Psychometric properties and factor structure of the Swedish version of the Perceived Stress Scale.

Eklund, Mona; Bäckström, Martin; Tuvesson, Hanna

Published in:
Nordic Journal of Psychiatry

DOI:
10.3109/08039488.2013.877072

2014

Running head: Swedish Perceived Stress Scale

Psychometric properties and factor structure of the Swedish version of the Perceived Stress Scale

Mona Eklund, Ph.D., Professor [1)]

Martin Bäckström, Ph.D., Professor [2)]

Hanna Tuvesson, Ph.D., Assistant Professor [3)]

[1)] Department of Health Sciences, Lund University, Lund, Sweden

[2)] Department of Psychology, Lund University, Lund, Sweden

[3)] School of Health Science, Blekinge Institute of Technology, Karlskrona, Sweden

Address for correspondence:

Mona Eklund, Department of Health Sciences, Occupational Therapy and Occupational Science,

Lund University, PO Box 157, SE-221 00 Lund, Sweden.

Tel +4646 2221957

Fax +4646 2221959

E-mail: mona.eklund@med.lu.se

Abstract

Background: The Perceived Stress Scale (PSS) measures general stress and the Swedish version, although used in several studies, has not been extensively evaluated for psychometric properties.

Aims: This study aimed to investigate psychometric properties and the factor solution of the Swedish 14-item version when used with two samples, namely a mixed Internet sample of women and men (N=171) and another of women with stress-related disorders (N=84). Classical test theory, including confirmatory factor analysis, was employed.

Results: The *factor structure* supported a two-factor model for the PSS and confirmed other language versions of the PSS, although one items showed a low item-total correlation. The PSS showed to be feasible with the investigated samples and the results indicated no ceiling or floor effects and good *internal consistency* of the PSS. Several aspects of construct validity were shown. An association of -.66 between the PSS and a measure of coping indicated good *concurrent validity*. *Criterion validity* was demonstrated through a statistically significant difference ($p<.001$) between the women with stress-related disorders and the Internet sample. *Predictive validity* of the PSS could be demonstrated in a short-term perspective. Based on the sample with stress-related disorders, *sensitivity to change* was shown through a statistically significant stress reduction ($p<.001$) from entering work rehabilitation to discharge.

Conclusions: The Swedish version of the PSS showed satisfactory psychometric properties and may be recommended for use with people with and without known stress-related disorders.

Keywords: Coping, homogeneity, reliability, validity.

Introduction

Stress-related health conditions are an urgent issue in most western countries (1-4). Stress is thus an important risk factor to assess in actions to prevent stress-related disorders and as an outcome measure in rehabilitation projects. Therefore, reliable and valid measures of perceived stress are needed.

The Perceived Stress Scale (PSS) has been developed to measure general stress (5). It is a self-report scale with both negative items (expressing stress) and positive items (expressing coping capacity). The intention was to capture the conceptualization of stress as outlined by Lazarus and Folkman (6): that stress is an appraisal of something threatening and that people cope with stress more or less effectively. The original 14-item version of the PSS showed satisfactory internal consistency and adequate concurrent and predictive validity (5), but a 10-item version was also launched (7). That study also concluded that two factors lay behind the scale, one formed by the negative items and one by the positive. The PSS has been translated into several languages. A Swedish version of the 14-item PSS was found to have adequate internal consistency and concurrent validity (8). More recently, the PSS has been translated into, for example, Chinese (9), Spanish (10) and Japanese (11). Adequate internal consistency and support for a two-factor solution based on the negative and positive items, respectively, has been demonstrated (9, 10). Leung and colleagues also found the 10-item version to have better psychometric properties than the 14-item version (9).

The 14-item version, but not the 10-item version, has been extensively used in Sweden (12-14). No methodological study seems to have been undertaken since the mid-1990s, however, and criterion validity, sensitivity to change and the factor structure do not seem to have been investigated at all. Criterion validity is a vitally important property if the scale is to be used to identify risk groups and sensitivity to change is a necessary

characteristic if it is to be used as an outcome measure. Against this background, the aim of this study was to investigate if previous findings of adequate internal consistency and concurrent validity of the Swedish 14-item PSS could be replicated, and to examine its factor solution, criterion validity, predictive validity and sensitivity to change when used with people with and without stress-related ill-health.

Methods

The study included two samples. One was an Internet sample consisting of participants who during 2012 spontaneously visited a website with different psychological tests and inventories. The other was a sample of women with stress-related ill-health who, in 2008 - 2010, participated in a work rehabilitation programme. The ethical vetting board at Lund University approved that study (Nos. 922/2004, 149/2008) and the principles of informed consent and voluntary participation were followed. No ethical approval was needed for the Internet sample since they visited the website and filled out the form on their own initiative.

Participants

The Internet sample consisted of 171 subjects: 108 women aged 31.5 years (SD = 11.6) and 63 men aged 33.6 years (SD=11.0). They had no known illness. Two subjects were found to be multivariate outliers; therefore the final N in this sample was 169.

The sample of women with stress-related disorders comprised women participating in a quasi-experimental evaluation study of a work rehabilitation programme (13), re-analysed for the present study. The women were on sick leave for their stress-related disorder, codes F34 and F43 according to the ICD-10 (15). A physician, most often specialized in primary health care, had made the diagnosis. The sample was selected from the registers of the Social Insurance Offices in two counties in southern Sweden and the data were collected during 2008 – 2009. Eighty-four women participated in either of two rehabilitation alternatives, and

for the purpose of the present study they were treated as one sample. All of them agreed to take part in the study. They were on average 45 years old (SD=10) and had been on sick leave for on average 12 months (SD=16). A majority had white-collar jobs. Data was collected at four measurement points – admission, discharge and follow-ups at 6 and 12 months after discharge. For the present study, the measurements at admission, discharge and the 12-month follow-up were used. Seven participants (8.3%) dropped out between admission and the discharge measurement.

Instruments

The Perceived Stress Scale

The PSS has 14 items and uses a five-point response scale from 0 (never) to 4 (very often). The items ask about how often a certain event has happened in the past month. Seven items have a negative character, and seven have a positive. The latter are reversed when calculating a total score (5). A higher score indicates more stress and the scores may range between zero and 56. The Swedish version (8) was used for the present study.

The Pearlin Mastery Scale

The Mastery scale, developed by Pearlin and colleagues (16, 17), was used to assess concurrent validity of the PSS. Mastery reflects a kind of coping capacity, specified as the view of one's ability to impact on circumstances that importantly influence one's life (16). The scale has seven items and a four-point response scale from 1 (strongly agree) to 4 (strongly disagree). Two items have a negative wording and are to be reversed. A higher score indicates more Mastery, thus more ability to influence and decide about important things in one's life, and scores may vary between 4 and 28. The Swedish version was used (18). A recent study indicated satisfying psychometric properties when analysed according to item-response theory, including good construct validity (19).

Data analyses

Classical test theory was employed. Samples and methods were used to test different psychometric properties of the PSS, as follows. When relevant, hypotheses were stated to guide the analyses.

Internal consistency: Cronbach's alpha was calculated for each sample separately, and the Corrected Item-Total Correlations (CITC) were checked. The limit for a satisfactory item-total correlation is >.30 (20).

Factor structure: The factor analysis was based on separate analyses in the two samples, starting out with the larger Internet sample. The inventory was originally introduced as a one-factor scale (5), but a two-factor model has later on been proposed, separating positively and negatively worded items (10). Models based on these two conceptualisations of the scale were tested and compared with Confirmatory Factor Analysis (CFA). In addition the fit of the best model was evaluated based on Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA). The first mentioned compares the estimated model with the independent model and the second investigates if there is covariation not accounted for in the models. It is highly likely that some of the variables will not fit perfectly to the models; to remedy this exploratory adjustment of the model can be tested if necessary. To estimate the models we used the MPlus program, and the models were tested with Maximum Likelihood estimator, but we also tested the CFA on the Internet sample with analysis or ordinal response variables (CATEGORICAL in Mplus). At last the factor structure was tested for invariance between the Internet sample and the stress sample.

Concurrent validity: Pearson's correlation test was used to calculate the association between the PSS and the Mastery scale in the sample with stress-related disorders. A strong

and negative correlation was hypothesized, since the measures partly measure an equivalent phenomenon (coping).

Criterion validity: The two-samples *t*-test was used to calculate differences between the sample with stress-related disorders and the Internet sample. It was hypothesized that the PSS would discriminate between the groups. Since stress has been shown to have different backgrounds in women and men (21), we also performed a comparison based on women only.

Predictive validity: Pearson correlations between PSS change and return to work were calculated in the group with stress-related disorders. It was hypothesized that there would be a moderate but statistically significant relationship between baseline stress scores and return to work rate after finished rehabilitation and at the 12-month follow-up. Return to work was estimated as the percentage of working hours compared to normal (before sick leave) working hours.

Sensitivity to change: Focusing on the women with stress-related disorders, it was hypothesized that their PSS scores would decrease from the point in time when they entered rehabilitation to completion of the programme, as analysed with a paired *t*-test.

Results

The distribution of responses to the PSS showed no empty response category on any of the items in any of the samples. The distribution of the total scale was somewhat left-skewed in the Internet sample (skewness = .40), while it was a bit right skewed in the female sample (skewness = -.53). No obvious ceiling or floor effect was seen on the scale in these samples.

Regarding *internal consistency*, the Cronbach's alpha was .84 in the sample with stress-related disorders and .90 in the Internet sample. The total–item correlation was quite low for item 12 in both samples, .25 in the Internet sample and .24 in the stress sample. All other CITCs were >.30.

The *factor analysis* was performed as follows: Item number 12 was excluded because of the low ICTC in the factor analyses, but not in any other analyses. Starting with the Internet sample, the one-factor model was first estimated; it had far from perfect fit, χ^2 (65) = 181.95. This model revealed that a large part of the covariation was accounted for (CFI = .85), but also that there were several significant relations not accounted for, RMSEA = .10. The two-factor model, with correlation between the factors, had considerably better fit, $\chi^2(64) = 118.53$ ($\Delta\chi^2$ (1) = 63.42, Δ p < .001), CFI = .95, and RMSEA = .071 (CI = .051 - .091). To conclude, the two-factor model, separating positive and negative items, was clearly better. The correlation between the two factors was .83, and all loadings were above .50[1] (see Table 1). Using both samples, we tested invariance between the Internet and the Stress sample, first by testing if the loadings were significantly different. After that we tested if the intercepts of the items were different (holding the mean of the factors equal), and the last model tested if the mean levels of the factors were different. It was found that a configural invariance baseline model had a χ^2 (128) = 227.61, CFI = .931, and RMSEA = .078 (C.I. .062 - .095). The weak invariance model, with all loadings set to equal, increased the χ^2 (139) = 237.20, CFI .932, RMSEA = .075 (C.I. .058 - .091) suggesting a non-significant decrease in fit ($\Delta\ \chi^2$ (11) = 9.59, Δp > .05) and the fit indices even indicated a somewhat better fit. The hypothesis of invariance thus seemed to hold, in other words the loadings on the two factors were not significantly different in the two groups. We also tested a strong invariance model, setting both the intercept and the factor loadings to be equal. This model revealed a somewhat worse fit, $\chi^2$150) = 255.30, CFI = .927, and RMSEA = .074 (C.I. .059 - .090). Generally some of the intercepts were higher for the stress sample (delta $\chi^2(11) = 18.10$, p > .05), but the difference did not reach the stipulated level of significance. This also suggests that if there were

[1] The same models were also estimated using WLSMV and setting all response variables as categorical. These estimations gave the same results; only the RMSEA was a bit higher. The method was not used for invariance testing because some of the response variables were not populated on all rating alternatives for all response variables.

difference in mean levels (intercepts) then these differences could be attributed to all of the items (or to a general factor). The last model also restrained the mean of the factors to be equal, χ^2 (152) = 274.73, CFI = .914, and RMSEA = .080 (C.I. .065 - .093), and this resulted in a significant decrease in fit (delta χ^2 (2) = 19.43, $p < .001$). In the strong model the means of the positive and negative factor were both higher ($p < .001$) for the stress sample (.408 and .320, respectively)[2]. To conclude, the two samples' factor structure was invariant but there were general differences in mean levels (as suggested above in relation to the total scale), e.g. the stress sample rated higher on all items.

Table 1 in here

Criterion validity was demonstrated through a statistically significant difference F(1, 251) = 12.517, $p < .001$, d = 0.46) between the women with stress-related disorders (mean = 29.59, SD = 7.37) and the Internet sample (mean = 24.8, SD = 11.14). The coping ability subscale and the stress subscale both discriminated between the samples, but none of them was significant after controlling for the other. Also when the comparison was based on women only, the women with stress-related disorders scored higher than the female Internet sample (mean = 26.83; SD =10.7), F(187) = -3.938, p = .049, d = 0.29). It was found that age was related to stress in both samples, somewhat stronger in the stress sample (r = -.309, p = .006) than the Internet sample (r = -.222, p = .004). Adding age as a covariate to the difference between the two samples strengthened the discrimination between groups, F(1,249) = 25.205, $p < .001$. Also when comparing only women, adding age as a covariate strengthened the discrimination, F(1,186) = 16.704, $p < .001$.

[2] Cohen and Williamson suggested a 10-item version (7)-that excluded 4 items (4, 5, 12, 13). Estimating a model with this version of the inventory resulted in acceptable fit, RMSEA = .08, CFI = .96. Items 4, 5 and 13 did not create problems for the models estimated based on data from the Swedish version.

With respect to *concurrent validity*, the correlation between the PSS and the Mastery scale was r = -.66, p < .001, in the stress sample. Using the two factors as subscales of the PSS, the correlation was r = -.69, p<.001, for the coping ability subscale and r = -.51, p<.001, for the stress subscale (p<.001).The partial correlation between the Mastery Scale and the coping ability scale was not significant when controlling for the stress scale. Controlling for age did not change these correlations.

Predictive validity was indicated by a statistically significant correlation between perceived stress at baseline and return to work rate among the women with stress-related disorders after rehabilitation completion (r = -.31; p = .004). No statistically significant relationship was found, however, at the 12-month follow-up (r = -.122; p > .05). The relationship was somewhat stronger for the stress subscale (r = -.32, p = .005) than the coping subscale (r = .23, p = .036), and the partial correlation between the stress subscale and follow-up work rate after controlling for the coping subscale was almost significant (r = -.211, p = .059). The stress subscale was even significantly related with return to work rate at the 12-month follow-up (r = -.239; p = .032). Controlling for age did not change these correlations.

Regarding *sensitivity to change*, as based on the sample with stress-related disorders, the change from entering work rehabilitation to discharge was statistically significant (p<.001, d=0.51). The perceived stress was reduced from 30 (SD = 8.6) to 25.5 (SD = 9).

Discussion

The PSS seems usable with the two types of samples used for this study, specifically women on sick leave for stress-related disorders and a mixed Internet sample of men and women without known illness. The response distribution did not show serious deviation from normality and all response alternatives were used, thus no obvious ceiling or floor effects were identified in these samples. This is important since it indicates that the scale can be used

with groups having higher and lower levels of stress symptoms than the here described samples.

By deleting item 12, which showed a low correlation with the total scale, we could confirm a two-factor model for the PSS, in the Internet sample as well as the sample of women with stress. This is in line with other language versions (5, 9, 10) and indicates that the Swedish version has retained the two factors of coping ability and stress. All other items were unproblematic, also items 4, 5 and 13 which have otherwise been deleted in the proposed 10-item version (7). Thus, unlike the study by Leung and colleagues (9) we did not find any indications that a 10-item version would be preferable. The 10-item version has otherwise been adopted in other countries and language versions as well (22-24) and nothing from the present study indicates the 10-item would not be feasible in Sweden.

The alpha values indicated homogeneity for both samples, also confirming findings from other language versions (5, 9, 10), as well as a previous study on the Swedish version (8).

The hypothesized relationship between the PSS and the Mastery scale was confirmed, ensuring concurrent validity for the PSS. Particularly the coping ability factor correlated strongly with Mastery. The results regarding other aspects of validity were also promising. The hypothesis concerning the ability of the PSS to discriminate between the sample with stress-related disorders and the Internet sample was confirmed. The means were approximately 30 and 25, and an effect size considered as medium-sized (25) was obtained, indicating a difference of clinical relevance and rendering support for criterion validity. Moreover, in the group of women on sick leave, there was an association between their stress level at admission and their return to work rate after completed rehabilitation, but not at the follow-up after 12 months. Hence, predictive validity could be established in a short-term but not in a long-term perspective. It seems that other forces than perceived stress lay behind the

women's return to work in the long run. The PSS was sensitive to change, as indicated by a highly significant difference between stress levels before and after work rehabilitation among the women with stress-related disorders. The mean scores were reduced from 30 to 25.5, which is considerable and corresponds to e medium-sized effect (25), and the PSS was thus able to detect a change that would be clinically important.

It is a common practice to include both positively and negatively worded items in rating scales measuring psychological constructs. This practice reduces an obvious response bias, but it also introduces a problem, e.g. that a scale supposed to measure one factor, when tested with CFA seems to measure two. With the PSS, there was a high correlation between the two factors, and their criterion validities did not indicate that the coping ability subscale (positive items) and the stress subscale (negative items) had different contents. Based on this it seems safe to collapse the two factors to one.

Based on this single study, we do not propose that item 12 should be deleted in the Swedish version of the PSS, and all validity and reliability analyses resulted in good values based on the entire 14-item version. Further psychometric testing is needed to see if the factor structure obtained in the present study can be confirmed, and if so, a shortened version might be recommendable. That would constitute a disadvantage when comparing findings based on the full version, but a solution to that would be to report mean scores instead of sum scores.

Methodological concerns

The test of factor structure invariance between the two samples was based on two rather small samples. It is safe to say that there were no major differences between the samples, but using larger samples some of the hypotheses of invariance might have been falsified. Another potential methodological concern is the fact that the Internet sample was self-chosen and their motives for visiting the website and characteristics other than age and sex are unknown.

However, since the included samples showed invariance regarding factor loadings, the question regarding the Internet sample is mainly whether the PSS mean value is a fair estimate of the population mean in Sweden. The present study gives no answer to this question.

Conclusion

The PSS showed to be feasible with the investigated samples. A two-factor structure was supported, suggesting that the Swedish version is equivalent to other language versions. Satisfactory homogeneity was also obtained. The established concurrent validity and criterion validity of the PSS, not previously shown for the Swedish version, together gave a picture of good construct validity. The scale could thus discriminate between people with and without troublesome stress, but it could not effectively predict return to work in a long-term perspective. In all, the Swedish version of the PSS showed satisfactory psychometric properties and may be recommended for use with people with and without known stress-related disorders.

Disclosure of interests

The authors have no conflicts of interest to declare.

References

1. European Foundation for the Improvement of Living and Working Conditions. Work-related disorders in Sweden. 2007. http://www.eurofound.europa.eu/ewco/surveys/SE0601SR01/SE0601SR01.pdf
2. Lander F, Friche C, Tornemand H, Andersen JH, Kirkeskov L. Can we enhance the ability to return to work among workers with stress-related disorders? BMC Public Health. 2009;9:372.
3. Uegaki K, Bakker I, de Bruijne M, van der Beek A, Terluin B, van Marwijk H, et al. Cost-effectiveness of a minimal intervention for stress-related sick leave in general practice: Results of an economic evaluation alongside a pragmatic randomised control trial. J Affect Disord. 2010;120:177-87.
4. Vaez M, Rylander G, Nygren A, Asberg M, Alexanderson K. Sickness absence and disability pension in a cohort of employees initially on long-term sick leave due to psychiatric disorders in Sweden. Soc Psychiatry Psychiatr Epidemiol. 2007;42:381-8.
5. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24:385-96.
6. Lazarus RS, Folkman S. Stress, appraisal, and coping. New York: Springer Pub. Co.; 1984.
7. Cohen S, Williamson G. Perceived stress in a probability sample in the U. S. In: Spacapan S, Oskamp S, editors. The social psychology of health. Newbury Park, CA: Sage; 1988. p. 31-67.

8. Eskin M, Parr D. Introducing a Swedish version of an instrument measuring mental stress. Reports from the Department of Psychology, No 813. Stockholm: Stockholm University; 1996.

9. Leung DY, Lam TH, Chan SS. Three versions of Perceived Stress Scale: Validation in a sample of Chinese cardiac patients who smoke. BMC Public Health. 2010;10:513.

10. Ramirez MT, Hernandez RL. Factor structure of the Perceived Stress Scale (PSS) in a sample from Mexico. Span J Psychol. 2007;10:199-206.

11. Mimura C, Griffiths P. A Japanese version of the perceived stress scale: Translation and preliminary test. Int J Nurs Stud. 2004;41:379-85.

12. Tuvesson H, Eklund M, Wann-Hansson C. Perceived stress among nursing staff in psychiatric inpatient care: The influence of perceptions of the ward atmosphere and the psychosocial work environment. Issues Ment Health Nurs. 2011;32:441-8.

13. Eklund M, Erlandsson L-K. Return to work outcomes of the Redesigning Daily Occupations (ReDO) program for women with stress-related disorders – A comparative study. Women Health. 2011;51:676-92.

14. Lesage FX, Berjot S. Validity of occupational stress assessment using a visual analogue scale. Occup Med (Lond). 2011;61:434-6.

15. WHO. The ICD-10 classification of mental and behavioural disorders. Geneva: World Health Organization; 1993.

16. Pearlin LI, Menaghan EG, Lieberman MA, Mullan JT. The stress process. J Health Soc Behav. 1981;22:337-56.

17. Pearlin LI, Schooler C. The structure of coping. J Health Soc Behav. 1978;19:2-21.

18. Brenner S-O, Arnetz B. Effekter av otrygghet i arbetet. Förlust av arbetet och arbetslöshet. Statens institut för psykosocial miljömedicin. Stressforskningsrapporter, No 170;1983.

19. Eklund M, Erlandsson LK, Hagell P. Psychometric properties of a Swedish version of the Pearlin Mastery Scale in people with mental illness and healthy people. Nord J Psychiatry. 2012;66:380-8.

20. Lindström M, Hariz G-M, Bernspång B. Dealing with real-life challenges: Outcome of a home-based occupational therapy intervention for people with severe psychiatric disability. OTJR : Occupation, Participation and Health. 2012;32:5-13

21. Asberg M, Nygren A, Leopardi R, Rylander G, Peterson U, Wilczek L, et al. Novel biochemical markers of psychosocial stress in women. PLoS One. 2009;4(1):e3590.

22. Barbosa-Leiker C, Kostick M, Lei M, McPherson S, Roper V, Hoekstra T, et al. Measurement Invariance of the Perceived Stress Scale and Latent Mean Differences across gender and time. Stress Health. Early online;2012 Oct 2.

23. 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.

24. Örücü MC, Demir A. Psychometric evaluation of perceived stress scale for Turkish university students. Stress Health. 2009;25:103-9.

25. Cohen J. A power primer. Psychol Bull. 1992;112:155-9.

Table 1. Loadings for the two-factor model with correlation between factors (ML estimation)

	Items	*Estimate*	*SE Estimate*	*Internet sample*	*Stress sample*
Factor 1	1	0.712	0.043	1.64	2.14
	2	0.571	0.056	2.02	2.33
	3	0.799	0.033	2.27	2.89
	8	0.708	0.043	1.99	2.47
	11	0.569	0.057	1.69	1.83
	14	0.857	0.027	1.40	1.84
	12			3.16	3.06
			Mean	2.02	2.37
			Sum	14.17	16.56
Factor 2	4	0.614	0.052	1.61	1.76
	5	0.570	0.057	1.76	2.21
	6	0.713	0.042	1.56	1.75
	7	0.749	0.038	1.48	1.94
	9	0.803	0.032	1.61	2.14
	10	0.868	0.025	1.38	1.79
	13	0.547	0.058	1.28	1.84
			Mean	1.53	1.92
			Sum	10.68	13.43
Total			Mean	1.78	2.11
			Sum	24.85	29.59