

Research Paper

## Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students

Deepesh Rathore<sup>1\*</sup>, Shailja Rana<sup>2</sup>, Prof. N. K. Chadha<sup>3</sup>

### ABSTRACT

**Objectives:** This study aimed to assess the reliability, validity, and factor structure of the GHQ-12 in the Indian undergraduate student population. Although it is the most widely used scale of psychological distress, the debate around its factor structure is still ongoing. **Method:** A sample of (n=432) Indian undergraduate students aged 17 to 22 years were taken. The English version of GHQ-12 was used. Internal consistency was assessed by Cronbach's alpha coefficient. Convergent validity was examined by the correlation coefficients with life satisfaction, subjective happiness, and a measure of general worry. A comparison of the different models of GHQ-12 proposed in the literature was made using Confirmatory Factor Analysis (CFA) with maximum likelihood estimates. **Result:** Cronbach's alpha (.784) reliability coefficient indicates adequate reliability. Correlation coefficients with life satisfaction ( $r = -.529$ ,  $p < .01$ ), subjective happiness ( $r = -.460$ ,  $p < .01$ ), and general worry ( $r = .424$ ,  $p < .01$ ) were also along the expected lines. Confirmatory Factor Analysis (CFA) showed a good fit to the data (Standardized Root Mean Square Residual (SRMR) = 0.0319, Comparative Fit Index (CFI) = 0.989, Tucker Lewis Index (TLI) = 0.982, Root Mean Square of Error Approximation (RMSEA) = 0.039). Hence, the findings indicated a seven items two-factor model consisting of Social Dysfunction and Anxiety/Depression.

**Keywords:** *General Health Questionnaire-12, Confirmatory Factor Analysis, Scale Validation, Indian Undergraduate Students*

As we enter the third year of the covid-19 pandemic, its impact on the mental health and well-being of people across different age groups, especially the middle adolescents (15 to 17 years) and late adolescents (18 to 21 years), have been severe (Sunitha & Gururaj, 2014). The disruption to routine social interactions and education due to

<sup>1</sup>Assistant Professor, Department of Psychology, Daulat Ram College, University of Delhi, Delhi, India  
<https://orcid.org/0000-0001-8511-3984>

<sup>2</sup>Assistant Professor, Department of Psychology, Zakir Husain College, University of Delhi, Delhi, India

<sup>3</sup>Professor, Department of Psychology, Daulat Ram College, University of Delhi, Delhi, India

\*Corresponding Author

Received: June 09, 2022; Revision Received: September 18, 2022; Accepted: September 30, 2022

## **Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

lockdown measures has increased their sense of isolation, uncertainty about the future, and affected their career aspirations. Compared to older adults, adolescents are more vulnerable. They tend not to seek help regarding their mental health issues due to peer pressure, lack of confidentiality, and knowledge to recognize mental health problems (Gaiha et al., 2020). Being a country with the world's largest adolescent population at 253 million, India cannot afford to carry the burden of untreated mental health problems. Therefore, there is a need to carry out a nationwide survey to estimate the prevalence of mental health problems, and hence there is a need for a valid screening tool in the Indian context.

The General Health Questionnaire (GHQ) is one of the most widely used mental health screening tools worldwide. The original version comprised 60 items, GHQ-60 (Goldberg, 1972); since then, many short versions have been developed, including GHQ-12, GHQ-28, and GHQ-30. Amongst all versions, the GHQ-12 is widely adopted due to its brevity and good psychometric properties as measured in clinical (Anjara et al., 2020; Mayhew et al., 2021) and nonclinical populations (Endsley et al., 2017; Liang et al., 2016). Along with this, it is a valuable predictor of mental health for different occupational groups (Campbell & Knowles, 2007; Laaksonen et al., 2007), clinical versus nonclinical population (Ruiz et al., 2017), gender differences in mental health (Van Droogenbroeck et al., 2018). Its different language versions have also been found to be effective (Graetz, 1991; Daradkseh et al., 2001; Winzer et al., 2014). To date, it has been translated into 38 different languages.

The GHQ-12 consists of 12 items, including both the positively phrased (6,7,10,11,12) and the negatively phrased (1,2,3,4,5,8,9) items. Three different scoring methods are available with the GHQ-12, dichotomous scoring (0-0-1-1), modified dichotomous scoring (0-1-1-1), and Likert scoring (0-1-2-3). Across all methods, a higher score reflects a greater degree of psychological distress.

There has been much debate regarding the factor structure of the GHQ-12. Some studies have claimed the GHQ-12 to be a unidimensional construct (Hankins, 2008; Ruiz et al., 2017) as Goldberg and Williams (1988) intended. Although other studies have pointed at a two-factor solution (Kalliath et al., 2004; Shevlin & Adamson, 2005; Winzer et al., 2014) and a three-factor solution (Padrón et al., 2012; Campbell et al., 2003; Tomás et al., 2015). Another issue surrounding the dimensionality of the GHQ-12 is the disagreements regarding the terminology and the qualitative meaning attached to the dimensions (Daradkeh et al., 2001; Hystad & Johnsen, 2020; Martin, 1999).

Although many previous studies have explored the reliability and validity of GHQ-12 in the Indian adult population (Endsley et al., 2017; Kashyap & Singh, 2017; Qin et al., 2018) and student population (Philip et al., 2021; Sriram et al., 1989), to the best of our knowledge, this is the first study of the GHQ-12 factor structure and psychometric properties on Indian undergraduate student population. Different models have been proposed in the literature, yet there is no consensus regarding the most appropriate factor structure of the GHQ-12. Moreover, the dimensions seem to vary from sample to sample and different population groups. Therefore, it becomes necessary to examine different factor structures that can be replicated in a wide variety of samples. Hence, the measure's widespread use and ongoing uncertainty have motivated us to conduct this study.

## METHODOLOGY

### *Sample*

The sample for this study comprised young undergraduate students of the University of Delhi, in the age group of 17 to 22 years of age. A total of 432 students (males= 149, females= 283) participated in this study. Convenience sampling method was used, and the data was collected online. Because the questionnaire was administered online, response pattern analysis was done to remove poor-quality data. That resulted in the exclusion of 10 respondents. The final sample comprised 422 respondents. The participants were briefed about the purpose of the study, anonymity, and confidentiality of the data. Informed consent was taken from all participants. The English-language version of the General Health Questionnaire (GHQ-12) containing 12 items was administered along with The Satisfaction with Life Scale, Subjective Happiness Scale, and the abbreviated version of the Penn State Worry Questionnaire. The scoring for GHQ-12 was done on a four-point Likert scale (0-1-2-3).

### *Measure*

- **GHQ-12:** The General Health Questionnaire – 12 (GHQ-12; Goldberg and Williams, 1988) comprises 12 items, which included both positively worded items (6, 7, 10, 11, 12) and negatively worded items (1, 2, 3, 4, 5, 8, 9). The Likert scoring method (0-1-2-3) was adopted to calculate the overall score. Where '0' indicated Often, '1' indicated Sometimes, '2' indicated Seldom, and '3' indicated Never. The total score obtained in this way ranged from 0 to 36, where a high score indicated mental illness, a cutoff point of 12 was taken, and any score higher than 12 was identified as representing some form of mental problems.
- **Satisfaction With Life Scale:** The Satisfaction with Life Scale (SWLS; Diener et al., 1985) comprises five items designed to measure the global cognitive judgment of one's satisfaction with life. Participants rate the statements on a 7-point Likert scale ranging from 1 "strongly disagree" to 7 "strongly agree." "A higher score reflects higher satisfaction with life.
- **Subjective Happiness Scale:** The Subjective Happiness Scale (SHS; Lyubomirsky & Lepper, 1999) comprises four items to measure one's global subjective happiness. Scoring is done on a 7-point Likert scale. Two items ask the respondents to rate the extent to which a brief description of happy and unhappy individuals characterizes them. The remaining two items ask the respondents to characterize themselves using both absolute ratings and ratings relative to peers.
- **The Penn State Worry Questionnaire Abbreviated:** The Penn State Worry Questionnaire-Abbreviated (PSWQ-A; Hopko et al., 2003) comprises eight items designed to measure an individual's level of worry regardless of the topic. The respondents rate the statements on a 5-point Likert scale ranging from 1 "not at all typical of me" to 5 "very typical of me." "The abbreviated version is derived from the original Penn State Worry Questionnaire (PSWQ; Meyer et al., 1990), which measures pathological worry in its generality, excessiveness, and uncontrollability.

### *Statistical analysis*

The obtained data were then analyzed for reliability and validity using IBM SPSS (version 23) software. Exploratory Factor Analysis (EFA) was carried out to explore the underlying factor structure of GHQ-12 using the principal component method with varimax rotation.

## Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students

Confirmatory Factor Analysis (CFA) was done using maximum likelihood estimation method to identify the best fit for the study sample in the models frequently cited in the literature namely, the original unidimensional model hypothesized by Goldberg and Williams (1988), two factor model by Andrich and Schoubroeck (1989) containing positively worded items (items 1, 3, 4, 7, 8, and 12) and negatively worded items (items 2, 5, 6, 9, 10, and 11), two factor model by Schmitz et al., (1999), Anxiety/Depression (items 1, 2, 6, 7, 10, and 11) and Social Performance (items 4, 5, 8, 9, and 12), two factor model by Politi et al., (1994), Dysphoria (items 2, 5, 6, 9, 10, and 11) and Social Dysfunction (items 1, 3, 4, 7, and 8), three factor model hypothesized by Worsley and Gribbin (1977), Anhedonia/Sleep disturbance (items 2 and 5), Social Performance (items 1, 3, 4, 7, 8, and 12), and Loss of Confidence (items 6, 9, 10, and 11), three factor model postulated by Farrell (1998), Anxiety (items 2, 5, 10, 11, and 12), Depression (items 1, 6, 7, 8, and 9) and Social Dysfunction (items 3 and 4), three factor model postulated by Martin (1999), Cope (items 1, 3, 4, 8, and 11), Stress (items 2, 5, and 7) and Depression (items 6, 9, and 12), three factor model hypothesized by Daradkeh et al. (2001), General Dysphoria (items 5, 6, 9, 10, and 11), Lack of Enjoyment (items 1, 2, 7, 8, and 12) and Social Dysfunction (items 3 and 4), and the three factor model postulated by Graetz (1991), Social Dysfunction (items 1, 3, 4, 7, 8, and 12), Anxiety/Depression (items 2, 5, 6, and 9), and Loss of Confidence (10 and 11). Individual model fit was examined using the commonly used goodness-of-fit statistics. Hu and Bentler (1999) have suggested an SRMR value less than .08 and an RMSEA value less than .06 for a good fit. MacCallum et al. (1996) have proposed that an RMSEA value below 0.08 indicates a good fit. More recently, Steiger (2007) has suggested the upper limit for a good fit to be 0.07. For SRMR, a value less than 0.05 is deemed a criterion for good fit (Byrne, 1998), but values less than 0.08 are also considered acceptable (Hu & Bentler, 1999). Hu and Bentler (1999) have recommended that for CFI, a value greater than 0.95 indicates a good fit, whereas values in the range of 0.90 to 0.95 are acceptable (Kline, 1998; McDonald & Ho, 2000). Finally, for TLI, a value greater than 0.95 indicates a good fit (Hu & Bentler, 1999). Moreover, a value in the range of 0.90 to 0.95 represents an adequate fit (Bentler & Bonett, 1980).

Two comparative fit measures were used to test the competing models, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). CFA was carried out with IBM SPSS AMOS (version 21) software.

## RESULTS

### *Descriptive statistics*

An overview of the descriptive statistics is presented in table 1. The final sample in the study comprised 422 respondents, out of which male respondents were 149 (35.30%), and female respondents were 273 (64.69%). In terms of age, 74 (49.7%) of the male respondents were in the age group of 17-18 years, 69 (46.3%) in the 19-20 years age group, and 6 (4%) in 21-22 years age group. Similarly, 145 (53.1%) of the female respondents were in the age group of 17-18 years, 118 (43.2%) in the 19-20 years age group, and 10 (3.6%) were in the 21-22 years age group. The male respondents obtained a mean score of 15.20 on GHQ-12 and a standard deviation of 7.67. The female respondents obtained similar scores with a mean score of 15.29 and a standard deviation of 7.45. The scores obtained were slightly positively skewed for both male (.402) and female (.353) respondents, with kurtosis of -.665 and -.607, respectively. Male and female respondents obtained similar mean scores on subjective happiness, 17.23 and 17.30, respectively. However, on the measures of life satisfaction and general level of worry, there is some difference, with male respondents scoring higher on

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

life satisfaction (20.23) but scoring low on worry (25.31). On the other hand, female respondents had scored lower on life satisfaction (19.89) but higher on worry (27.08).

**Table 1 Descriptive statistics**

Gender	Scales	M	SD	Median	Skewness	kurtosis
Male	GHQ-12	15.20	7.67	14	.402	-.665
	SHS	17.23	3.91	17	.137	.576
	SWLS	20.23	6.3	21	-.094	-.772
	PSWQ-A	25.31	7.77	26	-.044	-.714
Female	GHQ-12	15.29	7.45	15	.353	-.607
	SHS	17.30	3.95	17	-.223	.341
	SWLS	19.89	6.47	20	-.078	-.684
	PSWQ-A	27.08	7.47	28	-.347	-.544

*Note.* GHQ-12= General Health Questionnaire-12; SHS= Subjective Happiness Scale; SWLS= Satisfaction with Life Scale; PSWQ-A= Penn State Worry Questionnaire Abbreviated.

Table 2 shows the overall and individual item scores. An overall mean score of 15.26 (SD = 7.52) was obtained, higher than the cutoff point of 12. A total of 173 participants (41.2%) scored higher than the cutoff score. The mean score obtained on item 5 (1.54) was the highest, indicating that most students were under strain.

**Table 2 Descriptive statistics for GHQ-12 items**

GHQ-12 Items	M	SD	Scores			
			0	1	2	3
1. Been able to concentrate on what you are doing	1.53	0.882	13.5%	32.9%	40.8%	12.8%
2. Lost much sleep over worry	1.4	1.042	34.1%	0.2%	57.6%	8.1%
3. Felt you were playing a useful part in things	1.28	0.843	16.6%	47.4%	27.3%	8.8%
4. Felt capable of making decisions about things	0.97	0.913	36%	38.2%	18.7%	7.1%
5. Felt constantly under strain	1.54	0.995	18.7%	26.5%	36.5%	18.2%
6. Felt you could not overcome your difficulties	1.18	0.964	27.5%	39.1%	21.8%	11.6%
7. Been able to enjoy your normal day-to-day activities	1.43	0.882	14%	41.7%	31.8%	12.6%
8. Been able to face up to your problems	1.13	0.81	21.1%	51.4%	21.3%	6.2%
9. Been feeling unhappy and depressed	1.32	1.02	26.5%	29.1%	29.9%	14.5%
10. Been losing confidence in yourself	1.24	1.078	32.2%	28%	23.2%	16.6%
11. Been thinking of yourself as a worthless person	0.95	1.066	47.2%	22%	19%	11.8%
12. Been feeling reasonably happy, all things considered	1.29	0.834	13.7%	54%	21.6%	10.7%
Mean overall score	15.26	7.52				
Classified as cases (score $\geq 12$ ) %	41.2%					

**Reliability**

The reliability of GHQ-12 was assessed by calculating Cronbach’s alpha, inter-item correlation, and item-total correlation. Cronbach’s alpha was  $\alpha = .883$ , indicating good internal consistency.

**Exploratory Factor Analysis**

Table 3 shows the result of exploratory factor analysis using the principal component method with varimax rotation. First, sampling adequacy was assessed by the kaiser-Meyer-Olkin (KMO = 0.920) and Bartlett’s test of sphericity ( $p < .0001$ ); the result indicated that the sample size was adequate to run factor analysis. Next, a two-factor structure emerged based on the eigenvalue ( $>1$ ) and factor loadings observed in the rotated component matrix

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

obtained through varimax rotation with Kaiser normalization. As shown in Table 3, a two-factor solution emerged. The eigenvalue for both the factors was above 1. Factor 1 explained 44.29% of the variance, and factor 2 explained 10.53%. Overall, the cumulative variance of 54.83% was accounted for by both factors. In addition, item numbers 6, 7, 10, 11, and 12 showed cross-loadings on both factors.

**Table 3 Factor loadings**

GHQ-12 Items	Factor Loadings	
	Factor 1	Factor 2
Item no. 3	<b>.75</b>	.08
Item no. 4	<b>.74</b>	.11
Item no. 8	<b>.74</b>	.16
Item no. 1	<b>.63</b>	.23
Item no. 11	<b>.55</b>	<b>.53</b>
Item no. 10	<b>.52</b>	<b>.52</b>
Item no. 12	<b>.49</b>	<b>.46</b>
Item no. 5	.11	<b>.79</b>
Item no. 9	.31	<b>.76</b>
Item no. 2	-.00	<b>.71</b>
Item no. 6	<b>.48</b>	<b>.54</b>
Item no. 7	<b>.43</b>	<b>.45</b>
KMO	.920	
% of variance	44.29%	10.53%
Cumulative % of the variance	44.29%	54.83%
Eigenvalue	5.316	1.264

*Note. Factor loadings above .40 are in bold.*

After eliminating items with cross-loadings (Costello & Osborne, 2005), a seven items two-factor structure of the GHQ-12 emerged. Social Dysfunction (items 1, 3, 4, and 8) accounting for 44.194% of variance and Anxiety/Depression (items 2, 5, and 9) for 17.946%. The two factors accounted for the cumulative variance of 62.14%.

Furthermore, Cronbach's alpha was calculated for GHQ-7, and the value was  $\alpha = .784$ .

**Confirmatory factor analysis**

Figure 1 shows the standardized estimates and relationship between the two factors and the observed variables. The path coefficients ranged from .57 to .79, and the error variances ranged from .32 to .62. The covariance between Social Dysfunction and Anxiety/Depression was .57.

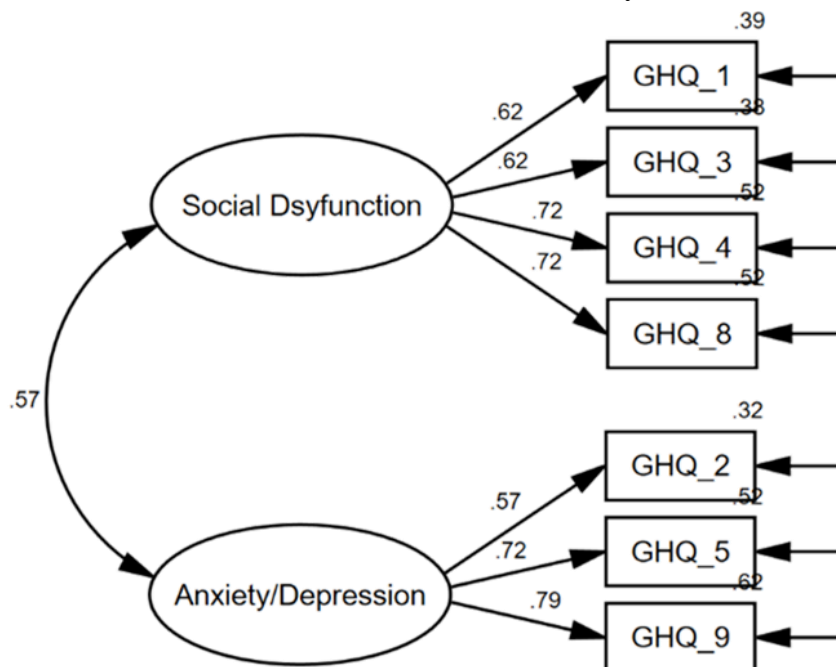
Table 4 presents the fit indices of nine different models most commonly cited in the literature and the exploratory factor analysis model of the present study. The comparative fit index (CFI) ranged from 0.889 to 0.947. The CFI values for the unidimensional model by Goldberg and Williams (1988) and the two-factor model by Schmitz et al. (1999) are less than the cutoff criteria for an adequate fit of 0.90. The Tucker-Lewis Index (TLI) ranged from 0.862 to 0.931. The TLI values for the unidimensional model by Goldberg and Williams (1988), the two-factor model by Schmitz et al. (1999), the three-factor model by Farrell (1998), and the three-factor model by Daradkeh et al. (2001) are less than the cutoff criteria (.90) for an adequate fit. The Standardized Root Mean Square Residual (SRMR)

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

values ranged from 0.049 to 0.061, indicating an adequate fit for all models with the model by Graetz (1991) in particular showing good fit (SRMR = 0.049) as per the criteria of SRMR less than 0.05. Furthermore, the Root Mean Square of Error Approximation (RMSEA) values ranged from 0.068 to 0.096, and the three-factor models by Worsley and Gribbin (1977), Graetz (1991), and Martin (1999) indicated adequate fit (RMSEA <0.08).

**Figure 1 GHQ-7: a two-factor model showing the standardized path coefficients and error variances**

In addition, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)



values were examined for the nine models. As shown in Table 4, the AIC value ranged from 192.15 to 310.02, and the BIC value ranged from 293.28 to 411.16. For both, the criterion smaller value represents a better comparative fit. However, the GHQ-7 model hypothesized in this study showed the least values for AIC (51.35) and BIC (112.03).

**Table 4 Goodness-of-fit indexes for GHQ-12 models**

Models	$\chi^2$	$\chi^2/df$	df	SRMR	CFI	TLI	RMSEA	AIC	BIC
<b>Unidimensional model</b>									
Goldberg and Williams (1988)	261.79	4.84	54	0.061	0.889	0.864	0.096	309.78	406.86
<b>Two-factor models</b>									
Andrich and Schoubroeck (1989)	194.82	3.67	53	0.053	0.924	0.905	0.080	244.84	345.96
Schmitz et al. (1999)	260.02	4.90	53	0.061	0.889	0.862	0.096	310.02	411.16
Politi (1994)	169.68	3.94	43	0.055	0.925	0.904	0.084	215.68	308.71
GHQ-7	21.357	1.64	13	0.0319	0.989	0.982	0.039	51.35	112.03
<b>Three-factor models</b>									
Worsley and Gribbin (1977)	159.73	3.13	51	0.050	0.942	0.925	0.071	213.73	322.94
Graetz (1991)	132.19	2.95	51	0.049	0.947	0.931	0.068	204.56	313.78
Farrell (1998)	228.57	4.48	51	0.057	0.905	0.877	0.091	282.57	391.78
Daradkeh et al. (2001)	199.08	3.90	51	0.054	0.921	0.894	0.083	253.08	362.29
Martin (1999)	142.15	3.46	41	0.053	0.935	0.913	0.077	192.15	293.28

Note. SRMR = standardized root mean square residual; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

**Validity**

The convergent validity of the GHQ-7 was assessed by calculating the correlation coefficients of the composite score of GHQ-7 with the three different constructs. As shown in Table 5, the GHQ-7 showed negative correlations with the subjective happiness scale ( $r = -.460, p < .01$ ) and satisfaction with life scale ( $r = -.529, p < .01$ ). On the other hand, a positive correlation was found between GHQ-12 and the Penn State worry questionnaire ( $r = .424, p < .01$ ). Overall, theoretically coherent results were obtained.

**Table 5 Correlations between GHQ-7, SHS, SWLS, and PSWQ-A**

	<b>GHQ-7</b>	<b>SHS</b>	<b>SWLS</b>	<b>PSWQ-A</b>
<b>GHQ-7</b>	—			
<b>SHS</b>	-.460**	—		
<b>SWLS</b>	-.529**	.491**	—	
<b>PSWQ-A</b>	.424**	-.373**	-.361**	—

\*\* $p < .01$

In addition, the correlation coefficients for the subscales of the GHQ-7 were also calculated. Table 6 shows the inter-correlation coefficients for the two subscales, anxiety/depression and social dysfunction, ranging between  $r = -.479, p < .01$  to  $r = .491, p < .01$ . In addition, the correlation of the subscales with the measures of subjective happiness, life satisfaction, and general worry was along the expected lines (Table 6).

**Table 6 Correlations between subscales of GHQ-7 and SHS, SWLS, and PSWQ-A**

	<b>Anx/Dep</b>	<b>Soc_Dysf</b>	<b>SHS</b>	<b>SWLS</b>	<b>PSWQ-A</b>
<b>Anx/Dep</b>	1				
<b>Soc_Dysf</b>	.421**	1			
<b>SHS</b>	-.358**	-.417**	1		
<b>SWLS</b>	-.412**	-.479**	.491**	1	
<b>PSWQ-A</b>	.387**	.330**	-.373**	-.361**	1

Note. Anx/Dep = Anxiety/Depression; Soc\_Dysf = Social Dysfunction

\*\* $p < .01$

**DISCUSSION**

The study examined the psychometric properties of the GHQ-12 in terms of reliability, factor structure, and validity in a sample of Indian undergraduate students. In addition, nine different models of the GHQ-12 widely cited in the literature were tested.

The reliability of the GHQ-12 was analyzed using Cronbach’s alpha as a measure of internal consistency. The obtained value indicated good reliability ( $\alpha = .883$ ), and the Cronbach's alpha is consistent with the values reported in other studies on young adults and adolescents (French & Tait, 2004; Padrón et al., 2012; Politi et al., 1994; Tait et al., 2002). However, Cronbach’s alpha for the GHQ-7 was slightly lower at  $\alpha = .784$  as the alpha value is affected by the length of the test (Tavakol & Dennick, 2011). Nevertheless, the revised alpha value indicates an acceptable level of reliability (Bland & Altman, 1997).

The EFA analysis suggested a seven items two-factor model. Social Dysfunction (items 1, 3, 4, and 8) and Anxiety/Depression (items 2, 5, and 9) were the factors that accounted for 62.14% of the variance.



## **Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

Next, the factor structure of GHQ-12 was explored by CFA, and our findings indicate that among the existing models cited in the literature, the three-factor model hypothesized by Graetz (1991) showed good fit upon examining the goodness of fit indices. The three factors were Anxiety/Depression, Social Dysfunction, and Loss of Confidence. Padrón et al. (2012), in a study on Spanish adolescents, identified three factors. Their data also showed a good fit for the model proposed by Graetz (1991). Similar results were obtained by Sánchez-López and Dresch (2008) in a study conducted on the Spanish general population and in a study of out-patients in Singapore (Gao et al., 2004).

Our study also indicated that upon comparing the one-factor model, two-factor models, and three factors models often cited in the literature, the three-factor models were more appropriate for the data as indicated by the model fit indices. In addition, the two measures of comparative fit, The Akaike information criterion (AIC) and the Bayesian information criterion (BIC), also indicated the same. However, the three-factor model proposed by Martin (1999) had lower values on these criteria than the model by Graetz (1991).

Furthermore, the fit indices of the two-factor GHQ-7 model were the most appropriate model as per the goodness-of-fit criteria proposed by Hu and Bentler (1999). The two-factor model with Social Dysfunction (four negatively worded items) and Anxiety/Depression (three positively worded items) is along the lines of the conceptualization of GHQ-12 by Andrich and Schoubroeck (1989).

Similar results were obtained by Wong and O'Driscoll (2016) in a sample of Hong Kong employees, where the result supported a seven items (GHQ-7) two-factor model consisting of Social Dysfunction (items 4, 7, and 8) and Anxiety/Depression (items 6, 9, 10, and 11). Both the factors are categorized in terms of positively and negatively worded items. In another study by Li et al. (2009) on Chinese adolescents. Where 12 items three-factor model demonstrated a good fit, but their findings also revealed that the eight items (GHQ-8) two-factor model represented the best-fit model.

The GHQ-7 also showed convergent validity indicated by correlations with the three different measures used in the study. Negative correlations were found between the subjective happiness scale and the satisfaction with life scale (Ruiz et al., 2017). Positive correlations with Penn state worry questionnaire (abbreviated). Hence, the obtained results were in the expected direction.

The findings of the present study shows that GHQ-12 is not a unidimensional construct. Multidimensionality was also shown in studies conducted by Campbell et al. (2003) and Zhong et al. (2022). However, the two-factor structure of GHQ-12 can be attributed to the division of positively and negatively worded items (Hystad & Johnsen, 2020).

### ***Finally, there were certain limitations in our study,***

- Our analysis used only the Likert method (0, 1, 2, 3) for scoring. There are multiple scoring methods used in the literature for scoring (Rey et al., 2014), one of them is the dichotomized scoring method (0, 0, 1, 1), and another one involves a slight variation (0, 1, 1, 1). Therefore, the results of our study do not apply to other scoring methods.
- Wording effects of the items on the factor structure of the GHQ-12 as identified by Molina et al. (2014) were not considered in the present study.

## Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students

- The samples for the present study were from the University of Delhi undergraduate courses. Although these students come from different regions of India, it is a possibility that they may differ from the general population of undergraduate students in India.
- The majority of the participants in the sample were females so this imbalance may affect the generalizability of the results. Although the sample size is adequate to carry out CFA, it can be further increased.
- Further studies on the Hindi and other regional language translations of GHQ-12 to improve screening in different parts of India can be carried out.

### CONCLUSION

The use of GHQ-12 as a unidimensional construct is not appropriate. The findings of the present study indicated GHQ-12 being a multidimensional measure. The seven items two-factor model showed the best fit for the data. The GHQ-7 showed adequate reliability and validity and hence can be used as a tool for screening mental health problems in the undergraduate student population in India.

### REFERENCES

- Abubakar, A., & Fischer, R. (2012). The factor structure of the 12-item General Health Questionnaire in a literate Kenyan population. *Stress and health: journal of the International Society for the Investigation of Stress*, 28(3), 248–254. <https://doi.org/10.1002/smi.1420>
- Andrich, D., & Van Schoubroeck, L. (1989). The General Health Questionnaire: A psychometric analysis using latent trait theory. *Psychological Medicine*, 19(2), 469–485. <https://doi.org/10.1017/S0033291700012502>
- Anjara, S. G., Bonetto, C., Van Bortel, T., & Brayne, C. (2020). Using the GHQ-12 to screen for mental health problems among primary care patients: psychometrics and practical considerations. *International journal of mental health systems*, 14, 62. <https://doi.org/10.1186/s13033-020-00397-0>
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *BMJ (Clinical research ed.)*, 314(7080), 572. <https://doi.org/10.1136/bmj.314.7080.572>
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS*. Mahwah, NJ: Lawrence Erlbaum.
- Campbell, A., Walker, J., & Farrell, G. (2003). Confirmatory factor analysis of the GHQ-12: can I see that again? *Australian and New Zealand Journal of Psychiatry*, 37(4), 475–483. <https://doi.org/10.1046/j.1440-1614.2003.01208.x>
- Campbell, A., & Knowles, S. (2007). A confirmatory factor analysis of the GHQ12 using a large Australian sample. *European Journal of Psychological Assessment*, 23(1), 2–8. <https://doi.org/10.1027/1015-5759.23.1.2>
- Cohen, R. J., & Swerdlik, M. E. (2005). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (6th ed.). New York: McGraw-Hill.
- Costello, A. B. & Osborne, J. (2005). "Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis," *Practical Assessment, Research, and Evaluation*, 10(7), 1-9. <https://doi.org/10.7275/jyj1-4868>

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

- Cheung, Y. B. (2002). A confirmatory factor analysis of the 12-item General Health Questionnaire among older people. *International Journal of Geriatric Psychiatry*, 17(8), 739–744. <https://doi.org/10.1002/gps.693>
- Daradkeh, T. K., Ghubash, R., & El-Rufaie, O. E. F. (2001). Reliability, validity, and factor structure of the Arabic version of the 12-item General Health Questionnaire. *Psychological Reports*, 89(1), 85-94. <http://dx.doi.org/10.2466/PRO.89.5.85-94>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of personality assessment*, 49(1), 71–75. [https://doi.org/10.1207/s15327752jpa4901\\_13](https://doi.org/10.1207/s15327752jpa4901_13)
- Endsley, P., Weobong, B., & Nadkarni, A. (2017). The psychometric properties of GHQ for detecting common mental disorder among community dwelling men in Goa, India. *Asian journal of psychiatry*, 28, 106–110. <https://doi.org/10.1016/j.ajp.2017.03.023>
- Farrell G. A. (1998). The mental health of hospital nurses in Tasmania as measured by the 12-item General Health Questionnaire. *Journal of advanced nursing*, 28(4), 707–712. <https://doi.org/10.1046/j.1365-2648.1998.00735.x>
- French, D. J., & Tait, R. J. (2004). Measurement invariance in the General Health Questionnaire-12 in young Australian adolescents. *European child & adolescent psychiatry*, 13(1), 1–7. <https://doi.org/10.1007/s00787-004-0345-7>
- Gao, F., Luo, N., Thumboo, J., Fones, C., Li, S. C., & Cheung, Y. B. (2004). Does the 12-item General Health Questionnaire contain multiple factors and do we need them? *Health and quality of life outcomes*, 2, 63. <https://doi.org/10.1186/1477-7525-2-63>
- Gaiha, S.M., Taylor Salisbury, T., Koschorke, M. (2020). Stigma associated with mental health problems among young people in India: a systematic review of magnitude, manifestations, and recommendations. *BMC Psychiatry*, 20, 538. <https://doi.org/10.1186/s12888-020-02937-x>
- Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness. Oxford U. Press.
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27(1), 191-197. <http://dx.doi.org/10.1017/S0033291796004242>
- Goldberg, D., & Williams, P. (1988). A user's guide to the General Health Questionnaire. Windsor, UK: NFER-Nelson.
- Graetz, B. (1991). Multidimensional properties of the General Health Questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, 26(3),132-138. <http://dx.doi.org/10.1007/BF00782952>
- Guan M. (2017). Measuring the effects of socioeconomic factors on mental health among migrants in urban China: a multiple indicators multiple causes model. *International journal of mental health systems*, 11, 10. <https://doi.org/10.1186/s13033-016-0118-y>
- Hankins M. (2008). The reliability of the twelve-item general health questionnaire (GHQ-12) under realistic assumptions. *BMC public health*, 8, 355. <https://doi.org/10.1186/1471-2458-8-355>
- Hopko, D. R., Stanley, M. A., Reas, D. L., Wetherell, J. L., Beck, J. G., Novy, D. M., & Averill, P. M. (2003). Assessing worry in older adults: confirmatory factor analysis of the Penn State Worry Questionnaire and psychometric properties of an abbreviated model. *Psychological assessment*, 15(2), 173–183. <https://doi.org/10.1037/1040-3590.15.2.173>

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

- Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Hystad, S. W., & Johnsen, B. H. (2020). The Dimensionality of the 12-Item General Health Questionnaire (GHQ-12): Comparisons of Factor Structures and Invariance Across Samples and Time. *Frontiers in psychology, 11*, 1300. <https://doi.org/10.3389/fpsyg.2020.01300>
- Kalliath, T. J., O'Driscoll, M. P., & Brough, P. (2004). A confirmatory factor analysis of the General Health Questionnaire - 12. *Stress and Health: Journal of the International Society for the Investigation of Stress, 20*(1), 11–20. <https://doi.org/10.1002/smi.993>
- Kashyap, G. C., & Singh, S. K. (2017). Reliability and validity of general health questionnaire (GHQ-12) for male tannery workers: a study carried out in Kanpur, India. *BMC psychiatry, 17*(1), 102. <https://doi.org/10.1186/s12888-017-1253-y>
- Kaur, M., & Kaur, R. (2019). Using the 12-item General Health Questionnaire (GHQ-12) to assess the Mental Health of Farmers of Punjab. *International Journal of Pure & Applied Bioscience, 6* (6), 905-912. <http://dx.doi.org/10.18782/2320-7051.7052>
- Kline, R. B. (1998). *Principles and Practice of Structural Equation Modeling*. New York, NY: The Guilford Press.
- Laaksonen, E., Martikainen, P., Lahelma, E., Lallukka, T., Rahkonen, O., Head, J., & Marmot, M. (2007). Socioeconomic circumstances and common mental disorders among Finnish and British public sector employees: evidence from the Helsinki Health Study and the Whitehall II Study. *International journal of epidemiology, 36*(4), 776–786. <https://doi.org/10.1093/ije/dym074>
- Li, W. H., Chung, J. O., Chui, M. M., & Chan, P. S. (2009). Factorial structure of the Chinese version of the 12-item General Health Questionnaire in adolescents. *Journal of clinical nursing, 18*(23), 3253–3261. <https://doi.org/10.1111/j.1365-2702.2009.02905.x>
- Liang, Y., Wang, L., & Yin, X. (2016). The factor structure of the 12-item general health questionnaire (GHQ-12) in young Chinese civil servants. *Health and quality of life outcomes, 14*(1), 136. <https://doi.org/10.1186/s12955-016-0539-y>
- Lyubomirsky, S., & Lepper, H. S. (1999). A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation. *Social Indicators Research, 46*, 137-155. <https://doi.org/10.1023/A:1006824100041>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods, 1*, 130-149. <https://doi.org/10.1037/1082-989X.1.2.130>
- Martin, A. J. (1999). Assessing the multidimensionality of the 12-item General Health Questionnaire. *Psychological Reports, 84*(3, pt 1), 927–935. <http://dx.doi.org/10.2466/PRO.84.3.927-935>
- Mayhew, E., Stuttard, L., & Beresford, B. (2021). An assessment of the psychometric properties of the GHQ-12 in an English population of autistic adults without learning difficulties. *Journal of Autism and Developmental Disorders, 51*(4), 1093–1106. <https://doi.org/10.1007/s10803-020-04604-2>
- McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64–82
- Molina, J. G., Rodrigo, M. F., Losilla, J. M., & Vives, J. (2014). Wording effects and the factor structure of the 12-item General Health Questionnaire (GHQ-12). *Psychological assessment, 26*(3), 1031–1037. <https://doi.org/10.1037/a0036472>

**Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

- Padrón, A., Galán, I., Durbán, M., Gandarillas, A., & Rodríguez-Artalejo, F. (2012). Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents. *Quality of life research: an international journal of quality-of-life aspects of treatment, care, and rehabilitation*, 21(7), 1291–1298. <https://doi.org/10.1007/s11136-011-0038-x>
- Philip, S., Molodynski, A., Barklie, L., Bhugra, D., & Chaturvedi, S. K. (2021). Psychological well-being and burnout amongst medical students in India: a report from a nationally accessible survey. *Middle East Current Psychiatry*, 28(1), [54]. <https://doi.org/10.1186/s43045-021-00129-1>
- Politi, P. L., Piccinelli, M., & Wilkinson, G. (1994). Reliability, validity, and factor structure of the 12-item General Health Questionnaire among young males in Italy. *Acta Psychiatrica Scandinavica*, 90(6), 432–437. <https://doi.org/10.1111/j.1600-0447.1994.tb01620.x>
- Qin, M., Vlachantoni, A., Evandrou, M., & Falkingham, J. (2018). General Health Questionnaire-12 reliability, factor structure, and external validity among older adults in India. *Indian journal of psychiatry*, 60(1), 56–59. [https://doi.org/10.4103/psychiatry.IndianJPsychiatry\\_112\\_17](https://doi.org/10.4103/psychiatry.IndianJPsychiatry_112_17)
- Rey, J. J., Abad, F. J., Barrada, J. R., Garrido, L. E., & Ponsoda, V. (2014). The impact of ambiguous response categories on the factor structure of the GHQ-12. *Psychological Assessment*, 26(3), 1021–1030. <https://doi.org/10.1037/a0036468>
- Rodrigo, M. F., Molina, J. G., Losilla, J. M., Vives, J., & Tomás, J. M. (2019). Method effects associated with negatively and positively worded items on the 12-item General Health Questionnaire (GHQ-12): results from a cross-sectional survey with a representative sample of Catalonian workers. *BMJ Open*, 9(11), e031859. <https://doi.org/10.1136/bmjopen-2019-031859>
- Ruiz, F. J., García-Beltrán, D. M., & Suárez-Falcón, J. C. (2017). General Health Questionnaire-12 validity in Colombia and factorial equivalence between clinical and nonclinical participants. *Psychiatry Research*, 256, 53–58. <https://doi.org/10.1016/j.psychres.2017.06.020>
- Sánchez-López, M., & Dresch, V. (2008). The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity, and factor structure in the Spanish population. *Psicothema*, 20(4), 839–843.
- Schmitz, N., Kruse, J., & Tress, W. (1999). Psychometric properties of the General Health Questionnaire (GHQ-12) in a German primary care sample. *Acta Psychiatrica Scandinavica*, 100(6), 462–468. <https://doi.org/10.1111/j.1600-0447.1999.tb10898.x>
- Shevlin, M., & Adamson, G. (2005). Alternative Factor Models and Factorial Invariance of the GHQ-12: A Large Sample Analysis Using Confirmatory Factor Analysis. *Psychological Assessment*, 17(2), 231–236. <http://dx.doi.org/10.1037/1040-3590.17.2.231>
- Sriram, T. G., Chandrashekar, C. R., Isaac, M. K., & Shanmugham, V. (1989). The General Health Questionnaire (GHQ): Comparison of the English version and a translated Indian version. *Social Psychiatry and Psychiatric Epidemiology: The International Journal for Research in Social and Genetic Epidemiology and Mental Health Services*, 24(6), 317–320. <https://doi.org/10.1007/BF01788035>
- Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, 42(5), 893–898. <https://doi.org/10.1016/j.paid.2006.09.017>

## **Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students**

- Sunitha, S., & Gururaj, G. (2014). Health behaviors & problems among young people in India: cause for concern & call for action. *The Indian journal of medical research*, 140(2), 185–208.
- Tait, R. J., Hulse, G. K., & Robertson, S. I. (2002). A review of the validity of the General Health Questionnaire in adolescent populations. *The Australian and New Zealand journal of psychiatry*, 36(4), 550–557. <https://doi.org/10.1046/j.1440-1614.2002.01028.x>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Tomás, J.M., Gutiérrez, M., Sancho, P., & Romero, I. (2015). Measurement invariance of the Satisfaction with Life Scale (SWLS) by gender and age in Angola. *Personality and Individual Differences*, 85, 182–186. <https://doi.org/10.1016/j.paid.2015.05.008>
- Van Droogenbroeck, F., Spruyt, B., & Keppens, G. (2018). Gender differences in mental health problems among adolescents and the role of social support: results from the Belgian health interview surveys 2008 and 2013. *BMC psychiatry*, 18(1), 6. <https://doi.org/10.1186/s12888-018-1591-4>
- Winzer, R., Lindblad, F., Sorjonen, K., & Lindberg, L. (2014). Positive versus negative mental health in emerging adulthood: a national cross-sectional survey. *BMC public health*, 14, 1238. <https://doi.org/10.1186/1471-2458-14-1238>
- Wong, K. C., & O'Driscoll, M. P. (2016). Psychometric properties of the General Health Questionnaire-12 in a sample of Hong Kong employees. *Psychology, health & medicine*, 21(8), 975–980. <https://doi.org/10.1080/13548506.2016.1140901>
- Worsley, A., & Gribbin, C. C. (1977). A factor analytic study of the twelve-item general health questionnaire. *Australian and New Zealand Journal of Psychiatry*, 11(4), 269–272. <https://doi.org/10.3109/00048677709159577>
- Zhong, X., Jin, X., Yan, L., Yang, L., Long, H., Wang, J., Wang, H., Liu, Y., Pu, J., Xie, P., & Ji, P. (2022). Reliability and Validity of General Health Questionnaire-12 in Chinese Dental Healthcare Workers During the COVID-19 Pandemic. *Frontiers in psychiatry*, 12, 792838. <https://doi.org/10.3389/fpsy.2021.792838>

### **Acknowledgement**

The author(s) appreciates all those who participated in the study and helped to facilitate the research process.

### **Conflict of Interest**

The author(s) declared no conflict of interest.

**How to cite this article:** Rathore, D., Rana, S. & Chadha, N. K. (2022). Psychometric Assessment of the GHQ-12 (General Health Questionnaire-12 items): Evaluating the Reliability, Validity, and Comparison of Factor Structure in Indian Undergraduate Students. *International Journal of Indian Psychology*, 10(3), 1355-1368. DIP:18.01.145.20221003, DOI:10.25215/1003.145