

# The 12-Item General Health Questionnaire (GHQ-12): Reliability, external validity and factor structure in the Spanish population

María del Pilar Sánchez-López and Virginia Dresch  
Universidad Complutense de Madrid

The purpose of this study was to analyze the internal consistency and the external and structure validity of the 12-Item General Health Questionnaire (GHQ-12) in the Spanish general population. A stratified sample of 1001 subjects, ages between 25 and 65 years, taken from the general Spanish population was employed. The GHQ-12 and the Inventory of Situations and Responses of Anxiety-ISRA were administered. A Cronbach's alpha of .76 (Standardized Alpha: .78) and a 3-factor structure (with oblique rotation and maximum likelihood procedure) were obtained. External validity of Factor I (Successful Coping) with the ISRA is very robust (.82; Factor II, .70; Factor III, .75). The GHQ-12 shows adequate reliability and validity in the Spanish population. Therefore, the GHQ-12 can be used with efficacy to assess people's overall psychological well-being and to detect non-psychotic psychiatric problems. Additionally, our results confirm that the GHQ-12 can best be thought of as a multidimensional scale that assesses several distinct aspects of distress, rather than just a unitary screening measure.

*El Cuestionario de Salud General de 12 Ítems (GHQ-12): fiabilidad, validez externa y estructura factorial en población española.* El objetivo de este estudio fue analizar la consistencia interna y la validez externa y estructural del Cuestionario de Salud General de 12 Ítems (GHQ-12) en la población general española. La muestra está compuesta por 1.001 sujetos, de 25 a 65 años de edad, de la población española. El GHQ-12 y el Inventario de Situaciones y Respuestas de Ansiedad fueron administrados. Se encontró un alfa de Cronbach de .76 (alpha estandarizado: 0.78) y una estructura de 3 factores (con rotación oblicua y el procedimiento de máxima verosimilitud). La validez externa del Factor I (Afrontamiento exitoso) con el ISRA fue muy alta (0.82; Factor II, .70; Factor III, .75). El GHQ-12 demostró una fiabilidad y validez adecuada en la población española. Por tanto, el GHQ-12 puede ser utilizado con eficacia para evaluar el bienestar psicológico y detectar problemas psiquiátricos no psicóticos. Además de ello, nuestros resultados confirman que el GHQ-12 se porta mejor como una escala multidimensional que evalúa varios aspectos del malestar psicológico, que utilizado como una medida de screening único.

The General Health Questionnaire (GHQ) is a self-administered screening questionnaire, designed for use in consulting settings aimed at detecting individuals with a diagnosable psychiatric disorder (Goldberg & Hillier, 1979). In its original version, it had 60 items (GHQ-60), which were reduced to 30 (GHQ-30), 28 (GHQ-28; in Spanish population see Gili, Ferrer, Roca, & Bernardo, 2000), and 12 items (GHQ-12) (Goldberg & Williams, 1988).

The 12-Item General Health Questionnaire (GHQ-12) is the most extensively used screening instrument for common mental disorders, in addition to being a more general measure of psychiatric well-being. Its brevity makes it attractive for use in

busy clinical settings, as well in settings in which patients need help to complete the questionnaire (Goldberg et al., 1997); its psychometric properties have been studied in various countries (Werneke, Goldberg, Yalcin, & Üstün, 2000) and with various types of population, for example, elderly people (Costa, Barreto, Uchoa, Firma, Lima-Costa, & Prince, 2006), and urological patients (Quek, Low, Razack, & Loh, 2001).

Although it has sometimes been considered unidimensional (Corti, 1994), various works have clearly revealed the existence of at least two factors in populations other than the Spanish one (Werneke, Goldberg, Yalcin, & Üstün, 2000). Factor analysis of the GHQ-12 has yielded 2- and 3-factor solutions, however, Gao, Luo, Thumboo, Fones, Li, and Cheung (2004) found that the 3-factor model proposed by Graetz (1991)—that is, three factors, namely Anxiety and Depression, Social Dysfunction, and Loss of Confidence—fit the data better than others models.

The customary type of scores used are a bimodel scale (0-0-1-1) and a 4-point Likert-type scale (0-1-2-3); the latter produces a more acceptable distribution of scores for parametric analysis (with less skew and kurtosis); Banks, Clegg, Jackson, Kemp,

Stafford and Wall (1980), recommended its use for the GHQ-12 to compare levels of psychiatric impairment within and between samples.

In the Spanish population, the psychometric properties of the GHQ-12 have been analyzed in adolescents (López-Castedo & Fernández, 2005), and puerperal women (Aguado, Navarro, Esteve, & Ascaso, 2003). González-Romá, Peiró, Luna, Baeza, Espejo and Muñoz (2003) analyses the factorial structure of the questionnaire (monofactorial and bifactorial model) in a reduced sample of Spanish adults (167 participants) and Claes & Fraccaroli (2002) show the factorial invariance (3 factors) of GHQ-12 in young people in 6 national contexts, including Spain. Also, Lobo and Muñoz (1996) refer to a study with an adult population in their chapter, but it is an unpublished study.

The purpose of this study is to analyze the internal consistency, factor structure, and external validity of the GHQ-12 in the Spanish general adult population, using a Likert-type scoring.

To assess external validity, we chose the Inventory of Situations and Responses of Anxiety-ISRA (Miguel-Tobal & Cano-Vindel, 2002) for several reasons:

1. The ISRA is an instrument that assesses anxiety; various previous works with the GHQ-12, as is reflected, found that anxiety is one of the principal components in diverse factor analyses with non-Spanish population (i.e., Graetz, 1991; Werneke, Goldberg, Yalcin & Üstün, 2000; Vanheule, & Bogarts, 2005).
2. On our previous investigations (with Spanish population), we found an important relation between the ISRA scores and various health indexes (i.e., Sánchez-López, López-García, Corbalán-Berna, & Dresch, submitted for publication).

In view of the previous results in non-Spanish population, which seem to confirm the multidimensionality of the GHQ-12 (i.e., Claes, & Fraccaroli, 2002; Gao, Luo, Thumboo, Fones, Li, & Cheung, 2004; and Graetz, 1991), we expect, firstly, that the ISRA will show a lower relation with the total GHQ-12 score than with each one of the factors that emerge, and secondly, that the ISRA will have a higher relation with the factor that, in turn, has a higher relation with anxiety.

## Methods

### Participants

The group was made up of 1001 subjects (601 women and 400 men), 50% aged between 25 and 44 years and 50% aged between 45 and 65 years (mean age= 41.75 years,  $SD= 10.95$ ), of various educational levels, selected from the general Spanish population.

### Instruments

The 12-Item General Health Questionnaire (GHQ-12) (Goldberg & Williams, 1988) consists of 12 items, each one assessing the severity of a mental problem over the past few weeks using a 4-point Likert-type scale (from 0 to 3). The score was used to generate a total score ranging from 0 to 36. The positive items were corrected from 0 (*always*) to 3 (*never*) and the negative ones from 3 (*always*) to 0 (*never*). High scores indicate worse health.

The Inventory of Situations and Responses of Anxiety-ISRA (Miguel-Tobal, & Cano-Vindel, 2002) is made up of 24 items that indicate the frequency with which certain anxiety-related reactions or thoughts occur, using a Likert-type score, ranging from 1 (*hardly ever*) to 5 (*almost always*), with high scores indicating high levels of anxiety.

### Procedure

The instruments were administered in 1-hour sessions and participation was voluntary. The statistical package SPSS, version 12.0, was used for data analysis. Data analysis techniques are described in the Results Section.

## Results

### Descriptive statistics

A mean GHQ-12 score of 8.52 ( $SD= 5.38$ ) was obtained in the general sample. As in numerous works, national and international, higher prevalence of mental-health problems are found in women, as measured with the GHQ-12 (for example, in the Spanish population, Cortés, Artacoz, Rodríguez-Sanz, & Borrell, 2004; Haro, & Pinto, 2006; Haro et al., 2006), it is useful to have some differential data for men and women. The women obtained a mean score of 9.30 ( $SD = 5.45$ ) and the men of 7.34 ( $SD= 5.05$ ). The difference was statistically significant ( $t= 5.83, p<.000$ ).

### Internal consistency

Cronbach's alpha was calculated to analyze internal consistency. We found an alpha value of .76 for the entire sample, Standardized Alpha: 0.78 (.75 in the group of women, Standardized Alpha: 0.77, and .76 in the group of men, Standardized Alpha: 0.78), indicating satisfactory internal consistency in all the groups. In table 1 is presented the item-scale analysis of the GHQ-12. The range of item-scale correlations is .7-.01, with item 11 being the one with the lowest correlation coefficient.

Table 1  
Item-scale analysis of the GHQ-12

	Adjusted Item-Scale Correlation	Cronbach Alpha if the item is eliminated
1. Able to concentrate	.48	.73
2. Loss of sleep over worry	.36	.75
3. Playing a useful part	.44	.74
4. Capable of making decisions	.46	.74
5. Felt constantly under strain	.41	.74
6. Couldn't overcome difficulties	.45	.74
7. Able to enjoy day-to-day activities	.39	.74
8. Able to face problems	.42	.74
9. Feeling unhappy and depressed	.57	.72
10. Losing confidence	.44	.74
11. Thinking of self as worthless	.01	.79
12. Feeling reasonably happy	.48	.73
Internal consistency of GHQ-12		
Entire sample	Alpha: 0.76	Standardized alpha: 0.78
Men	Alpha: 0.76	Standardized alpha: 0.78
Women	Alpha: 0.7	Standardized alpha: 0.77

*Factor structure*

Because of the undesirable properties of the orthogonal (varimax) rotation discussed by Graetz (1991), and his recommendations about the procedure to carry out the factor analysis of GHQ-12, oblique rotations (direct oblimin) were performed using the maximum likelihood procedure. Table 2 shows factor loadings after oblique rotation.

Three factors were obtained: Factor I is called «Successful Coping» Factor II is called «Self-esteem» and Factor III is called «Stress» Note that item 9, «feeling unhappy and depressed» presents loadings on two factors, positively on Factor II and negatively on Factor III.

Table 3 shows the eigenvalues and percentage of explained variance associated with each factor and their inter-correlations. First, all factors have eigenvalues that exceed the unit, a criterion frequently used to guide the number of meaningful factors. Second, the first factor is a major factor and accounts for more than one-third of the variance of the GHQ-12, whereas Factors II and III are minor factors. Together, all three factors account for 54.19% of the variance of the GHQ-12. Third, the factors are quite moderately inter-correlated; the correlation between Factor I and III is marginally higher than the other two correlations.

*Table 2*  
Maximum likelihood estimates of the oblique (direct oblimin) factor loadings for the 12-Item General Health Questionnaire

GHQ-Items	Factor loadings		
	I	II	III
1. Able to concentrate	.59	-	-
2. Loss of sleep over worry	-	-	-.63
3. Playing a useful part	.69	-	-
4. Capable of making decisions	.71	-	-
5. Felt constantly under strain	-	-	-.53
6. Couldn't overcome difficulties	-	.50	-
7. Able to enjoy day-to-day activities	.50	-	-
8. Able to face problems	.67	-	-
9. Feeling unhappy and depressed	-	.63	-.65
10. Losing confidence	-	.51	-
11. Thinking of self as worthless	-	.41	-
12. Feeling reasonably happy	.50	-	-
<b>KMO</b>	<b>Bartlett's sphericity</b>	<b>Determinant</b>	
.84	Chi square= 2717.34 <i>p</i> <.000	.06	

*Table 3*  
Eigenvalues, percentage of explained variance, inter-factor correlations and factor-total correlations for the GHQ-12

	Eigenvalue	Percentage explained variance	Inter-factor correlations		
			Factor I	Factor II	Factor III
Factor I	3.72	31.02	1.00		
Factor II	1.74	14.51	.10	1.00	
Factor III	1.03	8.66	-.42	-.33	1.00

*External validity*

External validity was analyzed by calculating the correlations between the total score of GHQ-12 and its three factors, and the ISRA's total score. Table 4 displays a correlation of .57 with the total GHQ-12 score, of .82 with Factor I, of .70 with Factor II, and of .75 with Factor III (all *ps*<.000).

Discussion

The reliability of the GHQ-12 in the general Spanish population is of .76, Standardized Alpha: 0.78, with little difference between men (.76, Standardized Alpha: 0.78) and women (.75, Standardized Alpha: 0.77). Although slightly lower than the index found in other populations (German, see Schmitz, Kruse, Heckrath, Alberti, & Tress, 1999; Australian, see Tait, Hulse, & Robertson, 2001; Iranian, see Montazeri, Harirchi, Shariati, Garmaroudi, Ebadi, & Fateh, 2003; and Arabic, see Daradkeh, Ghubash, & El-Rufaie, 2001), it is within the expected and acceptable values, as this instrument was designed for another population and in another language, and was subsequently adapted.

The functioning of item 11 should be confirmed in subsequent works with Spanish population to propose possible solutions, if necessary.

Regarding the factor structure, coinciding with various previous works; three factors emerged in the Spanish population: «Successful Coping», »Self-esteem», and «Stress». There are numerous studies that analyze the factor structure of the GHQ-12 with non-Spanish population, so it would be very protracted to compare in detail the data obtained by all of them. Therefore, the comparison is carried out with two studies that represent the majority tendencies of the results.

Gao, Luo, Thumboo, Fones, Li and Cheung (2004), after the confirmatory factor analysis of the proposed models, found that the model with the best fit was the 3-factor model proposed by Graetz (1991). Graetz re-examines the factor structure of the 12-item GHQ in a study in which the respondents were interviewed at yearly intervals on four separate occasions. The three factors proposed are basically similar to ours, with the same kind of loading, and with some small differences in the items that load on each factor. The percentages of variance explained by each factor are similar in the study of Graetz and in our study; the main difference lies in the order of the factors. Our Factor I is equivalent to their Factor II; our Factor II is equivalent to their Factor III, and our Factor III is equivalent to their Factor I. As noted by Werneke, Goldberg, Yalcin & Üstün, 2000, «factor analyzes in different

*Table 4*  
Pearson's correlation coefficients between GHQ-12 and ISRA

GHQ-12	ISRA	
	<i>r</i>	<i>p</i>
Overall scores	.57	.000
Factor I	.82	.000
Factor II	.70	.000
Factor III	.75	.000

setting including translation into different languages generally confirmed the original structure,” although «the ranking of the components may depend on the population under study» (p. 824). In contrast, Martin (1999) proposed an alternative structure based on an analysis of item content (Self-esteem, Stress and Successful Coping) and found that fits better than structures previously identified. Basically, our study corroborates the Martin’s results in general Spanish population.

Analysis of external validity reveals that the correlations of the GHQ-12 with the ISRA are significantly higher when the factors are analyzed independently than when the correlations with the total GHQ-12 score obtained are used. This finding corroborates the above-mentioned multidimensional properties of the questionnaire. The correlation of the first factor (Successful Coping) with the ISRA is especially important, as it indicates the weight of anxiety (at least, as measured by the ISRA), in this factor (items 1, 3, 4, 7, 8, and 12).

Despite the results of authors like Vanheule and Bogaerts (2005) and Graetz (1991), who conclude that «between-factor differences would suggest that GHQ has multidimensional properties that are not captured by a single severity score» (p. 133) and the results of this present work, confirming the multidimensional properties of the GHQ-12 and the greater external validity of each of its factors in comparison to the validity

of the total score, we provide the reliability and validity values of the total score because of the extensive use made of this test as a single severity score.

### Conclusions

The GHQ-12 displays adequate reliability and validity for use in the Spanish population. Its factor structure coincides, in the essential aspects, with that found in the more representative works with different kinds of populations. Between-factor differences suggest that the GHQ-12 has multidimensional properties that are not captured by a single severity score. The results of this work allow us to affirm that the GHQ-12 can be used effectively to assess the Spanish population’s overall psychological well-being and to detect non-psychotic psychiatric problems.

### Acknowledgements

The study shows the results of two plural-annual investigations financed by the Women’s Institute of the Ministry of Work and Social Affairs (Spain), reference numbers: 51/99 and 42/02.

Supported by the Programme AlBan, European Union Programme of High Level Scholarships for Latin America, identification number E03D01361BR.

### References

- Aguado, J., Navarro, P., Esteve, L., & Ascaso, C. (2003). Confirmatory factorial analysis of GHQ-12 in puerperal women [Análisis factorial confirmatorio del General Health Questionnaire (GHQ-12) en puérperas]. *IX Conferencia Española de Biometría*, 1-3.
- Banks, M.H., Clegg, C.W., Jackson, P.R., Kemp, N.J., Stafford, E.M., & Wall, T.D. (1980). The use of the General Health Questionnaire as an indicator of mental health in occupational studies. *Journal of Occupational Psychology*, 53, 187-194.
- Claes, R., & Fraccaroli, F. (2002). The General Health Questionnaire (GHQ-12): Factorial invariance in different language versions [II General Health Questionnaire (GHQ-12): Invarianza fattoriale in diverse versioni linguistiche]. *Bollettino di Psicologia Applicata*, 237, 25-35.
- Cortés, I., Artaco, L., Rodríguez-Sanz, M., & Borrell, C. (2004). Desigualdades sociales en la salud mental de la población ocupada [Social inequalities in mental health in the occupied population]. *Gaceta Sanitaria*, 18(5), 35, 1-9.
- Corti, L. (1994). For better or worse? Annual change in smoking, self-assessed health and subjective wellbeing (pp. 199-219). In N. Buck, J. Gershuny, & Colchester, S.J.: *Changing Households: The British Household Panel Survey 1990-1992*. University of Essex.
- Costa, E., Barreto, S.M., Uchoa, E., Firma, J.O.A., Lima-Costa, M.F., & Prince, M. (2006). Is the GDS-30 better than the GHQ-12 for screening depression in elderly people in the community? The Bambui Health Aging Study (BHAS). *International Psychogeriatrics*, 18, 493-503.
- 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, 85-94.
- Gao, F., Luo, N., Thumboo, J., Fones, C., Li, S., & Cheung, Y. (2004). Does the 12-item General Health Questionnaire contain multiple factors and do we need them? *Health and Quality of Life Outcomes*, 2, 63.
- Gili, M., Ferrer, V., Roca, M., & Bernardo, M. (2000). Psychiatric disorders and medical comorbidity in a community epidemiological study. *Psicothema*, 12, 131-135.
- Goldberg, D.P., Gater, R., Satorius, N., Üstün, T.B., Piccinelli, M., Gureje, O., & Rutter, M. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27, 191-197.
- Goldberg, D.P., & Hillier, V.F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139-145.
- Goldberg, D., & Williams, P. (1988). *A user’s guide to the General Health Questionnaire*. Windsor, UK: NFER-Nelson.
- González-Romá, V., Peiró, J.M., Luna, R., Baeza, J.A., Espejo, B., & Muñoz, P. (1991). A study of the factorial structure of the General Health Questionnaire (GHQ-12): Comparison of factorial models [Un estudio de la estructura factorial del Cuestionario de Salud General (GHQ-12): comparación de los modelos factoriales]. *Psicológica*, 12, 119-129.
- Graetz, B. (1991). Multidimensional properties of the General Health Questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, 26, 132-128.
- Haro, J. M., Palacín, C., Vilagut, G., Martínez, M., Bernal, M., Luque, I., Codony, M., Dolz, M., Alonso, J., & Grupo ESEMeD-España (2006). Prevalencia de los trastornos mentales y factores asociados: resultados del estudio ESEMeD-España [Prevalence of mental disorders and associated factors: Results of the ESEMeD-España study]. *Medicina Clínica*, 126, 445-51.
- Haro, J.M., & Pinto, A. (2006). El Pla Director: Anàlisi desde la perspectiva epidemiològica [The Director Plan: Analysis from the peidemiological perspective]. *Jornada Societat de Salut Pública de Catalunya i de Balears*. <http://www.academia.cat/societat/publica/docs/Ponent%20C.ppt> [electronic version].
- Lobo, A., & Muñoz, P.E. (1996). *General Health Questionnaire. Guide for the use of the different validated versions in Spanish language* [Cuestionario de Salud General GHQ. Guía para el usuario de las distintas versiones en lengua española validadas]. Barcelona: Masson.
- López-Castedo, A., & Fernández, L. (2005). Psychometric properties of the Spanish version of the 12-item General Health Questionnaire in adolescents. *Perceptual and Motor Skills*, 100, 676-680.
- Martin, A.J. (1999). Assessing the multidimensionality of the 12-Item General Health Questionnaire. *Psychological Reports*, 84, 927-935.

- Miguel-Tobal, J.J., & Cano-Vindel, A.R. (2002). *Inventory of Situations and Response of Anxiety (ISRA): Handbook* [Inventario de situaciones de respuestas de ansiedad (ISRA): manual]. Madrid: TEA Ediciones.
- 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.
- Quek, K.F., Low, W.Y., Razack, A.H., & Loh, C.S. (2001). Reliability and validity of the General Health Questionnaire (GHQ-12) among urological patients: A Malaysian study. *Psychiatry and Clinical Neurosciences*, 55, 509-513.
- Sánchez-López, M.P., López-García, J.J., Corbalán-Berna, F.J., y Dresch, V. (sometido a evaluación). Sociodemographic, psychological and health related factors associated with poor mental health in midlife Spanish women. *Women & Health*.
- Schmitz, N., Kruse, J., Heckrath, C., Alberti, L., & Tress, W. (1999). Diagnosing mental disorders in primary care: The General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90) as screening instruments. *Social Psychiatry and Psychiatric Epidemiology*, 34, 360-366.
- Tait, R., Hulse, G.K., & Robertson, S.I. (2002). A review of the validity of the General Health Questionnaire in adolescent populations. *Australian and New Zealand Journal of Psychiatry*, 36, 550-557.
- Vanheule, S., & Bogaerts, S. (2005). The factorial structure of the GHQ-12. *Stress & Health*, 21, 217-222.
- 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, 823-829.