Journal of Affective Disorders 238 (2018) 136–141

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

# Factor structure of the Geriatric Depression Scale and measurement invariance across gender among Chinese elders

Jiayue He, Xue Zhong, Shuqiao Yao*

*Medical Psychological Institute, Second Xiangya Hospital, Central South University, Changsha 410011, People's Republic of China*

ARTICLE INFO

*Keywords:*
Geriatric Depression Scale
Depression
Measurement invariance
Gender differences
Chinese elders

ABSTRACT

*Background:* The Geriatric Depression Scale (GDS) is commonly used in research and clinical settings for screening of depression. The current study aimed to examine the best-fit factor structure model of the GDS among Chinese elders and to evaluate the measurement invariance of the GDS across genders.

*Methods:* Participants included 1,553 elderly residents from the Hunan, Shandong, and Beijing provinces. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted in tandem to determine the structure of the GDS on a large scale. Multigroup CFA ($N = 1553$, 45.24% male, mean = 71.33 ± 8.06) was utilized to test the measurement invariance of the depressive symptom structure, which was generated by EFA and confirmed by CFA across gender.

*Results:* A three-factor model with 15 depression, 9 apathy and 4 vigor items presented the best fit indices. Measurement invariance of the new proposed model across gender was supported fully assuming different degrees of invariance.

*Limitations:* Our sample was entirely Chinese, and thus may not be representative of populations outside of China. Our results are based on a cross-sectional study, which did not take into consideration changes that may occur over time within individuals.

*Conclusions:* A three-factor model best fits the depressive symptom structure of the GDS among elderly Chinese, with measurement invariance across genders.

## 1. Introduction

The recent phenomenon of population aging has brought increased concern about the physical and mental health of the elderly. Depression is a common mental disorder in older populations (Fiske et al., 2009). Reported prevalence rates of depression in general population aged 75 years and above range from 4.6% to 9.3% (Luppa et al., 2012). Elderly individuals with clinically significant depressive symptoms are faced with a number of negative consequences such as social functional decline, marked disability, and decreased quality of life (Fiske et al., 2009).

The Geriatric Depression Scale (GDS), which was the first depression scale designed specifically for geriatric population (Yesavage et al., 1983), has become widely used for depression assessment among geriatric population (Adams et al., 2004).It has been translated from its original English version into more than 30 different languages such as Chinese (Chiu et al., 1994; Lai et al., 2005), Greek (Fountoulakis et al., 1999), Japanese (Onishi et al., 2004;Schreiner et al., 2001), Italian (Incalzi et al., 2003), Turkish (Ertan and Eker, 2000), and Korean (Jang et al., 2005). Relative to other depression assessment tools used in elderly population, the GDS has many advantages. Firstly it is easy for subjects to complete owing to the items being written in simple language with a yes/no response format. Secondly, the GDS does not contain somatic symptom items such as loss of appetite and sleep problems (Kessler,et al., 1992). These somatic complaints are likely to be shared by those with medical disorders, older adults, or people from certain cultural groups (Lewis-Fernández et al., 2005; Kalibatseva and Leong, 2011), making them less valid indicators for depression. Finally, previous studies suggested that the reliability and validity of the GDS was satisfactory in elderly populations from different countries (Burke et al., 1995; Adams et al., 2004; Mui et al., 2003; Malakouti et al., 2006). For example, the GDS has shown good reliability and validity among Iranian elders (Malakouti et al., 2006), and both long and short forms of the GDS have been reported to be reliable for assessing

*Abbreviations:* GDS, the Geriatric Depression Scale; EFA, exploratory factor analysis; CFA, confirmatory factor analysis; SD, standard deviation; SPSS, Statistical Product and Service Solutions; RMSEA, root-mean-square error of approximation; TLI, Tucker–Lewis index; CFI, comparative fit index; WLSMV, weighted least squares with mean and variance adjustment; BIC, Bayesian information criterion; PCA, Principal Component Analysis

* Corresponding author.
*E-mail address:* shuqiaoyao@csu.edu.cn (S. Yao).

https://doi.org/10.1016/j.jad.2018.04.100
Received 15 November 2017; Received in revised form 1 March 2018; Accepted 8 April 2018
Available online 17 May 2018

depression among community-dwelling elderly Asian immigrants (Mui et al., 2003).

Sheik et al. (1991) proposed that the 30 GDS items be categorized into five symptom groups: depression (9 items); vigor (6 items); positive (6 items); agitation (3 items); and social withdrawal (2 items). They found that four items (i.e. #1 "satisfied," #3 "life empty," #14"memory," and #17 "feel worthless") did not fit any of these factors. Since then, a number of other studies employing exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) have yielded different GDS factor structures. To date, studies have reported two (Chiu et al., 1994; Incalzi et al., 2003; Brown et al., 2007), three (Salamero and Marcos, 1992), four (Havins et al., 2012), five (Adams et al., 2001), six (Parmalee et al., 1989; Abraham et al., 1994; Adams et al., 2004), seven (Huang et al., 2017), and nine (Salamero and Marcos, 1992) factors depending upon different versions of the GDS, different language, and different samples (Kim et al., 2013). The five-factor solution differed substantially from Sheikh's original classification scheme.

Kim et al.(2013) conducted a meta-analysis in which they investigated the GDS factor structure with over 14,669 participants, from 26 GDS studies. Using EFA, their meta-analysis concluded that the full sample supported in a four-factor structure and that three factors of dysphoria, social withdrawal-apathy-cognitive impairment, and positive mood were commonly observed across different languages. However, this meta-analysis did not include articles using CFA. In order to examine the factor structure of a measurement tool throughly, researchers should do EFA firstly, and then apply CFA, just as Kim et al. (2013) recommended.

To develop the Chinese version of GDS, it is necessary for us to examine which factor structure model is more suitable for Chinese elderly, for which will be beneficial for increasing the comparability of results from different studies, and providing a standardized scoring method for them. Results of previous GDS factor analytic studies in Chinese populations have been mixed (Mui, 1996; Chiu et al., 1994; Lai, 2009; Lai et al., 2005; Lai et al., 2010). Chau et al..(2006) reported that the single-factor model of the GDS-30 fitted the data appropriately among Chinese elderly population, but Lai et al. (2005) reported a four-factor model in a sample of Chinese elderly population. Meanwhile, Lai concluded that despite the sociocultural differences among the Chinese elderly population in three locations, many of the depressive symptoms reported were related to several major common constructs, including positive mood, negative mood and agitation (Lai et al., 2005). However, these studies only employed CFA to analyze the factor structure of the GDS, which might miss an opportunity to uncover a Chinese-specific structure through the use of EFA (Kim et al., 2011). Thus, in accordance with the recommendation of Kim et al. (2013), we employed EFA to explore the structure of the GDS in one sample, and then used CFA to compare the EFA-derived model with other competing models demonstrated to be superior in previous studies in the other sample.

General population surveys since the 1970s have shown that women report higher levels of depression symptoms and disorders than men (Nolen-Hoeksema, 2001). This trend was confirmed with results from GDS, with women having significantly higher total GDS scores than men. However, it has not been established whether the GDS has measurement equivalence (i.e. similar effectiveness and accountability) across gender groups.

The main aims of this study were (1) to examine the factor structure of the 30-item GDS in a large representative elderly Chinese sample; and (2) to test the measurement invariance of the 30-item GDS across genders. We employed EFA to explore the structure of the GDS in sample 1, and then used CFA to compare the EFA-derived model with other competing models from previous studies in sample 2. After that, we examined measurement invariance of the best fitting model across gender.

## 2. Methods

### 2.1. Participants

A total of 1553 residents of the Hunan, Beijing, and Shandong provinces in China, including 704 men (45.24%) and 852 women (54.76%) were involved in this study. All participants were between 60 years old and 99 years old. The mean (standard deviation, SD) ages of the men and women were 71.33 (8.06) years and 70.64 (7.74) years, respectively.

### 2.2. Study design

The data were collected from participants in a district activity center by well-trained psychology postgraduate researchers. Administering researchers offered support for participants who lacked formal education or were visually impaired. All of the questionnaires were returned immediately upon completion. All participants provided informed consent, and the Ethics Committee of the Second Xiangya Hospital of Central South University approved the study.

### 2.3. Measurement

The Chinese version of the GDS. The GDS is a self-report scale consisting of 30 items answered using a yes/no response format. Twenty items represent a depressed response with a "yes" answer, and 10 items indicate a depressed response with a "no" answer. Values on the scale range from 0 to 30, with higher values indicating more symptoms of depression. Participants with a total score of 4 or less are considered normal. Those who score between 5 and 9 on the scale are considered to have mild depressive symptoms, and those who score 10 or higher are considered have moderate to severe depressive symptoms. The Chinese version of GDS has been validated (Chiu et al., 1994) and utilized extensively in Chinese studies (Chau et al., 2006). In our study, the scale showed high internal consistency with a Cronbach's alpha of 0.82.

### 2.4. Data analysis strategy

#### 2.4.1. Missing data

Of 1582 participants who were involved in the study, 29 failed to respond to all of the GDS items and were thus excluded from the analysis leaving an effective study cohort of 1553.

#### 2.4.2. Analytic steps

Our analyses contained three steps. Firstly, EFA was conducted on a randomly split-half of the whole sample ($N = 770$) to identify the best fitting factor model of the GDS in the present sample. Secondly, CFA was conducted on the remaining half of the sample ($N = 783$) to test the fit of several competing models, including the one generated from our EFA. Finally, measurement invariances of the best fitting model from the CFA were assessed in the full sample across gender.

*Stage1: EFA.* Descriptive statistics and EFA were performed in the SPSS, version 17 program (IBM, 2009). We used principal component factor analysis with orthogonal varimax rotation for the 30 items of the GDS. Items with a factor loading of 0.4 or greater were considered to contribute to the factor.

*Stage 2: CFA.* A series of CFAs were specified and estimated in Mplus 5.1 software (Muthen and Muthen, 1998–2007). Given that the items had binary response categories, maximum likelihood estimation was used for five or fewer response categories. Thus, the robust weighted least squares with mean and variance adjustment (WLSMV) estimator was used (Flora and Curran, 2004).

Our three-factor model generated from EFA was compared with other alternative models of the GDS-30 that were shown to be best fitting across previous studies (see Table 1). Model 1 represented