

# Factor Structure of the 12-item General Health Questionnaire (GHQ-12) Among Korean University Students

Boram Lee <sup>a</sup> , Yang Eun Kim <sup>b</sup> 

<sup>a</sup> Woosong University, Department of Early Childhood Education; <sup>b</sup> Woosong University, Department of Global Child Education, Deajeon, South Korea

## Abstract

**Background:** The prevalence of mental health issues among university student populations is a growing concern. Therefore, a reliable, standardized instrument is important for identifying students' symptoms. The General Health Questionnaire (GHQ-12), a 12-item self-report measure designed to screen for mental disorders in general practice and community settings, is a promising instrument. Although the GHQ-12's underlying factor structure has been investigated internationally in a variety of settings, the best factor structure is still unclear in South Korea, particularly in university settings. Therefore, this study investigated the GHQ-12's factor structure for a sample of South Korean university students.

**Methods:** In this research, 504 undergraduate students participated; they were aged 18-28 years and attended a four-year university course in South Korea. The collected data were subjected to confirmatory factor analysis (CFA), which tested previously proposed factor structures for the GHQ-12, including single-factor, correlated two-factor, and three-factor models.

**Results:** Confirmatory factor analysis indicated that Graetz's three-factor model, representing anxiety and depression, social dysfunction, and loss of confidence, fitted the data better than a unidimensional model or correlated two-factor models. Reliability analysis showed that the total GHQ-12 had adequate internal consistency.

**Conclusions:** The current study suggests that Graetz's three-factor solution provided the best fit to our data and that the Korean version of the GHQ-12 is a robust measure of general psychological distress symptoms. Moreover, our results further indicate the potential utility of also using the overall GHQ-12 score as a measure of general psychological distress, thus yielding significant advantages in both research and university settings.

## ARTICLE HISTORY

Received: May 18, 2020

Accepted: Jun 22, 2020

**KEYWORDS:** confirmatory factor analysis, factor structure, general health questionnaire, mental health, university students

## INTRODUCTION

In South Korea (hereafter "Korea"), various populations suffer from one or more mental disorders (e.g., depression and anxiety) [1-4]. Among these, university students are at high risk of developing mental health issues and conditions that might affect their mood, thinking, and behavior. Severe mental health issues are likely to affect students' ability to perform life activities such as working, studying, and interacting with community members [5]. In the Korean Council for University Education's 2018 survey on the psychological status and school adaptation of university students (conducted with 2600 students), 74.5% were identified with potential risk for anxiety and, 43.2% complained of difficulties caused by depression; in addition, 14.3% were found to be at risk for potential suicide [6]. From 2013 to 2017, there was an increase in the number of patients with mental health disorders among people in

their 20s, according to a survey by the Health Insurance Review and Assessment Agency [7]. Therefore, addressing mental health disorders is an important issue for Korea.

Most mental health disorders have their peak onset during young adulthood (the ages of 18-25 years), coinciding with students' time at university [8]. The transition from high school to tertiary education is a life-changing experience that can be extremely stressful for some. In addition to coping with academic pressure, students must deal with financial and career pressures, as well as stressful tasks of separation and individuation from their family of origin [8]. In this context, many university students experience their first onset of mental health problems or the exacerbation of their symptoms. However, mental disorders among young adult students often go unrecognized, resulting in delays in adequate treatment. Given that the university student

**Corresponding author:** Yang Eun Kim, E-Mail: yekim@wsu.ac.kr

**To cite this article:** Lee B, Kim YE. Factor Structure of the 12-item General Health Questionnaire (GHQ-12) Among Korean University Students. Psychiatry and Clinical Psychopharmacology 2020;30(3):248-253, DOI: 10.5455/PCP.20200518112914

population is particularly at risk for mental disorders, increased efforts need to include early screening for and identification of mental disorders.

One widely used screening instrument for common mental disorders is the 12-item General Health Questionnaire (GHQ-12, Goldberg, 1972) [9]. The GHQ-12 has demonstrated strong psychometric properties, indicated by internal consistency, retest reliability, and validity in various populations, including general adults [10, 11], primary care patients [12], multicultural populations [13, 14], elderly individuals [15], and adolescents [16]. Despite its adequate psychometric properties, the GHQ-12 does present areas of concern, particularly in its underlying factor structure, because its results have been inconclusive. The GHQ-12 was originally posited as having a unidimensional structure, with all 12 items loading on a single latent variable; a few empirical studies have supported this assumption [17, 18]. However, Banks et al. [19] found little evidence for the same and suggest that its structure does not provide sufficient information, in part, because of its simplicity. Several underlying dimensions of the GHQ-12 have been explored, with results suggesting a two - or three-factor structure. For example, in a survey with Australian teachers, using exploratory factor analysis (EFA), Andrich and van Schoubroeck [13] suggested that the GHQ-12 contains two dimensions in which positively worded items form one factor and negatively worded items form another. In Britain, Hu et al. [10] repeatedly found that Andrich and van Schoubroeck's two-factor model provided the best fit. In a study using Korean general adults, Park et al. [11] identified two factors, namely, general dysphoria and social dysfunction. In his study of young Australians, Graetz [20] proposed a three-factor structure of anxiety and depression, social dysfunction, and loss of confidence. With confirmatory factor analysis (CFA), it has been found in a number of studies that Graetz's three-factor model provides a significantly better fit to data than the one - and two-factor models [14, 21, 22].

In general, the GHQ-12's factor structure may differ because of sample characteristics (particularly, age) and practical issues such as mode of administration, original intent, and psychometric properties of candidate instruments because the instrument's psychometric properties can vary among population groups or cultures [23]. Thus, systematically assessing an instrument's psychometric properties prior to its widespread use within a specific population is important because the GHQ-12's factor structure might have significant implications for both clinicians and researchers. Additionally, as previously mentioned, often, the onset of mental disorders occurs during young adulthood. Therefore, university students are an important target population for mental health screening and treatment. Korean university students are of particular concern because this population is at significant risk for mental disorders; furthermore, the prevalence of disorders is also higher among them than the general population [6, 7]. A validated measurement of mental health is thus required for understanding susceptibility

to and etiology and treatment of pathological disorders. Furthermore, considering the GHQ-12 factor structures' varying outcomes in the studies reviewed above, the present study examined the factorial structure and psychometric properties of the GHQ-12's Korean version and evaluated its use among Korean university students.

## METHODS

### Participants

During the 2019-2020 academic year, 504 university students (i.e., 167 males; 337 females) were recruited from a four-year program in a private university in Korea. They belonged to varied disciplines, including architecture, culinary arts, design, education, social work, and public health. Participants' mean age was 20.2 years (SD = 1.63), and the age range was 18-28 years, with the majority of students (85.5%) being between 19 and 22 years old. The mean age of males was 20.6 years (SD = 2.0,6) and that of females was 20 years (SD = 1.34).

### Procedure

After receiving ethical approval from the Institutional Review Board (Protocol Code: 1041.549.200407-SB-91), six classes of each major were randomly selected during the 2019-2020 academic year. Through an arrangement with academic instructors, students were invited to complete questionnaires during scheduled class hours. Each time that the questionnaire was administered, the principal investigator personally provided instructions, explained the study's purpose, and collected participants' consent forms. Students were informed that their participation was strictly voluntary and that all information collected would be kept confidential. Written consent forms were obtained from all study participants. All participants were provided with paper questionnaires, the completion of which took an average of 10 minutes.

### Measure

**General Health Questionnaire-12:** The Korean version of the General Health Questionnaire (GHQ-12), translated by Park et al., was used to assess the level of psychological distress. Park et al. validated the Korean version of this scale in a study of Korean adults. The scale has 12 items, which are rated on a 4-point Likert-type scale; it assesses diagnosable psychiatric disorders in the community and non-psychiatric clinical settings [9]. Notably, the GHQ-12 consists of six positively worded items (e.g., "Have you recently been able to face up to your problems?") and six negatively worded items (e.g., "Have you recently been feeling unhappy and depressed?"); the positively worded items (i.e., 3 = More than usual; 0 = Not at all) and negatively worded items (i.e., 0 = Not at all; 3 = More than usual) have different response scales. Responses to all items were summed to a total score ranging from 0-36, with higher scores indicating poor mental health.

### Statistical Analysis

All statistics were analyzed using IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY, USA) and AMOS 20.0 (IBM Corp.) [24]. The GHQ-12's reliability was examined as between-item correlations and expressed using Cronbach's alpha.

Subsequently, CFA was performed based on four pre-existing models. The selection of competing models of the GHQ-12 for CFA was based on the review of relevant literature. The unidimensional model was used because of the GHQ-12's original unidimensional design. Other researchers have found support for two-factor solutions [10, 13]. The solution that has received more support, along with CFA studies, is Graetz's three-factor structure. Thus, we tested (1) the unidimensional model from the original GHQ-12; (2) the two-factor model proposed by Andrich and Van Schaubroeck –positively worded items (1, 3, 4, 7, 8, and 12) forming one factor and negatively worded items (2, 5, 6, 9, 10, and 11) forming another; (3) the two-factor model identified by Park et al: general dysphoria (items 2, 5, 6, 8, 9, 10, and 11) and social dysfunction (items 1, 3, 4, 7, and 12); and (4) Graetz's three-factor model: anxiety and depression (items 2, 5, 6, and 9); social dysfunction (items 1, 3, 4, 7, 8, and 12); and loss of confidence (items 10 and 11).

A CFA method was chosen as appropriate for investigating the GHQ-12's factor structure among Korean university students. CFA's goals are to test how well the data fit an a priori hypothesized model based on empirical evidence and to compare competing models' goodness of fit [25]. To assess models' fit to data, the following goodness-of-fit indices were used: the chi-square ( $\chi^2$ ) and its related degrees of freedom ( $df$ ); comparative fit index (CFI); goodness-of-fit index (GFI); root mean square error of approximation (RMSEA); and standardized root mean square residual (SRMR). In general, a chi-square per degree of freedom ( $\chi^2/df$ ) ratio of 5 or less indicates a good model fit [26]. For the CFI and GFI, values equal to or greater than 0.9 are considered to indicate acceptable model fit [27, 28]. RMSEA values of less than 0.05 indicate close fit [29]. However, Browne and Cudeck [30] suggested that RMSEA values up to 0.08 represent reasonable error fit of approximation. Finally, a value of SRMR less than 0.08 signifies adequate model fit [28]. In addition, the chi-square difference test was used to determine whether models differed significantly from one another.

## RESULTS

### Descriptive Statistics

Table 1 displays mean scores for the overall GHQ-12 and its individual items.

A mean GHQ-12 score of 12.59 (SD = 5.27) was obtained in the current sample, which was slightly high cutoff the cutoff point of 12.

**Table 1.** Descriptive statistics for GHQ-12 items

GHQ-12 items	Mean	SD
1. Able to concentrate	1.17	0.74
2. Loss of sleep over worry	0.95	0.83
3. Playing a useful part	1.39	0.86
4. Capable of making decisions	1.05	0.75
5. Felt constantly under strain	1.13	0.79
6. Couldn't overcome difficulties	0.72	0.75
7. Able to enjoy day-to-day activities	0.84	0.78
8. Able to face problems	2.05	0.81
9. Feeling unhappy and depressed	0.77	0.77
10. Losing confidence	0.86	0.81
11. Thinking of self as worthless	0.53	0.77
12. Feeling reasonably happy	1.13	0.82
Mean overall score	12.59	5.27

### Reliability and correlations analysis of GHQ-12

Table 2 shows each item's reliability and correlation with the overall scale and the Cronbach's alpha after eliminating the corresponding item. A Cronbach's alpha of 0.81 was observed for the GHQ-12's overall score, indicating our sample's acceptable level of internal consistency. Cronbach's alpha coefficients were also calculated on the 12 items that constitute the three subscales of the GHQ-12. For anxiety and depression, social dysfunction, and loss of confidence, Cronbach's alpha was 0.71, 0.80, and 0.54, respectively. Correlations with the overall scale ranged from 0.40 to 0.72, with item 5 ("Felt constantly under strain") having the lowest correlation coefficient. When an item was deleted, the Cronbach's alpha coefficients ranged from 0.75 to 0.83, indicating that each item was necessary and equally important.

**Table 2.** Correlation between items and overall GHQ-12 scale

GHQ-12 items	Correlation of item with overall scale	Cronbach's alpha if the item is eliminated
1. Able to concentrate	0.41	0.70
2. Loss of sleep over worry	0.55	0.76
3. Playing a useful part	0.49	0.79
4. Capable of making decisions	0.62	0.81
5. Felt constantly under strain	0.40	0.80
6. Couldn't overcome difficulties	0.58	0.83
7. Able to enjoy day-to-day activities	0.64	0.78
8. Able to face problems	0.47	0.80
9. Feeling unhappy and depressed	0.72	0.75
10. Losing confidence	0.70	0.82
11. Thinking of self as worthless	0.63	0.77
12. Feeling reasonably happy	0.60	0.81
Cronbach's alpha	0.81	

### Confirmatory Factor Analysis

A summary of model fit indices for alternative CFA models is displayed in Table 3. Underlying models' specification must be based on theoretical deliberations or evidence. First, the original one-factor model was tested to examine whether the GHQ-12 can be best understood as a unidimensional index of psychological distress. As shown in Table 3, the one-factor model fit the data poorly because none of the indices approached an acceptable level ( $\chi^2 = 427.2$ ,  $df = 54$ ;  $\chi^2 / df = 7.9$ ;  $p = 0.08$ ; CFI = 0.80; GFI = 0.89; RMSEA = 0.117 (90% CI = 0.106-0.127); SRMR = 0.096). The two-factor model, specifying a positively worded items factor and a negatively worded items factor, was found to fit the data marginally better than the one-factor model, as evidenced by the decreased value of chi square and the improved CFI, GFI, RMSEA, and SRMR; however, fit indices did not meet accepted fit criteria except for GFI ( $\chi^2 = 338.1$ ,  $df = 53$ ;  $\chi^2 / df = 6.4$ ;  $p = 0.11$ ; CFI = 0.85; GFI = 0.91; RMSEA = 0.103 (90% CI = 0.093-0.113); SRMR = 0.095). Another correlated two-factor model identified by Park et al. suggested a reasonable fit to the data on all indices but RMSEA and SRMR ( $\chi^2 = 243.8$ ,  $df = 52$ ;  $\chi^2 / df = 4.7$ ;  $p = 0.13$ ; CFI = 0.90; GFI = 0.93; RMSEA = 0.090 (90% CI = 0.067-0.088); SRMR = 0.081). In contrast, fit indices for Graetz's three-factor model provided an excellent fit to data and all fit indices were within recommended values ( $\chi^2 = 189.1$ ,  $df = 51$ ;  $\chi^2 / df = 3.7$ ;  $p = 0.17$ ; CFI = 0.93; GFI = 0.96; RMSEA = 0.073 (90% CI = 0.048-0.071); SRMR = 0.067). Additionally, Graetz's three-factor model proved better than all the models that were tested because the chi-square difference between the one-factor model ( $\chi^2 (3) = 238.1$ ,  $p < 0.001$ ), the two-factor model, specifying a positively worded items factor and a negatively worded items factor ( $\chi^2 (2) = 149.0$ ,  $p < 0.001$ ), and another correlated two-factor model forming general dysphoria and social dysfunction factors ( $\chi^2 (1) = 54.7$ ,  $p < 0.001$ ), differed statistically.

**Table 3.** Goodness-of-fit indices for GHQ-12 models in CFA

Models <sup>1</sup>	k	$\chi^2$	df	$\chi^2 / df$	p	CFI	GFI	RMSEA (90% CI)	SRMR
Model 1	12	427.2*	54	7.9	0.08	0.80	0.89	0.117 (0.106-0.127)	0.096
Model 2	12	338.1*	53	6.4	0.11	0.85	0.91	0.103 (0.093-0.113)	0.095
Model 3	12	243.8*	52	4.7	0.13	0.90	0.93	0.090 (0.067-0.088)	0.081
Model 4	12	189.1*	51	3.7	0.17	0.93	0.96	0.073 (0.048-0.071)	0.067

k=number of items; df=degrees of freedom; CFI=comparative fit index; GFI=goodness of fit index; RMSEA=root mean square error of approximation; SRMR=standardized root mean residual. Model 1: Unidimensional model (original); Model 2: Two-factor model (Andrich and Van Schaubroeck, 1989); Model 3: Two-factor model (Park, Kim, and Cho, 2012); Model 4 : Three-factor model (Graetz, 1991)

All items loaded significantly onto their respective factors. Standardized factor loadings were from 0.30 to 0.52, and all loadings were statistically significant ( $p < 0.001$ ). In the three-factor model, factors proposed by Graetz were highly correlated. The correlation was 0.80 between the first and second factors and 0.78 between the second and third factors. The correlations between the first and third factors were the strongest, at 0.91.

### DISCUSSION

This study assessed the factorial structure and psychometric properties of the GHQ-12's Korean version to use it for the Korean university student population. To the best of our knowledge, this is the first study of the GHQ-12 factor structure, to use a sample of Korean university students. Our findings offer evidence of the scale's validity as a reliable measure to assess mental health disorders among university students; the GHQ-12 Korean version is reliable, with satisfactory internal consistency for Korean university student samples. This is consistent with the values reported in other university populations [31, 32]. However, the low-reliability scores for the loss of confidence subscale ( $< 0.70$ ) may be because this subscale has only two items (items 10 and 12). Cronbach's alpha is strongly affected by the number of items, and scores that have a low number of items associated with them tend to have lower reliability [33]. The reliability of the loss of confidence subscale merits further investigation.

Next, the current study employed the CFA approach to testing the competing structural and theoretical conceptualizations of the GHQ-12. Consistent with findings from previous studies, Graetz's three-factor model that divided the items into anxiety and depression, social dysfunction, and loss of confidence constructs was found to provide the best fit to the data in the current study sample. In relation to previous research with a Korean adult sample by Park et al., the correlated two-factor solution provided an inadequate fit to the data in our study. The inconsistencies between the findings of Park et al. and ours may arise in part from the use of different statistical techniques and sample heterogeneity. For example, the two-factor solution presented by Park et al. emerged from an EFA of data from a nonclinical adult sample. Also, the age of the sample (18 to 64), though typical of questionnaire validation studies, was broader than that of the university students in our sample, whose ages ranged from 18 to 23.

Interestingly, items 4, 5, 8, 9, 10, and 12 showed relatively low factor loadings (ranging between 0.30 and 0.38). The low factor loadings might be due to item selection, or the low reliability of the observed variables may be a consequence of inadequate wording of the items, which would result in high measurement error and a small percentage of common variance [34]. However, there are situations in which retaining items with weak factor loadings are important and in which the unreliability problem is unavoidable, although the items are well written [34]. Decisions on removing items with weak or low

factor loadings are complex, as they involve considering both the pros and cons of reducing the number of items on an established questionnaire. On the one hand, item removal may make the instrument shorter, more precise, and more reliable, but on the other hand, comparisons between results obtained with the newly altered scale and those obtained by administering the original version may become impossible [35]. Despite the potential weakness of the scale that may arise from retaining every item, the original 12-item GHQ is the most widely used worldwide; thus, maintaining the original version seems desirable for comparative purposes. Moreover, removing the items with weak factor loadings may not represent an effective solution in the case at hand for two reasons: (1) the three-factor solution emerged as the most appropriate one in reproducing the observed data, as all loadings associated with their respective factors were significant at  $p < 0.001$ ; (2) internal consistency values were adequate for the overall scale, and no indication of item removal emerged as appropriate. Additionally, there are no clear criteria or consensus cutoff values for factor loadings. According to Bernard [36], however, a number of researchers have compromised on the cutoff point, with values from 0.30 to 0.49 considered worth including in the findings. Moreover, Child [37] suggested that as a rule of thumb, loadings with values equal to or greater than 0.30 are acceptable. Therefore, we deemed items with factor loadings of 0.30 or greater to be worthy of retention.

In addition, our results showed that all three factors were mutually highly correlated. Such high correlations, also documented by other studies, [14, 22] suggest that even if there are indeed three factors, in practice, distinguishing between them might be difficult because signs and symptoms are similar and common among the factors. For example, anxiety and depression can lead to sleep problems. Depression can, in turn, lead to more dysfunction and lower self-esteem [14]. These unclear boundaries between factors of psychological distress, together with evidence of strong correlation among factors, caused some authors to consider the practicality of using the overall score rather than overinterpreting the factors within the GHQ-12 [14, 22, 38]. Thus, the GHQ-12 should be used as a unitary measure, rather than a multidimensional construct.

In interpreting these findings, some limitations should be considered. First, the current findings' generalizability might be limited, given that this sample's demographics consisted of undergraduate students and that the factorial structure might differ among various clinical populations. In future research, therefore, replication with more heterogeneous samples is warranted. Moreover, the age range was quite restrictive, with only 9.4% of the sample belonging to the age category of 23-28 years. Although the age range is consistent with related research (e.g., Graetz), future studies would benefit from considering more mature students to further assess GHQ-12 factors' practical usefulness.

In conclusion, our findings provide evidence of the GHQ-12's valid factorial structure as applied to Korean university

students. The findings also suggest that the three-factor structure suggested by Graetz best fits the data from our sample. However, all three factors were mutually highly correlated, suggesting that the GHQ-12 is best understood as a unitary measure.

**Conflicts of Interest:** The authors have no potential conflicts of interest to disclose.

**Funding:** This work was supported by Woosong University.

## REFERENCES

- [1] Choi JH, Ju S, Kim KS, Kim M, Kim HJ, Yu M. A study on Korean university students' depression and anxiety. *Ind J Sci Technol* 2015;8(S8):1-9.
- [2] Cho MJ, Lee JY, Kim BS, Lee HW, Sohn JH. Prevalence of the major mental disorders among the Korean elderly. *J Korean Med Sci* 2011;26(1):1-10.
- [3] Jung S, Lee D, Park S. Subtypes of suicidal ideation in Korean adolescents: a multilevel latent profile analysis. *Aust N Z J Psychiatry* 2019;53(2):158-167.
- [4] Lee KJ, Kim JI. Relating factors for depression in Korean working women: Secondary Analysis of the Fifth Korean National Health and Nutrition Examination Survey (KNHANES V). *Asian Nurs Res* 2015;9(3):265-270.
- [5] Giamos D, Lee AYS, Suleiman A, Stuart H, Chen SP. Understanding campus culture and student coping strategies for mental health issues in five Canadian colleges and universities. *Can J Higher Ed* 2017;47(3):120-135.
- [6] Oh HY. Psychological states of university students. In Chang HS, editor. *Strategies for reducing the risk of psychological harm in university students. Proceedings of the 57<sup>th</sup> University Education Policy Conference; 2018 Mar 30; Seoul, Korea: Korean Council for University Education: 2002. p. 25-50.*
- [7] Health Insurance Review and Assessment Agency, Ad Hoc Committee on Health Insurance. *What age group does mental health affect the most? Seoul: Health Insurance Review and Assessment Service; 2018 Dec. 15 p.*
- [8] Pedrelli P, Nyer M, Yeung A, Zulauf C, Wilens T. College students: Mental health problems and treatment considerations. *Acad Psychiatry* 2015;39(5):503-511.
- [9] Goldberg DP. *The detection of psychiatric illness by questionnaire. Maudsley monograph no. 21. Oxford University Press, Oxford, 1972.*
- [10] Hu Y, Stewart-Brown S, Twigg Z, Weich S. Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychol Med* 2007;37(7):1005-1013.
- [11] Park JI, Kim YJ, Cho MJ. Factor structure of the 12-item General Health Questionnaire in the Korean general adult population. *J Korean Neuropsychiatr Assoc* 2012;51:178-184.
- [12] Schmitz N, Kruse J, Tress W. Psychometric properties of the General Health Questionnaire (GHQ-12) in a German primary care sample. *Acta Psychiatr Scand*

- 1999;100(6):462-468.
- [13] Andrich D, van Schoubroeck L. The general health questionnaire: a psychometric analysis using latent trait theory. *Psychol Med* 1989;19(2):469-485.
- [14] Padrón A, Galán I, Durbán M, Gandarillas A, Rodríguez-Artalejo F. Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents. *Qual Life Res* 2012;21(7):1291-1298.
- [15] Namjoo S, Shaghghi A, Sarbaksh P, Allahverdipour H, Pakpour AH. Psychometric properties of the General Health Questionnaire (GHQ-12) to be applied for the Iranian elder population. *Aging Ment Health* 2017;21(10):1047-1051.
- [16] Glozah FN, Pevalin DJ. Factor structure and psychometric properties of the General Health Questionnaire (GHQ-12) among Ghanaian adolescents. *J Child Adolesc Ment Health* 2015;27(1):53-57.
- [17] Hankins M. The factor structure of the twelve item General Health Questionnaire (GHQ-12): the result of negative phrasing? *Clin Pract Epidemiol Ment Health* 2008;24(4):10.
- [18] Fernandes HM, Vasconcelos-Raposo J. Factorial validity and invariance of the GHQ-12 among clinical and nonclinical samples. *Assessment* 2013;20(2):219-229.
- [19] Banks MH, Clegg CW, Jackson PR, Kemp NJ, Stafford EM, Wall TD. (1980). The use of the General Health Questionnaire as an indicator of mental health in occupational studies. *J Occupational Psychology* 1980;53(3):187-194.
- [20] Graetz B. Multidimensional properties of the General Health Questionnaire. *Soc Psychiatry Psychiatr Epidemiol* 1991;26(3):132-138.
- [21] Liang Y, Wang L, Yin X. The factor structure of the 12-item general health questionnaire (GHQ-12) in young Chinese civil servants. *Health Qual Life Outcomes* 2016;14:136.
- [22] Gao F, Luo N, Thumboo J, Fones C, Li SC, Cheung YB. Does the 12-item General Health Questionnaire contain multiple factors and do we need them? *Health Qual Life Outcomes* 2004;2:63.
- [23] Kim JH, Park EY. The Factor Structure of the Center for Epidemiologic Studies Depression Scale in stroke patients. *Top Stroke Rehabil* 2012;19(1):54-62.
- [24] Arbuckle JL. IBM SPSS Amos 20.0 user's guide. Mount Pleasant: Amos Development Corporation; 2011.
- [25] Byrne BM. Factor analytic models: viewing the structure of an assessment instrument from three Perspectives. *J Pers Assess* 2005;85(1):17-32.
- [26] Munro BH, Duffy ME, Brancato V, Newton S, Talbot L. *Statistical methods for health care research*. 5<sup>th</sup> Edition. Philadelphia: Lippincott Williams & Wilkins Co; 2005.
- [27] Marsh HW, Hau KT, Wen Z. In search of golden rules: comment on hypothesis testing approach cutoff values for fit indexes and dangers in overgeneralizing Hu & Bentler's (1999) findings. *Struct Equ Model* 2004;11(3):320-341.
- [28] Fan X, Sivo SA. Sensitivity of fit indices to model misspecification and model types. *Multivar Behav Res* 2007;42(3):509-529.
- [29] Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An empirical evaluation of the cutoff fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol Methods Res* 2008;36(4):462-494.
- [30] Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociol Methods Res* 1992;21(2):230-258.
- [31] Daradkeh TK, Ghubash R, ElRufaie OEF. Reliability, validity, and factor structure of the Arabic version of the 12-item General Health Questionnaire. *Psychol Rep* 2001;89(1):85-94.
- [32] Zulkefly NS, Baharudin R. Using the 12-item General Health Questionnaire (GHQ-12) to assess the psychological health of Malaysian college students. *Glob J Health Sci* 2010;2(1):73-80.
- [33] Tredoux C, Durrheim K. 3rd ed. *Numbers, hypotheses and conclusion: a course in statistics for the social sciences*. Lansdowne: UCT Press; 2013.
- [34] Ximénez C. Recovery of weak factor loadings when adding the mean structure in confirmatory factor analysis: A simulation study. *Front Psychol* 2016;6:1943.
- [35] Bottesi G, Ghisi M, Altoè G, Conforti E, Melli G, Sica C. The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Compr Psychiatry* 2015;60:170-181.
- [36] Bernard HR (2000), *Social research methods: Qualitative and quantitative approaches*, Sage, London, 2000.
- [37] Child D. *The essentials of factor analysis*. London, UK: A & C Black; 2006.
- [38] French DJ, Tait RJ. Measurement invariance in the General Health Questionnaire-12 in young Australian adolescents. *Eur Child Adolesc Psychiatry* 2004;13(1):1-7.