# Factorial Validity and Reliability of 12 items General Health Questionnaire in a Bhutanese Population

*Tshoki Zangmo*[*]

**Abstract**

The aim of this study is to test the factorial structure and the internal consistency of the 12-items General Health Questionnaire. A sample of 6861 Bhutanese completed the GHQ-12. Internal consistency was assessed by Cronbach's alpha coefficient. The factorial structure was extracted with an exploratory factorial analysis (EFA). The EFA run on the data yield to a one- factor structure without rotation and two factor structures after rotation. Cronbach's alpha showed a very good internal consistency of the scale (α= 0.88). Cluster analysis resulted in two clusters. Overall, the findings support that the GHQ-12 is a reliable and valid instrument for measuring minor psychological distress in a Bhutanese sample.

## Background

The General Health Questionnaire (GHQ) was developed in England as a screening instrument to identify psychological distress in primary care settings (Goldberg & Blackwell, 1970). It was originally designed as a 60-item instrument but several shortened versions are currently available, including the GHQ-30, the GHQ-28, the GHQ-20 and the GHQ-12. The shortest version of the questionnaire (GHQ-12) has been extensively validated and used in a number of countries and in different languages (Politi et al, 1994; Chan, 1993). Since this version is brief, simple and easy to complete, and its application as a screening tool in research settings is well documented, it was used to examine and test its

[*] Information and Media Officer, Centre for Bhutan Studies and GNH Research. Correspondence: tshokey@gmail.com

psychometric properties and factor structure (i.e. one, two or three factors) in a Bhutanese sample.

**Objective**

The general objective of this study was to assess the underlying dimensional structure of the GHQ-12 item and to evaluate its validity and reliability as an effective measure of distress as an overall score. In particular, the study attempt to explore the factorial structure through factor analysis and apply cluster analysis to test the presence of any subgroups.

**Data**

The data for the analysis was taken from the Second Gross National Happiness survey 2010 carried out by the Centre for Bhutan Studies, an autonomous research institute in Bhutan. The primary purpose for the survey was to collect information on the living conditions of the Bhutanese in general. The survey was conducted by personal interview with a sample size of 7,142 representative of the population by region and districts.

However, for the current study the sample size reduced to 6,861 due to the missing values in the questions of interest. A number of variables of the GHQ-12 contained missing values at random and further, don't knows were also decided to be re-categorized as missing since they did not provide any additional information. In terms of missing values at random which made (n=182,) 2.5% of the sample proportion, no demographics differences were observed.

**Instrument**

GHQ-12 consists of 12 questions that are rated on a four-pointer scale and has three types of response codes. Some are coded; 1) less than usual 2) no more than usual 3) rather more than usual 4) much more than usual while others followed; 1) much more than usual 2) no more than usual 3) rather more than usual 4) Not at all and few are coded as; 1) much less than usual 2) less than usual 3) same as usual 4) more than usual. All the scores are recoded in to 0-3 Likert scale from most negative to most positive value for all the 12 items. Though there are six positively and six negatively worded items, it must be noted that a higher value always indicates a positive response in that particular item and a lower value indicates otherwise.

## Descriptive Statistics

Table below presents the mean scores (M), standard deviation (SD), and minimum and maximum value.

Table 1a. Mean scores, standard deviation and minimum and maximum

| GHQ | Sample | M | SD | Min | Max |
|---|---|---|---|---|---|
| GHQ1 | 6861 | 0.8877715 | 0.6570523 | 0 | 3 |
| GHQ2 | 6861 | 0.7423116 | 0.9016026 | 0 | 3 |
| GHQ3 | 6861 | 0.8580382 | 0.5835112 | 0 | 3 |
| GHQ4 | 6861 | 0.8775689 | 0.6040628 | 0 | 3 |
| GHQ5 | 6861 | 0.8016324 | 0.8770662 | 0 | 3 |
| GHQ6 | 6861 | 0.8924355 | 0.8553227 | 0 | 3 |
| GHQ7 | 6861 | 0.9070106 | 0.6199634 | 0 | 3 |
| GHQ8 | 6861 | 0.941991 | 0.6166911 | 0 | 3 |
| GHQ9 | 6861 | 0.7210319 | 0.8366648 | 0 | 3 |
| GHQ10 | 6861 | 0.6926104 | 0.8052955 | 0 | 3 |
| GHQ11 | 6861 | 0.6124472 | 0.8426061 | 0 | 3 |
| GHQ12 | 6861 | 0.9255211 | 0.5687362 | 0 | 3 |

## Factor analysis

### *Testing assumptions*

Prior to the extraction of factors, several tests were carried out to assess the suitability of the data for factor analysis. The adequacy of the correlation matrix of the GHQ-12 item was checked and it was observed that there was a strong and statistically significant correlation [1] between the variables (±0.2 to ±0.7).

[1] Table i, Appendix I

The data also met the Kaiser-Meyer-Olkin (KMO) criteria for sampling adequacy as 0.91 which is greater than the suggested minimum of 0.6. The Bartlett's test for sphericity was significant ($x^2 = 35760, p < 0.001$). Taken together, these tests suggest that the data meets the minimum standards for factor analysis.

Table 1b. Kaiser-Meyer-Olkin (KMO) criteria for sampling adequacy

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .910 |
| Bartlett's Test of Sphericity | Approx. Chi Square | 35760.438 |
| | Df | 66 |
| | Sig. | 0.000 |

**Factor extraction**

In terms of extraction methods, principle component analysis was adopted as it is most commonly suggested in establishing preliminary solutions in exploratory factor analysis[2]. PCA is also recommended when no prior theory or model exists[3].

Although, initial unrotated[4] solution resulted in a single factor solution, the factor pattern was not clear and some variables seem to cross load into a second factor. So, in order to get more interpretable and simplified solution; factors scores were rotated using orthogonal (varimax) rotation. The data was experimented with other methods such as common factor analysis and other rotation methods which also resulted in two factor solutions. However, PCA orthogonal varimax rotation was opted as it produces uncorrelated factors which seem sensible when testing for any underlying diverse structures in GHQ-12 item.

[2] Pett et al. (2003)
[3] Gorsuch RL. Factor Analysis. Hillsdale, NJ: Erlbaum; 1990.
[4] Table ii, Appendix I

The Kaiser criterion [5] (eigenvalue>1 rule) along with examination of the scree[6] plot clearly suggested two factors to be extracted. There weren't any strong cross loadings between the factors. The two factors explained 59% of the total variance. Table 2 shows the eigenvalues, percentage of explained variance associated with each factor and their loadings on GHQ-12 items.

Table 2. Varimax orthogonal rotated solution of the factors (n=6861)

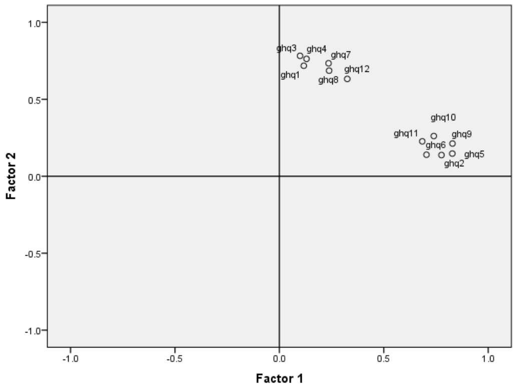
| | Variable | Factor 1 | Factor 2 |
|---|---|---|---|
| | Have you… | | |
| GHQ1 | been able to concentrate? | 0.117 | 0.718 |
| GHQ2 | lost much sleep over worry? | 0.776 | 0.138 |
| GHQ3 | felt that you were playing a useful part ? | 0.099 | 0.782 |
| GHQ4 | felt capable of making decisions? | 0.13 | 0.763 |
| GHQ5 | felt constantly under strain? | 0.828 | 0.147 |
| GHQ6 | felt that you couldn't overcome difficulties? | 0.704 | 0.14 |
| GHQ7 | been able to enjoy normal day to day activities? | 0.235 | 0.734 |
| GHQ8 | been able to face upto your problems? | 0.238 | 0.686 |
| GHQ9 | been feeling unhappy and depressed? | 0.829 | 0.212 |
| GHQ10 | been losing confidence? | 0.739 | 0.261 |
| GHQ11 | been thinking of yourself as a worthless person? | 0.684 | 0.227 |
| GHQ12 | been feeling reasonably happy, all things considered? | 0.325 | 0.632 |
| | Eigen values | 5.21 | 1.878 |
| | % of variance explained by each factor | 43.41% | 15.65% |
| | % of variance explained by two factors: 59.063% | | |

The cutoff to define the item as representing the factor was chosen with the factor loading > 0.50. GHQ2, GHQ5, GHQ6,

[5] Table iii, Appendix I
[6] Graph 1, Appendix I

GHQ9, GHQ10 and GHQ11 loaded on factor 1 while the rest loaded on factor 2. The negatively worded items form the first factor and positively worded items form the second.

The grouping of the variables into two factors is also obvious in the graph 2.
Graph 2: Factor plot in rotated space




Further analysis of the factorial structure by various demographics[7] also identified the same two factor models.

***Goodness of fit***

As a measure of goodness of fit for the two factor solutions, the KMO Measure of Sampling Adequacy for each individual variable was investigated. As observed in the diagonal of the matrix (see Table 3, Appendix I), the correlation coefficients are well above the acceptable level of 0.6. This suggests that the matrix is suitable for factoring.

In addition, each variable has a reasonable amount in common with the other variables. The table of communalities shows communalities for all variables to be above the desired minimum of 0.5 and hence, do not suggest removal of any items.

[7] Factorial structure of the GHQ-12 is observed by sex and region in Table iv, Appendix I

Table 4. Communality

| | Initial | Extraction |
|---|---|---|
| GHQ1 | 1.000 | .529 |
| GHQ2 | 1.000 | .621 |
| GHQ3 | 1.000 | .622 |
| GHQ4 | 1.000 | .599 |
| GHQ5 | 1.000 | .707 |
| GHQ6 | 1.000 | .516 |
| GHQ7 | 1.000 | .594 |
| GHQ8 | 1.000 | .528 |
| GHQ9 | 1.000 | .732 |
| GHQ10 | 1.000 | .615 |
| GHQ11 | 1.000 | .520 |
| GHQ12 | 1.000 | .506 |

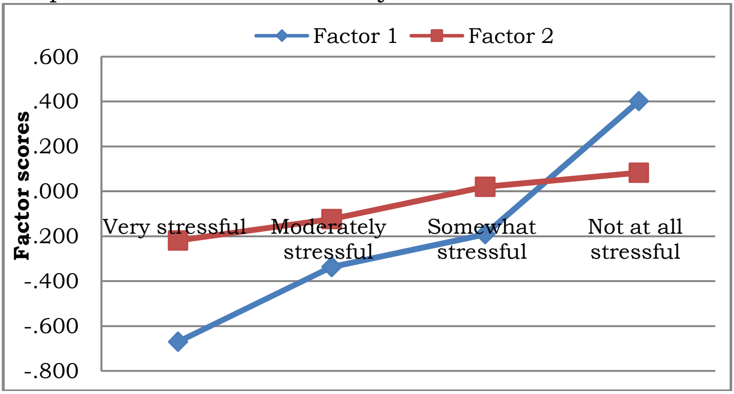
Finally, the model fit is tested investigating the residual matrix. Only 4 out of the 66 residuals (6%) are greater than 0.05 in absolute value, suggesting that this factor model is a good overall fit to the data as shown in Table 5, Appendix I.

***Factor validity***

As previously observed the variables that are similar in nature load together. For instance, all the negatively worded items load on factor 1. The negatively worded items seem to be interpreting some sort of mental distress in respondents. In the same way, positively worded items load on factor 2. The results ensure face validity of the suggested factors.

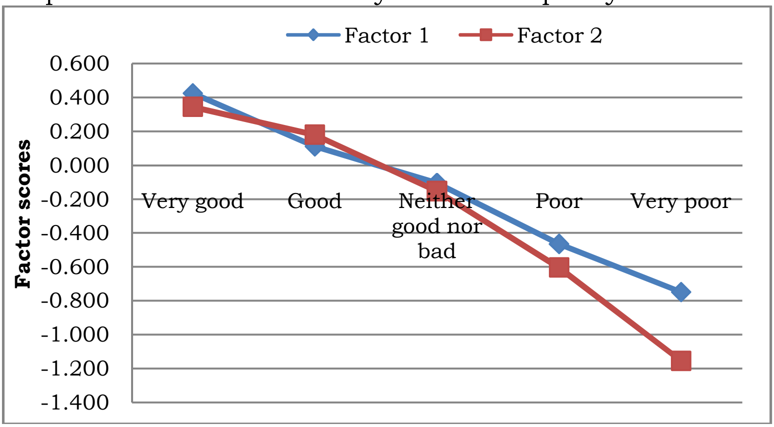
In order to test for construct validity, mean factor scores are compared amongst outcome variables such as level of stress and life quality (Graph 2 and 3). Both the outcome variables are categorical in nature. Here again, as assumed the factor scores increases (higher mental health) as the level of stress in the categories decreases.

Graph 2. Mean factor scores by level of stress




Similar results are observed with level of life quality. As expected, respondents who have "very good" life quality also have higher factors scores indicating higher mental health. The mean factors scores decreases as the categories of life quality decreases to "very poor".

Graph 3. Mean factor scores by level of life quality




Analysis of Variance (ANOVA) was conducted to determine whether various demographic groups differed in their judgement of the two factors. However, no significant differences were observed in any of the demographic [8] variables.

***Factor reliability***

In terms of the 12 items of GHQ, it was previously observed to show homogeneity with its mean inter-item correlation above 0.38. With respect to the two-factor model, the homogeneities (mean inter-item correlation) were 0.56 (r = 0.43-0.84, p < 0.01) and .47 (r = 0.39-0.78, p < 0.01) for first and second factors. Further, inter-correlations between the variables and its corresponding factors as shown in the tables below demonstrate reliability of the factors.

Table 6a. Correlations between the six variables and factor 1

| | Factor 1 score | GHQ 2 | GHQ 5 | GHQ 6 | GHQ 9 | GHQ 10 | GHQ 11 |
|---|---|---|---|---|---|---|---|
| Factor 1 score | 1 | | | | | | |
| GHQ2 | .776** | 1 | | | | | |
| GHQ5 | .828** | .661** | 1 | | | | |
| GHQ6 | .704** | .462** | .531** | 1 | | | |
| GHQ9 | .829** | .627** | .704** | .511** | 1 | | |
| GHQ10 | .739** | .485** | .533** | .490** | .610** | 1 | |
| GHQ11 | .684** | .436** | .463** | .425** | .528** | .612** | 1 |

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6b. Correlations between the six variables and factor 2

| | Factor 2 score | GHQ1 | GHQ3 | GHQ4 | GHQ7 | GHQ8 | GHQ12 |
|---|---|---|---|---|---|---|---|
| Factor 2 score | 1 | | | | | | |
| GHQ1 | .718** | 1 | | | | | |
| GHQ3 | .782** | .514** | 1 | | | | |
| GHQ4 | .763** | .456** | .552** | 1 | | | |
| GHQ7 | .734** | .468** | .492** | .480** | 1 | | |
| GHQ8 | .686** | .386** | .445** | .495** | .495** | 1 | |
| GHQ12 | .632** | .395** | .424** | .408** | .527** | .460** | 1 |

**. Correlation is significant at the 0.01 level (2-tailed).

[8] Table iv in the Appendix indicates that there are no significant difference in the factor scores when observed by both gender and region

***Interpretation of factors***

The item loadings on the first factor (e.g., lost sleep over worry, constantly under strain, unhappy or depressed) seem to indicate the psychological issues faced by individuals and might represent the construct psychological functioning. While those loadings on the second factor (e.g., able to concentrate, play useful part in things, capable of making decisions) seem to be a combination of variables representing the ability of an individual to perform normal social and emotional functions in life and may be expressed as social functioning.

Accordingly, two variables named psychological and social functioning are generated based on the mean of the six variables, which loads on the respective factors. The GHQ score is sum of all the 12 items ranging from 0 to 36. Higher score indicates a positive value.

Significant correlation coefficient between two factors, GHQ score and psychological and social functioning variables supported the convergent validity of the factors.

Table 7. The correlations between the factor scores of GHQ-12 and its factors

| | Factor 1 | Factor 2 | Psychological functioning | Social functioning | GHQ score |
|---|---|---|---|---|---|
| Factor 1 | 1 | | | | |
| Factor 2 | .000 | 1 | | | |
| Psychological functioning | .970** | .237** | 1 | | |
| Social functioning | .254** | .966** | .473** | 1 | |
| GHQ score | 0.788 | 0.615 | 0.911 | 0.794 | 1 |

**. Correlation is significant at the 0.01 level (2-tailed).

It seems there are significant differences between mean scores for the positively and negatively worded items. It is also observed that GHQ score is highly correlated with both the factors. The findings here are similar to those previously reported in other studies (Doi & Minowa, 2003; Montazeri et al., 2003).

When the correlation between the GHQ-12 and quality of life scores was investigated, as expected a significant positive correlation emerged indicating that those who were more distressed showed lower levels of global quality of life. Likewise, significant negative relationship resulted when correlated with stress level.

Table 8a. Spearman correlation

| n=6814 | Life quality | GHQ score | Stress level |
|---|---|---|---|
| Life quality | 1 | | |
| GHQ score | 0.305 | 1 | |
| Stress level | -0.1441 | -0.3526 | 1 |

P value=0.000

***Internal consistency***

The reliability of the measures was examined in relation to the instrument's internal consistency (Cronbach's alpha coefficients) and homogeneity (mean inter-item correlations). Cronbach's alpha coefficients of 0.70 or higher and mean inter-item correlations in the 0.20 to 0.60 range were deemed to indicate good reliability (Nunnally et. al, 1967). The alpha for the whole sample was found to be 0.87 and was the same for both males and females indicating satisfactory results. As previously observed, the 12 items of the GHQ showed homogeneity.

Table 8b. Interim covariance and Cronbach's α

| | National | Male | Female |
|---|---|---|---|
| Average interim covariance | 0.206 | 0.171 | 0.231 |
| Scale of reliability coefficient | 0.878 | 0.861 | 0.887 |

## Cluster Analysis

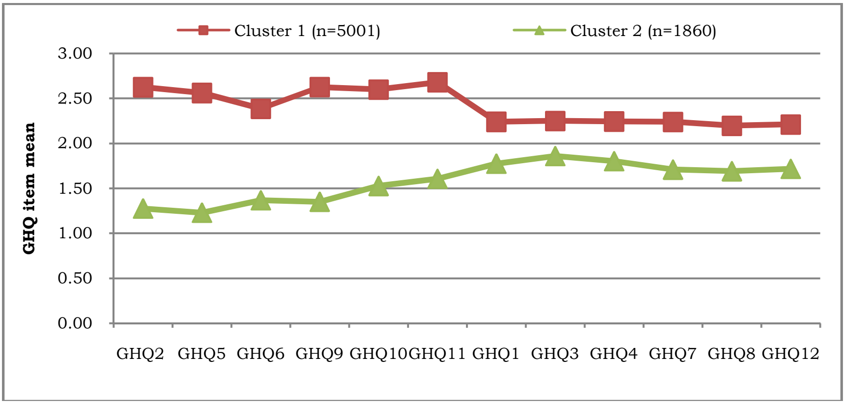
### *Clustering method*

Next, a hierarchical agglomerative clustering procedure is run to explore the possible number of clusters in the data. Specifically, average linkage with a squared Euclidean distance measure is used for the analysis[9]. Inspection of the agglomeration schedule for changes in the agglomerative coefficient and a visual inspection of the tree dendrogram for key cut points suggest two cluster solutions.

A K-means clustering procedure is next used to classify the data. Using the cluster centroids from the hierarchical clustering procedure, a two-cluster solution is specified as part of a K-means cluster analysis of the sample. Examination of the *F*-test and the mean squared results indicates that all variables in the procedure were statistically significant[10] (P value< 0.01).

### *Cluster analysis by GHQ-12*

Graph 4. Mean items scores by clusters




---

9 Other clustering methods such as Wards method, centroid clustering results in 2 cluster solutions.

10 Table v, Appendix I

One sample t test was found to be significant (P value<0.01) between the variable groups.

There seem to be a significant difference between the ways two clusters respond to GHQ-12 items. It clearly reveals two groups; high scorers corresponding to cluster 1 (5001 respondents) and low scorers corresponding to cluster 2 (1860 respondents). ANOVA revealed that the item variances[11] were significantly higher across all items for high scorers in comparison with the low scorers.

It must also be noted that the GHQ-12 items have been rearranged so that first six items are negatively phrased (factor 1) while the last six are positively phrased (factor 2). Variances were particularly high for cluster 1 for the negatively phrased items, suggesting a three-way interaction between individual item variance, group item valence and clusters as shown in graph 4. This may suggests that negatively phrased items may be affecting the responses.

For instance, high scorers (i.e. people with high self-reported mental health) seem to score much higher for negatively phrased items as compared to positively phrased items. Similarly, low scorers (i.e. people with high self-reported mental health) seem to score much lower for negatively phrased items.

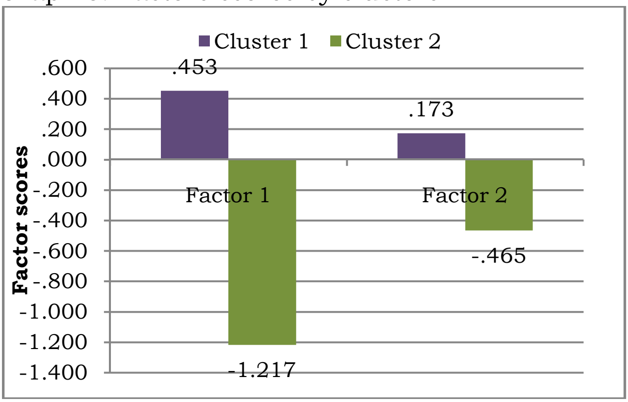
[11] Table vi, Appendix I

Table 9. Means, standard deviations and variances for GHQ-12 by clusters

| Group | n | Mean | SD | 95% CI Lower | Upper | Min | Max |
|---|---|---|---|---|---|---|---|
| Negatively phrased items | | | | | | | |
| Cluster 1 | 5001 | 2.578 | 0.381 | 2.568 | 2.589 | 1.5 | 3.0 |
| Cluster 2 | 1862 | 1.391 | 0.484 | 1.369 | 1.413 | 0.00 | 2.33 |
| Positively phrased items | | | | | | | |
| Cluster 1 | 5001 | 2.228 | 0.359 | 2.218 | 2.238 | 1.00 | 3.00 |
| Cluster 2 | 1862 | 1.756 | 0.500 | 1.734 | 1.779 | 0.00 | 3.00 |
| GHQ score | | | | | | | |
| Cluster 1 | 5001 | 28.838 | 3.384 | 28.744 | 28.932 | 21.0 | 36.0 |
| Cluster 2 | 1862 | 18.882 | 4.674 | 18.670 | 19.095 | 0.00 | 28.0 |

Consistent with the previous findings, the two clusters have significant differences in terms of factor scores as depicted in the graph 5.

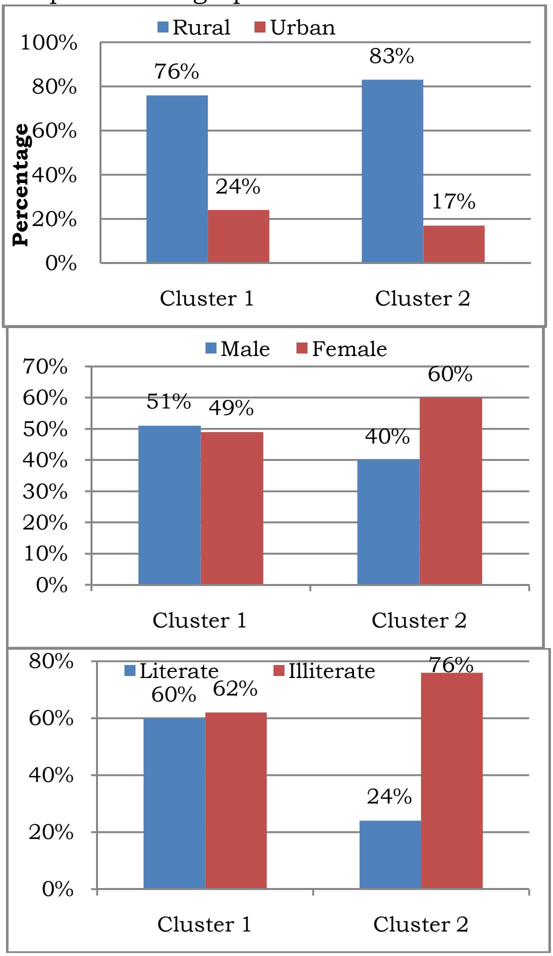
Graph 5. Factors scores by clusters




### ***Cluster profiling***

The graphs below outlines demographics cluster membership based on K-means clustering solution.

Graph 6: Demographic constituents of the two clusters




As previously observed, respondents in cluster 1 has significantly higher psychological wellbeing (lower levels of psychiatric distress). There were almost equal proportions of men and women in cluster 1.

Cluster 2 represents the low scores and so have lower psychological wellbeing (high levels of psychological distress).

With regards to cluster 2, it was observed that a high proportion of this cluster was women (60%) and respondents who are illiterate (76%).

In both the clusters most of the respondents seem to be from rural areas. This is perhaps due to the high proportion (77.8%) of rural respondents in the sample itself.

**Discussion**

The aim of this study was to explore the factorial structure underlying GHQ-12 and to apply a cluster analysis to GHQ-12 data to determine whether any subgroups of respondents could be identified.

The largest number of studies (Graetz, 1991 ; Politi et al. 1994; Martin, 1999; Campbell et al. 2003; Shevlin & Adamson, 2005) report similar factor structures to ours, namely one factor on which all of the positively worded items have high loadings, and the other on which the negatively worded items have high loadings.

Based on the validity, reliability and a good overall fit of the two factors model, it does not seem efficient to use GHQ-12 as uni-dimensional construct. However, considering the high inter-factor correlations also found in previous studies (Cheung, 2002; Kalliath et al. 2004; Werneke et al, 2000) suggest the use of GHQ-12 as a uni-dimensional score. Additionally, the Cronbach's alpha showed a very good internal consistency.

Studies have reported one (Lewis, 1992), two (Politi et al. 1994; Kalliath et al. 2004) and three factor solutions (Cheung, 2002; Campbell et al. 2003; Shevlin & Adamson, 2005). Some of this apparent inconsistency may be methodological in origin, including differences in setting, sample size and composition, weighting of item scores and methods of analysis. Exploratory Factor Analysis is prone to divergent findingsresulting from a neglect of method effects, specifically a response bias on the items expressing negative mood states. It might be due to the ambiguous response options for indicating the absence of negative mood states. To test the presence of substantial response bias on the negatively phrased items, further analysis needs to be carried out.

The results of the cluster analysis identified two groups of respondents, high and low scorers, both with different response patterns to the GHQ-12, but particularly the low scorers who tended to score low on negatively worded. This again indicates that the presence of a response bias lies at the heart of explaining the various factor models proposed for the GHQ-12, the uni-dimensional model accounting for differences in variance might perhaps provide the best fit.

## Conclusion

The factor structure of the GHQ-12 in a Bhutanese population is given by a variant of a two-factor solution, corresponding to positively worded and negatively worded items. However, the present study provides support for the view that valid measures of mental health should include items that assess both thereby using GHQ-12 on a uni-dimensional scale. While a two-dimensional model may not offer much of an advantage over a one- factor model when screening for psychiatric disorders, the former approach may come into its own in studying the determinants and consequences of well-being.

Hence, future studies should test the factor invariance of the two-factor model of the GHQ-12. For instance, this test could consider demographic (e.g., gender, age) and sociocultural (e.g., ethnic group, dialects) variables. It would also be important to examine the criterion related validity of the GHQ-12, considering some relevant psychiatric symptoms (e.g., suicidal ideation, negative affects) and indicators of work-related stress (e.g., fatigue, burnout). Finally, it would be recommended to establish the sensitivity and specificity of this measure.

**Limitations**

Although an important strength of the study was the nationally representative large sample size, it must be understood that a large sample size may also cause hindrance to significance tests. For instance, in a chi-square test, even a small difference might lead to significance and likewise in correlation coefficients, a large sample size would make it easy to achieve significance. Hence, it is recommended that an analysis on a subset of sample must be carried out in future to achieve further validity.

Another weakness lies in the GHQ-12 itself. All self-report questionnaires are prone to method variance, namely the tendency for people to respond the same way to similarly worded items. This may contribute to the aggregation of responses to positively and negatively worded items. This would require testing whether the 'positive' items on the GHQ-12 do indeed correlate with items of positive mental health or positive emotionality on other personality scales.

With regard to data values missing at random, demographics differences were only analysed. However, it is important to consider if the missing values are completely missing at random or whether the probability that a certain item missing has any relationship with other related variables. For example, people who are very depressed might be less inclined to report the GHQ-12 items, thereby leading to biasness in the estimates. It is recommended that in future such missing value analysis must be incorporated.

**Reference**

Bun Cheung, Y. (2002). A confirmatory factor analysis of the 12-item General Health Questionnaire among older people. *International Journal of Geriatric Psychiatry*, *17*(8), 739-744.

Chan, D. W. (1993). The Chinese General Health Questionnaire in a psychiatric setting: The development of the Chinese scaled version. *Social Psychiatry and Psychiatric Epidemiology*, *28*(3), 124-129.

Doi, Y., & Minowa, M. (2003). Gender differences in excessive daytime sleepiness among Japanese workers. *Social Science & Medicine*, *56*(4), 883-894.

Goldberg, D. P., & Blackwell, B. (1970). Psychiatric illness in general practice: a detailed study using a new method of case identification. *British Medical Journal*, *2*(5707), 439.

Gorsuch, R. L. (1990). Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*, *25*(1), 33-39.

Graetz, B. (1991). Multidimensional properties of the general health questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, *26*(3), 132-138.

Kalliath, T. J., O'Driscoll, M. P., & Brough, P. (2004). A confirmatory factor analysis of the General Health Questionnaire-12. *Stress and Health*, *20*(1), 11-20.

Lewis, G. (1992). Dimensions of neurosis. *Psychological Medicine-London-*, *22*, 1011-1011.

Montazeri, A., Harirchi, A. M., Shariati, M., Garmaroudi, G., Ebadi, M., & Fateh, A. (2003). The 12-item General Health Questionnaire (GHQ-12): Translation and validation study of the Iranian version. *Health and Quality of Life Outcomes*, 1(1), 66.

Nunnally, J. C., Bernstein, I. H., & Berge, J. M. T. (1967). *Psychometric theory* (Vol. 226). New York: McGraw-Hill.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research.* Sage.

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.

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.

Werneke, U., Goldberg, D. P., Yalcin, I., & Üstün, B. T. (2000). The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, *30*(04), 823-829.

**Appendix I**

Table 3: Anti-image correlation matrix

| | GHQ1 | GHQ2 | GHQ 3 | GHQ 4 | GHQ 5 | GHQ 6 | GHQ 7 | GHQ 8 | GHQ 9 | GHQ 10 | GHQ 11 | GHQ 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GHQ1 | .915[a] | | | | | | | | | | | |
| GHQ2 | | .911[a] | | | | | | | | | | |
| GHQ3 | | | .888[a] | | | | | | | | | |
| GHQ4 | | | | .897[a] | | | | | | | | |
| GHQ5 | | | | | .879[a] | | | | | | | |
| GHQ6 | | | | | | .947[a] | | | | | | |
| GHQ7 | | | | | | | .918[a] | | | | | |
| GHQ8 | | | | | | | | .929[a] | | | | |
| GHQ9 | | | | | | | | | .901[a] | | | |
| GHQ10 | | | | | | | | | | .906[a] | | |
| GHQ11 | | | | | | | | | | | .911[a] | |
| GHQ12 | | | | | | | | | | | | .934[a] |

Table 5. Residual analysis

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GHQ1 | | | | | | | | | | | |
| GHQ2 | 0.031 | | | | | | | | | | |
| GHQ3 | -0.022 | 0.023 | | | | | | | | | |
| GHQ4 | -0.045 | 0.014 | -0.013 | | | | | | | | |
| GHQ5 | 0.021 | -0.026 | 0.01 | 0.011 | | | | | | | |
| GHQ6 | 0.003 | -0.093 | 0.011 | 0.006 | -0.062 | | | | | | |
| GHQ7 | -0.039 | 0.018 | -0.033 | -0.039 | 0.012 | -0.001 | | | | | |
| GHQ8 | -0.059 | -0.002 | -0.037 | -0.021 | 0.007 | 0.012 | -0.023 | | | | |
| GHQ9 | 0.001 | -0.046 | -0.001 | -0.01 | -0.015 | -0.072 | 0.005 | -0.003 | | | |
| GHQ10 | -0.022 | -0.086 | -0.015 | -0.006 | -0.074 | -0.036 | -0.027 | -0.02 | -0.028 | | |
| GHQ11 | -0.032 | -0.092 | -0.025 | -0.01 | -0.093 | -0.054 | -0.036 | -0.015 | -0.051 | 0.043 | |
| GHQ12 | -0.036 | -0.001 | -0.027 | -0.035 | 0.005 | -0.013 | 0.001 | -0.013 | 0.007 | -0.02 | -0.012 |

Table i. Correlation matrix of the GHQ-12 items

| | GHQ 1 | GHQ2 | GHQ 3 | GHQ 4 | GHQ5 | GHQ6 | GHQ7 | GHQ8 | GHQ9 | GHQ 10 | GHQ 11 | GHQ 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GHQ1 | 1 | | | | | | | | | | | |
| GHQ2 | 0.227 | | | | | | | | | | | |
| GHQ3 | 0.514 | 0.216 | 1 | | | | | | | | | |
| GHQ4 | 0.456 | 0.216 | 0.552 | 1 | | | | | | | | |
| GHQ5 | 0.231 | 0.661 | 0.212 | 0.232 | 1 | | | | | | | |
| GHQ6 | 0.19 | 0.462 | 0.207 | 0.212 | 0.531 | 1 | | | | | | |
| GHQ7 | 0.468 | 0.295 | 0.492 | 0.48 | 0.308 | 0.261 | 1 | | | | | |
| GHQ8 | 0.386 | 0.253 | 0.445 | 0.495 | 0.292 | 0.278 | 0.495 | 1 | | | | |
| GHQ9 | 0.259 | 0.627 | 0.258 | 0.256 | 0.704 | 0.511 | 0.357 | 0.332 | 1 | | | |
| GHQ10 | 0.268 | 0.485 | 0.283 | 0.317 | 0.533 | 0.49 | 0.333 | 0.332 | 0.61 | 1 | | |
| GHQ11 | 0.224 | 0.436 | 0.236 | 0.28 | 0.463 | 0.425 | 0.286 | 0.312 | 0.528 | 0.612 | 1 | |
| GHQ12 | 0.395 | 0.312 | 0.424 | 0.408 | 0.348 | 0.277 | 0.527 | 0.46 | 0.403 | 0.358 | 0.349 | 1 |

Table ii. Unrotated factor solution

| | Factor 1 |
|---|---|
| GHQ1 | .564 |
| GHQ2 | .672 |
| GHQ3 | .593 |
| GHQ4 | .603 |
| GHQ5 | .717 |
| GHQ6 | .620 |
| GHQ7 | .663 |
| GHQ8 | .633 |
| GHQ9 | .761 |
| GHQ10 | .727 |
| GHQ11 | .663 |
| GHQ12 | .663 |

Table iii. Eigen values of the corresponding factors

| | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|
| Component | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| Factor 1 | 5.210 | 43.414 | 43.414 | 5.210 | 43.414 | 43.414 |
| Factor 2 | 1.878 | 15.649 | 59.063 | 1.878 | 15.649 | 59.063 |
| Factor 3 | .713 | 5.943 | 65.007 | | | |
| Factor 4 | .663 | 5.525 | 70.532 | | | |
| Factor 5 | .605 | 5.039 | 75.571 | | | |
| Factor 6 | .547 | 4.561 | 80.132 | | | |
| Factor 7 | .492 | 4.098 | 84.230 | | | |

| | | | |
|---|---|---|---|
| Factor 8 | .453 | 3.776 | 88.006 |
| Factor 9 | .427 | 3.557 | 91.564 |
| Factor 10 | .393 | 3.276 | 94.839 |
| Factor 11 | .343 | 2.858 | 97.697 |
| Factor 12 | .276 | 2.303 | 100.000 |

Extraction Method: Principal Component Analysis.

Graph i. Scree plot

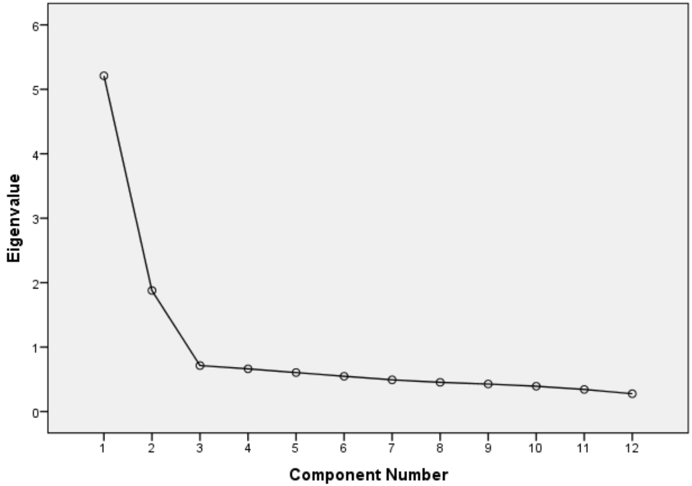


Table iv. Varimax orthogonal rotated solution of the factors
By gender

| | Male (n=3305) | | Female (n=3556) | |
|---|---|---|---|---|
| | Facto1 | Factor 2 | Factor 1 | Factor 2 |
| GHQ1 | .125 | .722 | .125 | .722 |
| GHQ2 | .766 | .123 | .766 | .123 |
| GHQ3 | .073 | .782 | .073 | .782 |
| GHQ4 | .112 | .766 | .112 | .766 |
| GHQ5 | .804 | .119 | .804 | .119 |
| GHQ6 | .658 | .106 | .658 | .106 |
| GHQ7 | .199 | .738 | .199 | .738 |
| GHQ8 | .213 | .699 | .213 | .699 |
| GHQ9 | .810 | .192 | .810 | .192 |
| GHQ10 | .717 | .245 | .717 | .245 |
| GHQ11 | .670 | .208 | .670 | .208 |
| GHQ12 | .288 | .629 | .288 | .629 |

By region

| | Rural | | Urban | |
|---|---|---|---|---|
| | Facto1 | Factor 2 | Factor 1 | Factor 2 |
| GHQ1 | .124 | .725 | .085 | .681 |
| GHQ2 | .776 | .147 | .772 | .089 |
| GHQ3 | .103 | .793 | .080 | .735 |
| GHQ4 | .141 | .772 | .083 | .722 |
| GHQ5 | .827 | .163 | .827 | .072 |
| GHQ6 | .712 | .142 | .667 | .121 |
| GHQ7 | .235 | .743 | .232 | .691 |
| GHQ8 | .260 | .690 | .122 | .661 |
| GHQ9 | .834 | .220 | .807 | .178 |
| GHQ10 | .745 | .270 | .713 | .216 |
| GHQ11 | .695 | .228 | .628 | .203 |
| GHQ12 | .332 | .638 | .295 | .605 |

Table v. Total variance explained by corresponding factors

ANOVA

| | Cluster Mean Square | df | Error Mean Square | df | F | Sig. |
|---|---|---|---|---|---|---|
| GHQ1 | 295.308 | 1 | .389 | 6859 | 759.680 | .000 |
| GHQ2 | 2470.941 | 1 | .453 | 6859 | 5457.533 | 0.000 |
| GHQ3 | 207.229 | 1 | .310 | 6859 | 667.785 | .000 |
| GHQ4 | 265.288 | 1 | .326 | 6859 | 813.099 | .000 |
| GHQ5 | 2405.668 | 1 | .419 | 6859 | 5746.584 | 0.000 |
| GHQ6 | 1404.817 | 1 | .527 | 6859 | 2666.346 | 0.000 |
| GHQ7 | 382.327 | 1 | .329 | 6859 | 1163.255 | .000 |
| GHQ8 | 348.017 | 1 | .330 | 6859 | 1055.798 | .000 |
| GHQ9 | 2207.242 | 1 | .378 | 6859 | 5834.512 | 0.000 |
| GHQ10 | 1563.105 | 1 | .421 | 6859 | 3715.447 | 0.000 |
| GHQ11 | 1565.475 | 1 | .482 | 6859 | 3248.873 | 0.000 |
| GHQ12 | 330.643 | 1 | .275 | 6859 | 1201.016 | .000 |

Table vi. Cluster scores

| | | Mean scores | |
|---|---|---|---|
| Measures | Mean (n=6861) | Cluster 1 (n=5001) | Cluster 2 (n=1860) |
| GHQ1 | 2.21 | 2.24 | 1.77 |
| GHQ2 | 2.25 | 2.62 | 1.27 |
| GHQ3 | 2.14 | 2.25 | 1.86 |
| GHQ4 | 2.12 | 2.24 | 1.80 |
| GHQ5 | 2.19 | 2.56 | 1.23 |
| GHQ6 | 2.1 | 2.38 | 1.37 |
| GHQ7 | 2.09 | 2.24 | 1.71 |
| GHQ8 | 2.05 | 2.20 | 1.69 |
| GHQ9 | 2.27 | 2.62 | 1.35 |
| GHQ10 | 2.3 | 2.60 | 1.52 |
| GHQ11 | 2.38 | 2.68 | 1.60 |
| GHQ12 | 2.07 | 2.21 | 1.71 |
| GHQ score (0-36) | 26.13 | 28.8386 | 18.8828 |