used CFA with a more appropriate estimation method for categorical data, the WLSMV estimator, whereas the remaining studies that used CFA employed the ML estimator, more appropriate for continuous indicators^{18,19,26}. The use of CFA has the advantage of estimating convergent and divergent validity adjusting for measurement error³⁵. Although ML is extensively used in CFA investigations, some problems arise because the method treats categorical variables as continuous ones: (1) the estimated correlations between variables are attenuated, especially when the variables have less than five categories and a high degree of asymmetry; (2) the method produces errors in the estimation of the variance, and also produces incorrect estimation of some parameters⁴³.

This study contributed to the PSS validation because we used a more adequate method to assess validity (CFA) – criteria to assess model fit were stringent and scalability evaluation was performed. Furthermore, PSS was evaluated in two cities in a middle-income country with contrasting socioeconomic indicators.

Our results suggest that the two-factor models showed better fit and internal consistency than one-factor models. Considering all criteria evaluated (fit indices, reliability, convergent validity, discriminant validity, and scalability), the reduced two-factor PSS10 version seems to be a more

appropriate scale for measuring perceived stress among pregnant women. Furthermore, the results were consistent in the two samples. However, these results should be corroborated by more studies evaluating the PSS dimensional structure among pregnant women from more diverse socioeconomic backgrounds, since our samples were mainly composed of middle-class women. It is possible that cultural differences and differences in understanding some items may have affected the scale validity.

Contributors

A. V. C. P. Yokokura reviewed the literature, performed the statistical analysis and interpreted the data, wrote the manuscript, and participated in the approval of the final version to be published. A. A. M. Silva and M. A. Barbieri conceived the study, collaborated with data analysis and interpretation, contributed to the writing of the text, and participated in the approval of the final version to be published. J. K. B. Fernandes reviewed the literature, collaborated with data analysis and interpretation, contributed to the writing of the text and to the approval of the final version to be published. C. M. Del-Ben and F. P. Figueiredo collaborated with data analysis and interpretation, H. Bettioli conceived the study, contributed to the writing of the paper, collaborated with interpretation of the data, and participated in the approval of the final version to be published.

Acknowledgments

The authors would like to thank the pregnant women who participated in the research, to the Maranhão State Research Foundation (FAPEMA) and the Brazilian National Research Council (CNPq).

References

- Ribeiro J, Honrado A, Leal I. Contribuição para o estudo da adaptação portuguesa das escalas de ansiedade, depressão e stress (EADS) de 21 itens de Lovibond e Lovibond. *Psicol Saúde Doenças* 2004; 5:229-39.
- Huizink AC, Robles de Medina PG, Mulder EJH, Visser GHA, Buitelaar JK. Psychological measures of prenatal stress as predictors of infant temperament. *J Am Acad Child Adolesc Psychiatry* 2002; 41:1078-85.
- Bjelica A. Pregnancy as a stressful life event and strategies for coping with stress in women with pregnancy induced hypertension. *Med Pregl* 2004; 57:367-8.
- Segato L, Andrade A, Vasconcellos DIC, Matias TS, Rolim MKSB. Ocorrência e controle do estresse em gestantes sedentárias e fisicamente ativas. *Rev Educ Fís* 2009; 20:121-9.
- Araújo MA, Albertini R, Guimarães FP. Incidência de sintomas de estresse em primíparas: vivências e relatos de mulheres. *Polêmica* 2010; 9:64-73.
- Rodrigues OMPR, Schiavo RA. Stress na gestação e no puerpério: uma correlação com a depressão pós-parto. *Rev Bras Ginecol Obstet* 2011; 33:252-7.
- Woods SM, Melville JL, Guo Y, Fan MY, Gavin A. Psychosocial stress during pregnancy. *Am J Obstet Gynecol* 2010; 202:61.e1-61.e7.
- Mulder EJ, Robles de Medina PG, Huizinc AC, Van Den Bergh BR, Buitelaar JK, Visser GH. Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Hum Dev* 2002; 70:3-14.
- Cookson H, Granell R, Joinson C, Ben-Shlomo Y, Henderson AJ. Mother's anxiety during pregnancy associated with asthma in the children. *J Allergy Clin Immunol* 2009; 123:847-53.
- Chaaya M, Osman H, Naassan G, Mahfoud Z. Validation of the Arabic version of the Cohen Perceived Stress Scale (PSS-10) among pregnant and postpartum women. *BMC Psychiatry* 2010; 10:111.
- Cohen S, Karmack T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983; 24:385-96.
- Machado WL, Damásio BF, Borsa FC, Silva JP. Dimensionalidade da Escala de Estresse Percebido (Perceived Stress Scale, PSS10) em uma amostra de professores. *Psicol Reflex Crit* 2014; 27:38-43.
- Campo-Arias A, Bustos-Leiton GJ, Romero-Chaparro A. Internal consistency and dimensionality of the perceived stress scale (PSS-10 and PSS-14) in a sample of female university students in Bogotá, Colombia. *Aquichan* 2009; 9:271-80.
- Leung DYP, Lam TH, Chan SSC. Three versions of Perceived Stress Scale: validation in a sample of Chinese cardiac patients who smoke. *BMC Public Health* 2010; 10:513.
- Roberti JW, Harrington LN, Storch EA. Further psychometric support for the 10-item version of the Perceived Stress Scale. *Journal of College Counseling* 2010; 9:135-47.
- Wongpakaran N, Wongpakaran T. The Thai version of the PSS-10: an investigation of its psychometric properties. *Biopsychosoc Med* 2010; 12:6.
- Trigo M, Canudo N, Branco F, Silva D. Estudo das propriedades psicométricas da Perceived Stress Scale (PSS) na população portuguesa. *Psychologica* 2010; 53:353-78.
- Ramírez MTG, Hernández RL. Factor structure of the Perceived Stress Scale (PSS) in a sample from Mexico. *Span J Psychol* 2007; 10:199-206.
- Andreou E, Alexopoulos EC, Lionis C, Varvogli L, Gnardellis C, Chrousos GP, et al. Perceived Stress Scale: reliability and validity study in Greece. *Int J Environ Res Public Health* 2011; 8:3287-98.
- Mimura C, Griffiths P. A Japanese version of the perceived stress scale: translation and preliminary test. *Int J Nurs Stud* 2004; 41:379-85.
- Wang Z, Chen J, Boyd JE, Zhang H, Jia X, Qiu J, et al. Psychometric properties of the Chinese version of the perceived stress scale in police-women. *PLoS One* 2011; 6:e28610.
- Remor E. Psychometric properties of a European Spain version of the Perceived Stress Scale (PSS). *Span J Psychol* 2006; 9:86-93.
- Lesage F, Berjot S, Deschamps F. Psychometric properties of the French versions of the perceived stress scale. *Int J Occup Med Environ Health* 2012; 25:178-84.
- Luft CDB, Sanches SO, Mazo GZ, Andrade A. Versão brasileira da Escala de Estresse Percebido: tradução e validação para idosos. *Rev Saúde Pública* 2007; 41:606-15.
- Cohen S, Williamsom GM. Perceived stress in a probability sample of United States. In: Spacapan S, Oskamp S, editors. *The social psychology of health: claremont symposium on applied social psychology*. Newbury Park: Sage; 1988. p. 31-67.
- Reis RS, Hino AAF, Rodriguez-Añez CR. Perceived Stress Scale: reliability and validity study in Brazil. *J Health Psychol* 2010; 15: 107-14.
- Faro A. Análise fatorial confirmatória das três versões da Perceived Stress Scale (PSS): um estudo populacional. *Psicol Reflex Crit* 2015; 28:21-30.
- Dias JCR, Silva WR, Maroco J, Campos JA-DB. Escala de Estresse Percebido aplicada a estudantes universitárias: estudo de validação. *Psychology, Community & Health* 2015; 4:1-13.
- Lee EH. Review of the psychometric evidence of the Perceived Stress Scale. *Asian Nurs Res (Korean Soc Nurs Sci)* 2012; 6:121-7.

30. Ribeiro JP, Marques T. A avaliação do stresse: a propósito de um estudo de adaptação da escala de percepção de stresse. *Psicol Saúde Doenças* 2009; 10:237-48.
31. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation* 2005; 10:173-8.
32. Ullman JB. Structural equation modeling. In: Tabachnick BG, Fidell LS, editors. *Using multivariate statistics*. 5th Ed. Boston: Pearson Education; 2007. p. 653-771.
33. Silva AAM, Simões VMF, Barbieri MA, Cardoso VC, Alves CMC, Thomaz EBA, et al. A protocol to identify non-classical risk factors for preterm births: the Brazilian Ribeirão Preto and São Luís prenatal cohort (BRISA). *Reprod Health* 2014; 11:2-9.
34. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999; 6:1-55.
35. Brown TA. *Confirmatory factor analysis for applied research*. 2nd Ed. New York: The Guilford Press; 2015.
36. Browne MW, Cudeck R. Alternate ways of assessing model fit. In: Bollen KA, Long JS, editors. *Testing structural equation models*. Newbury Park: Sage; 1993. p. 136-62.
37. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990; 107:238-46.
38. Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate data analysis*. 7th Ed. Harlow: Pearson; 2009.
39. Raykov T, Marcoulides GA. *Introduction to psychometric theory*. New York: Routledge; 2011.
40. Ribeiro MR, Alves MT, Batista RF, Ribeiro CC, Schraiber LB, Barbieri MA, et al. Confirmatory factor analysis of the WHO violence against women instrument in pregnant women: results from the BRISA Prenatal Cohort. *PLoS One* 2014; 9:e115382.
41. Ribeiro MR, Silva AA, Alves MT, Batista RF, Ribeiro CC, Schraiber LB, et al. Effects of socioeconomic status and social support on violence against pregnant women: a structural equation modeling analysis. *PLoS One* 2017; 12:e0170469.
42. Hardouin JB, Antignac AB, Sébille V. Nonparametric item response theory using Stata. *Stata J* 2011; 11:30-51.
43. Léon DAD. *Análise fatorial confirmatória através dos softwares R e Mplus* [Monografia de Graduação]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2011.
44. Barrett P. Structural equation modeling: adjudging model fit. *Pers Individ Dif* 2007; 42:815-24.
45. Buiatti CL. Validação da escala de PSM-Public Service Motivation por meio de modelagem de equações estruturais [Dissertação de Mestrado]. Ribeirão Preto: Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo; 2007.
46. Hewitt PL, Flett GL, Mosher SW. The perceived stress scale: factor structure and relation to depression symptoms in a psychiatric sample. *J Psychopathol Behav Assess* 1992; 14:247-57.
47. Valentini F, Damásio BF. Variância média extraída e confiabilidade composta: indicadores de precisão. *Psicol Teor Pesqui* 2016; 32:1-7.
48. American Educational Research Association; American Psychological Association; National Council on Measurement in Education. *Standards for educational and psychological testing*. Washington DC: American Psychological Association; 2014.
49. Fornell C, Larcker DF. Evaluating structural equations models with unobservable variables and measurement error. *J Mark* 1981; 18:39-50.
50. Tinsley HEA, Tinsley DJ. Uses of factor analysis in counseling psychology research. *J Couns Psychol* 1987; 34:414-24.
51. Carter SR. Using confirmatory factor analysis to manage discriminant validity issues in social pharmacy research. *Int J Clin Pharm* 2016; 38:731-7.
52. Pico-Alfonso MA. Psychological intimate partner violence: the major predictor of posttraumatic stress disorder in abused women. *Neurosci Biobehav Rev* 2005; 29:181-93.

Resumo

O objetivo deste estudo foi avaliar a estrutura dimensional, confiabilidade, validade convergente, validade discriminante e escalabilidade da Escala de Estresse Percebido (EEP). A amostra foi constituída por 1.447 gestantes de São Luís (Maranhão) e 1.400 de Ribeirão Preto (São Paulo), Brasil. Foram avaliadas as versões com 14 e 10 itens por meio da análise fatorial confirmatória, utilizando-se o estimador dos mínimos quadrados ponderados ajustados pela média e variância (WLSMV). Em ambas as cidades, os modelos bifatoriais (fatores positivo, medindo resiliência a situações estressoras, e negativo, medindo situações de estresse) apresentaram melhor ajuste do que os unifatoriais. Os modelos bifatoriais da escala completa (EEP14) e da reduzida (EEP10) apresentaram boa consistência interna (alfa de Cronbach ≥ 0.70). Todas as cargas fatoriais foram ≥ 0.50 , exceto as dos itens 8 e 12 da dimensão negativa e a do item 13 da dimensão positiva. As correlações entre ambas as dimensões do estresse com a violência psicológica foram na magnitude esperada (0,46-0,59), fornecendo evidência de validade de construto convergente adequada. As correlações entre as dimensões positiva e negativa das escalas foram em torno de 0,74-0,78, menores do que $< 0,85$, o que sugere adequada validade discriminante. A variância média extraída e a escalabilidade foram ligeiramente maiores para a EEP10 do que a para a EEP14. Os resultados foram consistentes em ambas as cidades. Em conclusão, a solução unifatorial não é recomendada para avaliar estresse em mulheres grávidas. A escala bifatorial reduzida com 10 itens parece ser mais apropriada para medir estresse percebido em gestantes.

Gestantes; Estresse Psicológico; Análise Fatorial

Resumen

El objetivo de este estudio fue evaluar la estructura dimensional, confiabilidad, validez convergente, validez discriminante y escalabilidad de la Escala de Estrés Percibido (EEP). La muestra estaba constituida por 1.447 gestantes de Sao Luis (Maranhão) y 1.400 de Ribeirão Preto (São Paulo), Brasil. Se evaluaron las versiones con 14 y 10 ítems, mediante un análisis factorial confirmatorio, utilizándose el estimador de los mínimos cuadrados ponderados ajustados por la media y variancia (WLSMV). En ambas ciudades, los modelos bifactoriales (factores positivo, midiendo resiliencia a situaciones estresantes, y negativo, midiendo situaciones de estrés) presentaron un mejor ajuste que el de los unifactoriales. Los modelos bifactoriales de la escala completa (EEP14) y de la reducida (EEP10) presentaron una buena consistencia interna (alfa de Cronbach ≥ 0.70). Todas las cargas factoriales fueron ≥ 0.50 , excepto las de los ítems 8 y 12 de la dimensión negativa y la del ítem 13 de la dimensión positiva. Las correlaciones entre ambas dimensiones del estrés con la violencia psicológica fueron en la magnitud esperada (0,46-0,59), proporcionando evidencia de validez del constructo convergente adecuada. Las correlaciones entre las dimensiones positiva y negativa de las escalas fueron en torno de 0,74-0,78, menores que $< 0,85$, lo que sugiere una adecuada validez discriminante. La variancia media extraída y la escalabilidad fueron ligeramente mayores para la EEP10 que para la EEP14. Los resultados fueron consistentes en ambas ciudades. En conclusión, la solución unifactorial no está recomendada para evaluar estrés en mujeres embarazadas. La escala bifatorial reducida con 10 ítems parece ser más apropiada para medir el estrés percibido en gestantes.

Mujeres Embarazadas; Estrés Psicológico; Análisis Factorial

Submitted on 06/Nov/2015

Final version resubmitted on 13/Mar/2017

Approved on 22/Mar/2017