Evaluating the construct validity of Stress Overload Scale-Short using exploratory structural equation modeling

Wenlong Mu and Wenjie Duan

Abstract
This study used exploratory structural equation modeling and confirmatory factor analysis to examine the construct validity of the Chinese version of Stress Overload Scale-Short, which included personal vulnerability and event load. The participants included 629 community residents and 495 university students. The results indicated a better goodness-of-fit using exploratory structural equation modeling compared with confirmatory factor analysis. The Stress Overload Scale-Short performed well in distinguishing individuals with more negative emotion symptoms from the general population. A moderation analysis demonstrated that social support moderates the effect of personal vulnerability on negative emotion symptoms. These results facilitated the application of Stress Overload Scale-Short in the current population.

Keywords
exploratory structural equation modeling, negative emotion symptoms, perceived support, psychometrics, Stress Overload Scale

Introduction
The World Health Organization claimed that stress is the “Health Epidemic of the 21st Century.” It considered its negative associations with physical health (Lagraauw et al., 2015), well-being (Saleh et al., 2017), academic performance (Elias et al., 2011), and work productivity (Siu, 2003) in various contexts. Most theories (e.g. Appraisal Theory and Conservation of Resources Theory) interpreted stress as an imbalanced status (McGrath, 1970), which occurs when “environmental demands tax or exceed adaptive capacity” of individuals (Cohen et al., 1995: 3). Stress-related symptoms are further induced when inadequate resources meet high demands (Cohen et al., 2016). For instance, a longitudinal survey including 360 firefighters found that increased job demands (e.g. physical workload) and decreased job resources (e.g. social support) predicted high levels of depression and sleeping disorders (Airila et al., 2014). Thus, stress level should be evaluated in terms of “demand” and “resource.” However, most extant stress assessments have measured only...
one aspect, such as personal resources, in the Resilience Scale, and environmental demands in the Weekly Stress Inventory. Such biased evaluations might lead to inaccuracy in explaining symptoms and promoting health (Amirkhan, 2012).

Amirkhan (2016) recently constructed a 10-item Stress Overload Scale-Short (SOS-S) to measure stress from event load (EL; objective items that measure excessive demands) and personal vulnerability (PV; subjective items that measure depleted resources; Amirkhan et al., 2015). A 1-month longitudinal comparison study on 285 community residents indicated the higher correlation coefficients of this instrument to depression and physical symptoms than that of classic stress measurements, such as the 14-item Perceived Stress Scale (Amirkhan, 2012). Specifically, the brevity of the 10-item instrument is considered advantageous for screening purposes and clinical applications. Moreover, Amirkhan (2016) conducted a two-wave design with 156 community participants and demonstrated that SOS-S can identify individuals at risk of developing stress-related symptoms.

Furthermore, SOS-S has been applied in different countries and populations, including America (Amirkhan, 2016), South Africa (Wilson et al., 2017), and China (Duan and Mu, 2017; Su and Guo, 2014). This provided empirical basis for its validity and effectiveness. All studies reported satisfactory psychometric properties (e.g. high internal consistency coefficients, clear and solid factor structure, and adequate incremental and discriminant validity) of the two-factor structure of SOS-S. For instance, Duan and Mu (2017) conducted a cross-sectional evaluation using community and student samples to demonstrate the high coefficient omega ($\omega > .87$) and the two-factor-related models of SOS-S. Strong factorial invariance was also confirmed across these samples (Duan and Mu, 2017). Although previous studies have suggested that the two-factor-related model showed better fit indices than the signal-factor and second-order two-factor models (Duan and Mu, 2017), the goodness-of-fit of the aforementioned models were all statistically acceptable. We cannot rule out the possibility that the characteristics of the samples and/or statistical method resulted in model differences. Moreover, an increasing number of studies have indicated that the confirmatory factor analysis (CFA) method has a limited effect on validating multifactor psychological scales (Marsh et al., 2014). The highly restrictive CFA method specifies that each item loads on only one factor, and nontrivial cross-loadings are restricted as zero. Results not only degraded the model fit, but it also overestimated factor correlations (Marsh et al., 2014). Exploratory structural equation modeling (ESEM) was recently developed to compensate for the shortcomings of CFA. The ESEM benefits from the merits of the exploratory factor analysis and CFA (Marsh et al., 2011) and tends to produce more realistically estimated factor correlations and better model fit than CFA (e.g. Duan et al., 2017; Joshanloo et al., 2016). Accordingly, this study initially aimed to further evaluate the construct validity of SOS-S using ESEM.

The SOS-S has two subscales, namely PV and EL, which can be adopted to screen individuals at different risk levels. According to the mean scores of a sample with the two subscales, the entire sample can be divided into four subgroups, that is, groups with high PV and high EL, high PV and low EL, low PV and high EL, and low PV and low EL (Amirkhan, 2016). These four subgroups are further determined as high risk (high PV, high EL), fragile (high PV, low EL), challenged (low PV, high EL), and low risk (low PV, low EL) groups. Existing studies have shown that individuals in the high-risk group suffered serious physical symptoms (e.g. sick days) and behavioral disorders (e.g. lost temper) (Amirkhan, 2016) and exhibited diminished well-being (Duan and Mu, 2017) than those from the other three groups. To validate the screening role of SOS-S, this study further examined its efficacy on the negative emotion symptoms. Moreover, Hobfoll’s conservation of resources theory implies that social support, as emotional and interpersonal support, can reduce loss of resources and increase
resistance to demands. Thus, social support is expected to moderate the relationship between stress and mental health. In a recent study, Duan and Mu (2017) found that individuals with high SOS-S scores reported increased depression levels. The relationship between SOS-S and depression was further moderated by social support. However, the different moderated roles of EL and PV have not been examined in the literature. As mentioned above, the EL subscale reflects external demands, whereas the PV subscale is associated with internal resources. Social support is a multidimensional construct that includes emotional support, self-esteem support, tangible aid, and informational support (Cutrona and Russell, 1990). The appropriateness of social support depends on the correspondence between the social support type and stress (Cohen and Syme, 1985; Cutrona and Russell, 1990). Therefore, it is meaningful to distinguish the possible difference between the two subscales (i.e. EL and PV) with the addition of social support as a buffer in dysfunction defense.

In summary, this study mainly aims to solidify the psychometric characteristics of SOS-S through ESEM, and to examine the different functions of the EL and PV subscales. This study expects that (a) ESEM reveals a clear two-factor-related structure of SOS-S and shows better goodness-of-fit than CFA, (b) SOS-S can be used to screen high-risk individuals who exhibit more negative emotion symptoms than the other three groups, and (c) perceived social support differently moderates the relationships between EL/PV and negative emotion symptoms.

**Methods**

**Study design**

The scale validation is based on two samples recruited from communities and colleges. The first sample included residents of five conveniently selected communities in Wuhan, China. The residents voluntarily participated in this survey, and they completed the questionnaire including demographic information and SOS-S. The second sample was recruited from eight different universities in Wuhan. Wuhan has 82 universities including key and ordinary universities, as well as secondary colleges. Stratified sampling was used to select eight universities. The participants were students who confirmed they were above 18 years old, and completed a questionnaire package, including demographic information, SOS-S, Multidimensional Scale of Perceived Social Support, and a Depression Anxiety Stress Scale (DASS; see below). An online questionnaire platform was used to collect data. Students voluntarily visited this platform and completed the aforementioned questionnaire. Participants signed consent forms before completing the questionnaires. An ethics approval was granted by the Human Subjects Ethics Sub-Committee of the Department of Sociology, Wuhan University.

**Participants**

The community sample comprised 363 males ($M=41.37$ years, $SD=6.605$, range: 24–58) and 266 females ($M=30.38$ years, $SD=9.144$, range: 19–58). The student sample ($N=495$) consisted of 38.8 percent male ($M=22.07$ years, $SD=1.381$, range: 18–25) and 61.2 percent female participants ($M=21.78$ years, $SD=1.479$, range: 18–25).

**Instruments**

**Stress assessment.** Stress level was measured by a 10-item SOS-S (Amirkhan, 2016), in which the EL subscale evaluated perceived environmental demands (five items, e.g. *In the past week, have you felt like things kept piling up*), and the PV subscale assessed perceived inadequacy to environmental demands (five items, e.g. *In the past week, have you felt like just giving up*). Participants were asked how they felt in the past week, and were required to rate statements from 1 (not at all) to 5 (a lot) in each subscale. The mean scores of EL and PV were calculated to reflect stress levels.

**Perceived social support.** Perceived social support was evaluated using the Multidimensional
Scale of Perceived Social Support (MPSS) on a 7-point Likert scale, with 1 meaning very strongly disagree, to 7 denoting very strongly agree. The MPSS was developed by Zimet et al. (1988) to assess subjective social support from family, friends, and others. Higher mean scores reflect a greater degree of perceived social support. The MPSS showed the psychometric characteristics of the Chinese population (Zhou et al., 2015).

**Negative emotion symptoms.** Negative emotion symptoms were evaluated by the 21-item DASS. Lovibond and Lovibond (1995) indicated that the total mean score of DASS measures the negative emotion symptoms of the previous week. Respondents completed DASS on a 4-point scale from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time). High mean scores in DASS indicate severely negative emotion symptoms. The DASS displayed good reliability and validity when used in Chinese population (Wang et al., 2016).

**Analytical strategy**

First, CFA and ESEM were simultaneously conducted on the community sample to evaluate the factor structure of SOS-S. Target rotation was used in ESEM (Asparouhov and Muthen, 2009). To evaluate the model fits of CFA and ESEM, standardized root mean square residual (SRMR), comparative fit index (CFI), and root mean square error of approximation (RMSEA) were reported, in which CFI > .95, SRMR < .08, and RMSEA < .08 were used as evaluation criteria (Hu and Bentler, 1998). In addition, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were compared between CFA and ESEM. Models with smaller AIC and BIC were better than those with higher AIC and BIC (Joshihanloo and Jovanovic, 2016). We expected that ESEM would indicate a better model fit, and a smaller AIC and BIC than CFA. Second, statistics were calculated for a detailed description of the student sample. Reliability was evaluated using McDonald’s omega (ω).

Peterson’s correlations were calculated among these variables. PV and EL were expected to be positively related to negative emotion symptoms, and negatively related to social support. Third, we used the mean score of the student sample as a cut-point to divide PV/EL into high and/or low levels of PV and/or EL. The student sample was then divided into four subgroups: high risk (i.e. high PV, high EL), fragile (i.e. high PV, low EL), challenged (i.e. low PV, high EL), and low risk (low PV, low EL). The general linear model (GLM) was utilized to examine the main effect of factorial PV/EL for negative emotion symptoms. The main effect of factorial PV/EL for negative emotion symptoms was expected to be significant. Simple effect tests were then conducted to examine the screening effect of SOS-S. In line with previous studies, we expected that people in the high-risk group suffered from greater negative emotion symptoms than those in the other three groups (Amirkhan, 2016; Duan and Mu, 2017). Finally, latent moderated structural equations (LSM) were conducted to examine the moderated role of social support between PV/EL and negative emotion symptoms. The variable scores were mean-centered before a model estimation. The product of SOS-S and social support was then included as an antecedent of negative emotion symptoms along with SOS-S and social support. Social support measured by MPSS can provide emotion and self-esteem support, thereby enhancing personal resources. Thus, we expected that the buffering effect of social support in the relationship between PV and negative emotion symptoms to be greater than that between EL and negative emotion symptoms. Data analyses were conducted using JASP 0.8.12 (JASP Team, 2017) and M-plus 7.4 (Muthén and Muthén, 1998–2012).

**Results**

**ESEM and CFA results**

We initially computed the CFA model with maximum likelihood (ML) estimation. Each item was theoretically loaded on a corresponding PV or EL
subscale. The CFA results showed a marginally acceptable goodness-of-fit ($\chi^2=243.708$, $df=34$, $CFI=.926$, $SRMR=.050$, $RMSEA=.099$, $AIC=17,558.995$, $BIC=17,696.763$). ESEM was then conducted with ML estimation by specifying all items load on PV and EL latent factors. As expected, the fit indices of the ESEM model significantly improved ($\chi^2=128.390$, $df=26$, $CFI=.964$, $SRMR=.027$, $RMSEA=.079$, $AIC=17,459.677$, $BIC=17,632.998$), and the values of AIC and BIC were smaller than those in CFA. The factor loadings of the two models were all significant on their target factors (loadings > .49, $p<.001$), as displayed in Table 1. In addition, the ESEM model ($r=.674$) obtained a smaller factor correlation between PV and EL than that in CFA ($r=.794$), which meant that ESEM has greater factor distinctiveness (Joshanloo et al., 2016; Joshanloo and Jovanovic, 2016).

### Table 1. Standardized factor loadings ($N=629$).

<table>
<thead>
<tr>
<th></th>
<th>CFA</th>
<th>ESEM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Personal vulnerability</td>
<td>Event load</td>
</tr>
<tr>
<td>Item 1</td>
<td>.683***</td>
<td>.554***</td>
</tr>
<tr>
<td>Item 2</td>
<td>.709***</td>
<td>.606***</td>
</tr>
<tr>
<td>Item 3</td>
<td>.776***</td>
<td>.643***</td>
</tr>
<tr>
<td>Item 4</td>
<td>.748***</td>
<td>.746***</td>
</tr>
<tr>
<td>Item 5</td>
<td>.684***</td>
<td>.783***</td>
</tr>
<tr>
<td>Event load</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 6</td>
<td>.657***</td>
<td>.220**</td>
</tr>
<tr>
<td>Item 7</td>
<td>.645***</td>
<td>−.134***</td>
</tr>
<tr>
<td>Item 8</td>
<td>.623***</td>
<td>−.292***</td>
</tr>
<tr>
<td>Item 9</td>
<td>.806***</td>
<td>.223***</td>
</tr>
<tr>
<td>Item 10</td>
<td>.791***</td>
<td>.366***</td>
</tr>
</tbody>
</table>

CFA: confirmatory factor analysis; ESEM: exploratory structural equation modeling. *$p<.05$, **$p<.01$, ***$p<.001$.

Descriptive statistics, correlations, and internal reliability

Means, standard deviation (SD), McDonald’s omega (ω), and Peterson’s correlations are displayed in Table 2. All scales in the current sample indicated good internal reliability (ω > .807). As expected, both PV and EL were positively related to negative emotion symptoms ($r_{PV}=.645$, $p<.001$; $r_{EL}=.526$, $p<.001$), but negatively associated to perceived social support ($r_{PV}=-.165$, $p<.001$; $r_{EL}=-.107$, $p<.05$). PV also showed higher correlation to negative emotion symptoms than EL.

Screening role of SOS-S

The student sample was further divided into four subgroups (i.e. 180 in low risk, 32 in fragile, 85 in challenge, and 198 in high risk) based on the mean scores of the PV ($M=2.42$) and EL ($M=2.96$) subscales. GLM analyses of variance suggested the significant main effect for PV, with the high group reporting serious negative emotion symptoms ($F(1, 491)=71.850$, $p<.001$). Significant main effect of EL was also observed, with the high group showing serious negative emotion symptoms ($F(1, 491)=51.722$, $p<.001$). The interaction of PV and EL for negative emotion symptoms was insignificant ($F(1, 489)=.220$, $p=.248$). Simple effect tests suggested that negative emotion symptoms in the high-risk group ($M=1.12$, $SD=0.24$) were significantly higher than those in the fragile ($M=0.73$, $SD=0.38$), challenge ($M=0.66$, $SD=0.39$), and low risk ($M=0.38$, $SD=0.24$) groups.
Latent moderated structural equations of social support

Two moderation models were constructed to examine the different functions of PV and EL. The moderated role of social support between PV and negative emotion symptoms was initially examined. Results showed that PV ($B = .699$, $p < .001$) and PSS ($B = -0.157$, $p < .001$) significantly influenced negative emotion symptoms and the PV × PSS interaction ($B = -0.083$, $p < .05$) was significant (Figure 1). These results indicated that social support moderated the relationship between PV and negative emotion symptoms. Thus, the negative effect of PV on negative emotion symptoms was reduced when individuals perceived that they had high levels of social support. The moderation model explained 59.49 percent of variance in the negative emotion symptoms.

The moderated role of social support between EL and negative emotion symptoms was also examined. Results showed that EL ($B = .577$, $p < .001$) and PSS ($B = -0.236$, $p < .001$) were significantly related to negative emotion symptoms. However, the PV × PSS interaction was insignificant ($B = -0.066$, $p = .75$). These results demonstrated that social support did not moderate the effect of PV on negative emotion symptoms.

Discussion

This research aimed to solidify the psychometric properties of SOS-S through ESEM, and to examine the different functions of the EL and PV subscales. Findings confirmed the construct validity of the SOS-S in the context of China. The ESEM resulted in satisfactory fit indices, which substantiated a theory-based two-factor scale for assessing the stress levels of Chinese people. Furthermore, evidence was obtained for the screening effect of SOS-S. We found that people in the high-risk group exhibited more serious negative emotion symptoms than those in the other three groups. Finally, moderation analysis indicated that social support moderated the relationship between PV and negative emotion symptoms, instead of EL.

Consistent with previous studies that compared the performance of CFA and ESEM in evaluating the psychometric properties of a multifactorial scale (Joshanloo and Jovanovic, 2016; Marsh et al., 2014), ESEM performed better in capturing the factor structure of SOS-S than CFA. This yielded a better fit and smaller factor correlation. In the ESEM model, items were free to load on all factors, which meant that cross-loadings were not specified to zero. Thus, ESEM was less restrictive than CFA, which specifies cross-loadings at zero. This provides ESEM results with better fit and more realistic factor correlation (Marsh et al., 2014; Morin et al., 2013). This study supported the construct validity of SOS-S, and the psychometric properties of this scale were further solidified.

For the practical application of SOS-S, we found that individuals in the high-risk group suffered more serious negative emotion symptoms than those in the other three groups.

<table>
<thead>
<tr>
<th></th>
<th>Personal vulnerability</th>
<th>Event load</th>
<th>Perceived social support</th>
<th>Negative emotion symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal vulnerability</td>
<td>–</td>
<td>.666***</td>
<td>-.165***</td>
<td>.645***</td>
</tr>
<tr>
<td>Event load</td>
<td>–</td>
<td>-.107*</td>
<td>.526***</td>
<td>.254***</td>
</tr>
<tr>
<td>Perceived social support</td>
<td>–</td>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Negative emotion symptoms</td>
<td>–</td>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.421</td>
<td>2.958</td>
<td>5.198</td>
<td>0.746</td>
</tr>
<tr>
<td>SD</td>
<td>0.843</td>
<td>0.876</td>
<td>0.999</td>
<td>0.518</td>
</tr>
<tr>
<td>McDonald’s omega</td>
<td>.825</td>
<td>.807</td>
<td>.931</td>
<td>.934</td>
</tr>
</tbody>
</table>

*p < .05, ***p < .001.

Table 2. Descriptive statistics and coefficient omega for each scale in student sample (N = 495).
Results indicated that SOS-S performed well in screening high-risk individuals who exhibited more negative emotion symptoms from the general population. Chinese individuals generally tended to suppress and deny their psychological symptoms (Cheung, 1982; Ryder et al., 2008). Based on the screening effect, SOS-S can be used as a diagnostic tool to distinguish symptomatic from healthy people, and to provide a reference for health professionals. Specially, GLM analyses of variance showed the significant main effects of PV and EL on negative emotion symptoms, but the interaction of PV and EL was insignificant. The insignificant interaction effect of PV and EL demonstrated that PV and EL, although highly correlated, have unique and discrete contributions to stress overload. Thus, PV and EL affect negative emotion symptoms differently. Therefore, the effects of PV and EL should be alleviated variably.

Moderation analysis further confirmed the different roles of PV and EL in influencing negative emotion symptoms, and specific support-based interventions should be used to buffer the negative effect of different stress. Results suggested that social support only significantly moderated the relationship between PV and negative emotion symptoms, instead of EL. This result can be attributed to the specific type of social support measured in this study. Previous studies have shown that the appropriateness of social support depends on the correspondence between social support type and stress (Cohen and Syme, 1985; Cutrona and Russell, 1990). In this study, social support was measured by MSPSS, which subjectively assesses social support. Social support provides self-esteem and emotional support to individuals with compromised resources, or bolsters their perceived ability to cope with impinging demands, thereby reducing PV (Cobb, 1976; Cohen and Wills, 1985; Zimet et al., 1988).

Figure 1. Moderation effect of perceived social support between personal vulnerability and negative emotion symptoms (N = 495).

PV: personal vulnerability; PSS: Multidimensional Scale of Perceived Social Support; DASS: negative emotion symptoms.
However, EL may be mainly reduced by tangible aid (Cutrona and Russell, 1990). Future studies may focus on different types of social support to examine the above hypothesis.

However, this study has several limitations. First, this study only collected psychological symptoms to evaluate the screening effect of SOS-S. Future studies should consider investigating the screening effect for physical symptoms (e.g. headache, upset stomach, and insomnia) and behavioral disorders (e.g. losing temper, overeating, and increased alcohol use). Second, data were collected using a cross-sectional design, and thus the predictive validity of SOS-S could not be examined. Third, demographic information on marital status, education level, and occupation were not collected from community residents; thus, their effects were not controlled in the analysis. Fourth, community residents were recruited at convenience from selected communities, and students were recruited from a specific city; thus, further studies should collect diverse and random samples to enhance representativeness. Finally, the student sample was collected through an online questionnaire platform. Although previous studies have reported the equivalence between Internet and paper-and-pencil data collection methods, students were self-selected to participate in this survey, which may lead to bias (Weigold et al., 2013).

**Declaration of Conflicting Interests**
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by National Social Science Foundation of China (Grant No. 17CSH073) and Wuhan University Humanities and Social Sciences Academic Development Program for Young Scholars “Sociology of Happiness and Positive Education” (Grant No. Whu2016019).

**References**


Duan W, Li J and Mu W (2017) Psychometric characteristics of strengths knowledge scale and strengths use scale among adolescents. *Journal*
of Psychoeducational Assessment. Epub ahead of print 5 May. DOI: 10.1177/0734282917705593.
JASP Team (2017) JASP (version 0.8.1.2). Available at: https://jasp-stats.org/