

This is the peer reviewed version of the following article:

Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents

Alicia Padrón, Iñaki Galán, María Durbán, Ana Gandarillas, Fernando Rodríguez-Artalejo

Qual Life Res. 2012 Sep;21(7):1291-8.

which has been published in final form at

<https://doi.org/10.1007/s11136-011-0038-x>

- Padrón A, Galán I, Durbán M, Gandarillas A, Rodríguez-Artalejo F. Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents. *Qual Life Res.* 2012; 21(7): 1291-1298. Impact Factor: 2.486 (Q2). (doi: <https://doi.org/10.1007/s11136-011-0038-x>)

Confirmatory factor analysis of the General Health Questionnaire (GHQ-12) in Spanish adolescents

Authors: Alicia Padrón^a, Iñaki Galán^{b,c}, María Durbán^d, Ana Gandarillas^e, Fernando Rodríguez-Artalejo^{c,f}

^a Public Health Centre, Area 10. Consejería de Sanidad de Madrid. C/ Francisco Gasco Santillán 2, portal B 2º, Getafe (Spain)

^b National Centre for Epidemiology. Instituto de Salud Carlos III. Madrid. C/ Monforte de Lemos 5, Madrid (Spain)

^c Department of Preventive Medicine and Public Health. School of Medicine. Universidad Autónoma de Madrid/ IdiPAZ. C/Arzobispo Morcillo 2, Madrid (Spain)

^d Department of Statistics and Econometrics. Universidad Carlos III de Madrid. C/ Madrid 126. Getafe (Spain).

^e Department of Epidemiology. Consejería de Sanidad de Madrid. C/ Julián Camarillo 4, Madrid (Spain).

^f CIBER of Epidemiology and Public Health (CIBERESP). C/ Melchor Fernandez Almagro 3-5, Madrid (Spain).

Declaration of conflict of interest: None

Word count: 2921.

Address for correspondence: Iñaki Galán, MD, PhD.

National Centre for Epidemiology

Instituto de Salud Carlos III, Madrid

C/ Monforte de Lemos 5, Madrid (Spain)

Tfn: 918222679/FAX: 913877815/ igalan@isciii.es

Abstract

Purpose. To evaluate the factor structure of the 12-item General Health Questionnaire (GHQ-12) in a population of Spanish adolescents.

Methods. Cross-sectional study among 4,146 individuals (mean age 16.3 years). The students completed a health questionnaire which included the GHQ-12. A confirmatory factor analysis (CFA) was conducted with four factor-structure models. Three of the models were theory-driven: unidimensional, two-dimensional (positive and negative questions), model proposed by Graetz (anxiety and depression, social dysfunction, loss of confidence); and the fourth model was based on our exploratory factor analysis (EFA).

Results. The Cronbach's alpha for internal consistency was 0.82. A three-dimensional structure was identified in the EFA. The first factor included items 1, 2, 5, 6, 7 and 9 of the GHQ-12; the second, items 3, 10 and 11; and the third, items 4, 8 and 12. The three factors together explained 53.7% of the variance. The model with the best fit in the CFA was the three-dimensional model proposed by Graetz, followed by the three-dimensional model derived from the EFA. These two models had acceptable goodness-of-fit indices.

Conclusions. In an adolescent population from Southern Europe, the GHQ-12 showed high internal consistency. The factor structure that best fitted the data was the Graetz three-dimensional model. However, the high correlations observed between factors suggest that the GHQ-12 should be used as a unidimensional scale, as currently done.

Keywords. Mental health, confirmatory factor analysis, adolescents, health survey, Spain.

Abbreviations. General Health Questionnaire (GHQ)

Confirmatory Factor Analysis (CFA)

Exploratory Factor Analysis (EFA)

Root Mean Square Error of Approximation (RMSEA)

Comparative Fit Index (CFI)

Goodness of Fit Index (GFI)

Expected Cross Validation Index (ECVI)

Akaike Information Criterion (AIC)

Introduction.

The prevalence of mental disorders among people aged 12-24 years ranges from 8% to 57%, according to different studies, and it is considered that an average of one in every five young people will suffer from some type of mental health problem [1]. Some psychological disorders that begin in adolescence persist into adulthood, but they may improve if treated early [2].

The General Health Questionnaire (GHQ) is a self-administered questionnaire to assess psychiatric disorders [3]. It was designed by Goldberg to be used in non-psychiatric clinical settings [3]. Nevertheless, the GHQ is also widely used as a health questionnaire in healthy community settings [4]. The shortest version consists of 12 items (GHQ-12) and has shown satisfactory validity and reliability in several studies [3,5-8].

Although the GHQ was originally aimed at the adult population, it has also been used with success in adolescents [9-14]. French and Tait [10] showed that children aged 11-15 years interpret the GHQ in a manner similar to adults.

The GHQ was designed originally as a unidimensional scale. However, several papers have explored its possible dimensions. Even Goldberg himself developed a scaled version of the GHQ. According to him "it was intended for studies in which an investigator requires more information than is provided by a single severity score". Since then various studies have shown different dimensions in the structure of the GHQ versions and, in particular, in the GHQ-12 [6,7,13-16]. The three-dimensional model proposed by Graetz consists of the following dimensions: anxiety and depression (which includes items 2, 5, 6, and 9), anhedonia and social dysfunction (items 1, 3, 4, 7, 8, and 12) and loss of confidence or self-esteem (items 10 and 11), and is usually

considered to be the most appropriate [9,10,16-19]. However, there is no consensus as to the validity or usefulness of the different dimensions.

Little information is available on the psychometric characteristics of the GHQ-12 in adolescents from Southern Europe. In Spain, one study has assessed its internal consistency and explored its dimensions [13]; to date, however, no confirmatory factor analysis (CFA) has been conducted to evaluate the factor structure of the GHQ-12 and to assess how well it fits with different theory-driven models formulated *a priori*. The objective of this study was to evaluate the factor structure of the GHQ-12 by using CFA in a population of Spanish adolescents.

Methods.

Study design and population. The study data were drawn from the Non-Communicable Diseases Risk Factor Surveillance System (*Sistema de Vigilancia de Factores de Riesgo de Enfermedades No Transmisibles-SIVFRENT*) targeting the adolescent population. This system monitors the principal health-related lifestyles in a representative sample of the population in the fourth year of Compulsory Secondary Education in the region of Madrid [20].

The study participants were selected through two-stage cluster sampling, with the first-stage units (schools) stratified by area (city of Madrid and other municipalities) and type of school (public or private). The questionnaire was self-administered in the classroom under the guidance of experienced field workers. The participation of schools and students was voluntary, after obtaining informed consent. For this study, we analyzed data from 4,146 adolescents that completed the survey in 2008 and 2009 in 94 schools which included 185 classrooms.

The response rate was 81.3% for all the selected schools and 90.9% for students within the schools (8.3% were absent on the day of the survey, 0.08% refused to participate, and 0.7% questionnaires were removed due to inconsistent responses). Thus, the overall study response (schools and students) was 73.9%. We collected data from 4,210 students. There was some relevant information missing from 1.5% of the questionnaires. The data from those students were excluded, so that the final analytical sample consisted of 4,146 individuals.

Questionnaire. The questionnaire included the GHQ-12, validated in Spanish by Lobo and Muñoz [3]. The GHQ-12 items can be classified as either positively or negatively worded. An item is considered to be positively worded (6 items) if it refers to health. The response options range from “More than usual” to “Much less than usual.” Items that refer to disease are considered to be negatively worded (6 items). The response options range from “Not at all” to “Much more than usual.” Responses are assigned to a 4-point Likert scale. The score for each response on the GHQ-12 ranges from 0 to 3, with the overall score ranging from 0 to 36. A cut-point of 12 or higher was used to classify subjects as “cases” [5].

Statistical analysis. We first calculated the distribution of the scores on each question and on the overall scale, estimating the percentage of individuals with the highest scores (ceiling effect) and the lowest scores (floor effect). Reliability or internal consistency was evaluated by the Cronbach’s alpha coefficient. EFA was subsequently conducted using the method of principal components analysis. The criteria for selecting the number of factors were the eigenvalues (minimum value close to 1) and their graphic representation (scree plot). An oblique rotation (promax=2) was used to identify and characterize these factors. We preferred Promax-rotation over the Oblimin

one because despite both are similar, Promax is computationally faster for large datasets [21]. To assign an item to a particular factor we selected factor loadings equal to or greater than 0.4 [22].

A CFA was then performed with four factor structure models: 1) unidimensional; 2) two-dimensional: positively and negatively worded items; 3) three-dimensional, as proposed by Graetz-[16]: anxiety and depression (four negative items: 2, 5, 6, and 9), anhedonia and social dysfunction (six positive items: 1, 3, 4, 7, 8, and 12), and loss of confidence or self-esteem (two negative items: 10, 11); 4) dimensionality obtained from the EFA.

We conducted an EFA before the CFA for various reasons. First, because few EFA have been carried on adolescents [6,13,14] and only one of them was done in Spain [13]. Second, because in most CFA in adolescents, the three-dimensional model proposed by Graetz [16] showed the best fit to the data; that model, however, was derived from a study in an adult population. Given that CFA only assesses models specified *a priori*, we decided to test a model obtained in our study population.

The selection of the rest of the models was based on a literature review. The model most strongly supported by studies with CFA is that of Graetz [9,10,16-19]. The unidimensional model was also used because the GHQ-12 was designed as an overall scale; in addition, some authors who identified several dimensions with the CFA recommended using a unidimensional model to obtain a uniform score [9,10]. Other authors have assessed the existence of two dimensions, one with the positively worded items and the other with the negatively worded items [23], and there has even been a

study of possible response bias for the negatively worded items [24]. Lastly, French and Tait [10] also suggested that a two-dimensional model of positive and negatively worded items can be obtained by grouping the first and third factors in the Graetz [16] model, which they found to be highly correlated.

The CFA parameters were estimated using the generalized weighted least squares method using asymptotic covariance matrix. The use of the asymptotic matrix relaxes the hypothesis on the distribution of the observed variables, and improves the fit and the statistics used in the tests [25]. To evaluate the models, we considered various indicators of goodness of fit. The Root Mean Square Error of Approximation (RMSEA) measures the discrepancy between the observed covariance matrix and the one adjusted by the model [26]. An RMSEA of less than 0.05 is considered to indicate a good model fit, and an RMSEA of less than 0.08 is considered a reasonable fit. The Comparative Fit Index (CFI) and the Goodness of Fit Index (GFI) evaluate the improvement in the fit in comparison with a simpler model [27], generally one that assumes independence among the items. These indices vary between 0 and 1, and values over 0.95 are considered acceptable [28]. The Expected Cross Validation Index (ECVI) also compares models [29], with a lower value indicating better model fit. The Akaike Information Criterion (AIC), which penalizes the likelihood function based on the number of parameters in the model, is used in a similar way. A model with lower AIC is considered more plausible than one with a higher AIC. We also assessed, for the best fitting model, the standard errors and t-values for the factor loadings.

The descriptive analyses and the EFA were performed with SPSS 16.0 for Windows (2008 SPSS Inc., Chicago, IL, USA). For the CFA, LISREL 8.8 for Windows (Scientific Software International Inc., Lincolnwood, IL, USA) was used.

Results.

The mean age of the study participants was 16.3 years (range 13.7 to 18.7 years) and 51.6% were females.

The mean GHQ-12 score was 10.9 (SD=5.3). Taking ≥ 12 as the cut-point, 38.5% of participants were considered as cases. Table 1 shows the mean score for each item and the percentage of responses in each response category. The items with the highest mean scores were 1, 2 and 5; of these, item 5 “Have you felt constantly under strain?” was notable for having the highest score. About 0.2% of the sample had zero points on the overall scale, and only one person had the maximum score.

The Cronbach’s alpha for internal consistency of the GHQ-12 was 0.82. Table 2 shows the correlation of each item with the overall scale, and the Cronbach’s alpha after eliminating the corresponding item. Item 4 “Have you felt capable of making decisions about things?” had the lowest correlation coefficient (0.358). For the rest of the items the correlations with the overall scale ranged between 0.492 and 0.742. The Cronbach’s alpha did not change substantially when eliminating the corresponding item.

The EFA suggests a three-dimensional structure. All the factor loadings that permit assignment of an item to a specific factor exceeded 0.4 (Table 3). The first factor includes items 1, 2, 5, 6, 7 and 9, and explains 34.7 % of the variance. The second

factor includes items 3, 10 and 11, and explains 10.9% of the variance. And the third one includes items 4, 8 and 12, and explains 8.2% of the variance. The three factors together explain 53.7 % of the variance.

In the CFA, all the models had an RMSEA of less than 0.08 and the values for the rest of the indicators were acceptable, but only Graetz's three-dimensional model [16] had an RMSEA of less than 0.06 (Table 4). In general, the best results for all the goodness-of-fit indices were obtained with this model. The second best goodness-of-fit indices were obtained with the model derived from the EFA, whereas the unidimensional model yielded the worst values.

For Graetz's three-dimensional model, the standard errors and t-values for factor loadings (Table 5) were statistically significant.

The factors in the three-dimensional model proposed by Graetz [16] were highly correlated (Figure 1). The correlation was 0.84 between the first and third factor, 0.74 between the second and third factor, and 0.72 between the first and second factor. Given the high correlation between factor 1 and 3, a nested model with two factors was also fitted. The chi-square difference was 314.3 with 2 degrees of freedom, so that the model with three factors was preferred. Similar results were obtained when the model was compared with other nested models.

Discussion.

In Spanish adolescents, our results show that the GHQ-12 has a high internal consistency and that the three-factor structure proposed by Graetz [16] yields the best

fit in the CFA. However, the high between-factor correlation observed may question the use of the GHQ-12 beyond a unidimensional scale.

The satisfactory value that we obtained for the Cronbach's alpha is consistent with the values in the range 0.81-0.88 reported in other healthy adolescent populations [6,10,11,13,14,30]. We did not explore the Cronbach's Alpha of the different subscales because its use could be questioned when the measurement scale is composed by very few items, as it happens with factors 2 and 3 of our Exploratory Factor Analysis, each of them comprising 3 items which is considered to be the minimum number for a scale aiming to explore a factor [31]. Moreover, we also consider that the reliability of one test (analyzed with the Cronbach's Alpha) is meaningful when it is composed (globally through its subscales) by a number of items combined to represent a global single summative score which is not the case of our present subscales. The mean GHQ-12 score is also similar to that found in other studies with non clinical adolescent populations [10,14]. To our knowledge, most of the articles that assessed the validity and/or reliability of the GHQ-12 in adolescents are based in healthy community samples, such as ours. The ceiling/floor effects of the overall scale were small.

A three-dimensional structure of the GHQ-12 was identified with EFA. The first dimension includes items 1, 2, 5, 6, 7 and 9. The second includes items 3, 10 and 11, and coincides with Graetz's [16] "loss of confidence" dimension plus item 3. The third dimension includes items 4, 8 and 12. This three-dimensional model explains 53.7% of the variance. Lopez-Castedo and Fernandez [13] obtained a two-dimensional model in an EFA of an adolescent population in Spain. The first factor (anxiety) included items 1, 2, 5, 6, 9, 10, and 11, and the second one (social dysfunction) included items 3, 4, 7,

8 and 12. This model is similar to the one obtained in our study. With the exception of items 3 and 7, our first two dimensions are the same as the “anxiety” dimension described by Lopez-Castedo and Fernandez [13]. However, their two-dimensional model explained less variance (37.8%) than ours did.

Politi *et al.* [6], in a study of 18-year-old males, identified two dimensions that explained 46.7% of the total variance. These were “general dysphoria”, which included items 2, 5, 6, 9, 10, 11 and 12, and “social dysfunction” which included items 1, 3, 4, 7, 8 and 12. This model is also similar to the model of Lopez-Castedo and Fernandez [13].

Li *et al.* [14] identified two factors in Chinese adolescents that explained 53.3% of the variance. The first included items 1, 3, 4, 7, 8 and 12, and the second included 2, 5, 6, 9, 10 and 11. Their model classified the items as positive or negative, and coincides with the model of Politi *et al.* [6]. With the exception of items 1, 3 and 7, our first two dimensions are the same as their second factor.

Few studies in adolescents have examined the GHQ-12 with CFA [9,10,14]. As far as we know, ours is the first such study in southern Europe. In the study of Li *et al.* [14], the models with the best fit were the two-dimensional model with 8 items, described by Kalliath *et al.* [32], followed by the three-dimensional model of Worsley and Gribbin [33] and Graetz’s three-dimensional model with 12 items [16]. In our study, Graetz’s model yielded the best fit in the CFA [16]. The results of the modification indices for the matrix of factor loadings and the measurement error covariance matrix showed that there was no possible additional path(s) that would significantly improve the fit of the hypothesized structural model.

This result is consistent with other studies in adolescents [9,10], but the validity and usefulness of the different dimensions are subject to debate. The high correlations between factors in the CFA, that have also been assessed in other studies [9,10], suggests they are not independent and can be explained by signs and symptoms common to all the factors. For example, anxiety and depression can cause insomnia. In turn, depression can lead to lower self-esteem and social functioning. Furthermore, chronic stress promotes depression. The blurred boundaries between the factors of psychological distress, together with the statistical evidence of high correlations between the dimensions, would support the use of the GHQ-12 only to obtain an overall score, without assigning too much value to the different dimensions.

Although a possible response bias between the positively and negatively worded statements has been investigated, we know of no study, including ours, in which the two-dimensional model of positively and negatively phrased items yields the best fit in the CFA. French and Tait [10] in a CFA with adults found that the model with the best fit was Graetz's three-dimensional model [16], but the correlations between the first and third dimensions (negatively phrased statements) were much higher than with the second dimension (positively phrased statements). For this reason, they suggested the possibility of a bias in the negatively worded items. But in our study the correlations were only slightly higher between the first and third factor. Cheung [17] in an older population, obtained even higher correlations between the second and third factors than between the first and second. However, we find interesting that this important issue related with the possible response bias, that is beyond our current objectives, should be evaluated thoroughly in further analysis of the data.

In our study, the model derived from the EFA did not yield the best fit in the CFA. This is not uncommon [34] because the possible sources of model misfit are not in the EFA. In CFA it is usually assumed that each item loads on only one factor, whereas in EFA each item may depend on all the common factors, making it difficult to interpret the results. Furthermore, EFA does not allow evaluation of the model because identification of the factors is based on arbitrary cut-points, and it is not possible to use goodness-of-fit criteria. Li *et al.* [14] obtained similar results in the context of the GHQ-12.

Our study has some limitations. Specifically, although the response rate was quite high, it is possible that student absence on the day of the survey was related to some health problem that could affect the GHQ-12. However, given that only 8.3% of the students were absent during the survey, it could have only a modest impact on the results of the psychometric characteristics of the GHQ-12. The main strength of the study was its large size and the fact that it was representative of all students in upper Secondary Education in Madrid. This allowed considerable variation in the GHQ-12 score, so that the results can be extrapolated to similar populations.

Conclusions

We conclude that, in an adolescent population from Southern Europe, the GHQ-12 has shown good psychometric characteristics. Also the factor structure that best fitted the data in CFA was the Graetz three-dimensional model [16]. However, the high correlations observed between factors suggest that the GHQ-12 should be used as a unidimensional scale, as currently done.

Acknowledgements.

Reference list.

1. Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: a global public-health challenge. *Lancet*, 369, 1302-1313.
2. Fombonne, E. (1998). Increased rates of psychosocial disorders in youth. *European Archives of Psychiatry and Clinical Neuroscience*, 248, 14-21.
3. Lobo, A., & Muñoz, P.E. (1996). *Cuestionario de salud general GHQ (General Health Questionnaire). Guía para el usuario de las distintas versiones en lengua española validadas*. Barcelona: Masson.
4. Goldberg, D. P., & Blackwell, B. (1970). Psychiatric illness in general practice. A detailed study using a new method of case identification. *British Medical Journal*, 1, 439-443.
5. 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, 191-197.
6. 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, 432-437.
7. Sanchez-Lopez, M. P., & Dresch, V. (2008). The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity and factor structure in the Spanish population. *Psicothema*, 20, 839-843.
8. 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-R) as screening instruments. *Social Psychiatry and Psychiatric Epidemiology*, 34, 360-366.

9. Sweeting, H., Young, R., & West, P. (2009). GHQ increases among Scottish 15 year olds 1987-2006. *Social Psychiatry and Psychiatric Epidemiology*, 44, 579-586.

10. French, D. J., & Tait, R. J. (2004). Measurement invariance in the General Health Questionnaire-12 in young Australian adolescents. *European Child and Adolescent Psychiatry*, 13, 1-7.

11. Tait, R. J., French, D. J., & Hulse, G. K. (2003). Validity and psychometric properties of the General Health Questionnaire-12 in young Australian adolescents. *Australian and New Zealand Journal of Psychiatry*, 37, 374-381.

12. Banks, M. H. (1983). Validation of the General Health Questionnaire in a young community sample. *Psychological Medicine*, 13, 349-353.

13. Lopez-Castedo, A., & Fernandez, L. (2005). Psychometric properties of the Spanish version of the 12-item General Health Questionnaire in adolescents. *Perceptual and Motor Skills*, 100, 676-680.

14. 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, 3253-3261.

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

16. Graetz, B. (1991). Multidimensional properties of the General Health Questionnaire. *Social Psychiatry and Psychiatric Epidemiology*, 26, 132-138.
17. Cheung, Y. B. (2002). A confirmatory factor analysis of the 12-item General Health Questionnaire among older people. *International Journal of Geriatric Psychiatry*, 17, 739-744.
18. 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.
19. 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, 231-236.
20. Consejería de sanidad. (2009) Informe sobre hábitos de salud en la población juvenil de la Comunidad de Madrid . Boletín Epidemiológico de la Comunidad de Madrid 15 nº2. Retrieved from <http://www.madrid.org/cs/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-disposition&blobheadername2=cadena&blobheadervalue1=filename%3DFebrero2009.pdf&blobheadervalue2=language%3Des%26site%3DPortalSalud&blobkey=id&blobtable=MungoBlobs&blobwhere=1220583125549&ssbinary=true>
Accessed 23 December 2010.
21. Tabachnick, B., & Fidell, L. (2007). *Using Multivariate Statistics*. 2 nd edition. Boston: Pearson.

22. Watson, R., & Thompson, D. R. (2006). Use of factor analysis in Journal of Advanced Nursing: literature review. *The Journal of Advanced Nursing*, 55, 330-341.
23. Andrich, D., & van Schoubroeck, L. (1989). The General Health Questionnaire: a psychometric analysis using latent trait theory. *Psychological Medicine*, 19, 469-485.
24. Hankins, M. (2008). The factor structure of the twelve item General Health Questionnaire (GHQ-12): the result of negative phrasing?. *Clinical Practice and Epidemiology in Mental Health*, 4, 10.
25. Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Lincolnwood, USA: Scientific Software International.
26. Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. Scott-Long (Eds.), *Testing structural equation models* (pp.136-162). Newbury Park: Sage Publications.
27. Bentler, P. M. (1990). Comparative fit indexes in structural models 2. *Psychological Bulletin*, 107, 238-246.
28. Hu, Y., Stewart-Brown, S., Twigg, L., & Weich, S. (2007). Can the 12-item General Health Questionnaire be used to measure positive mental health. *Psychological Medicine*, 37, 1005-1013.
29. Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariation structures. *Multivariate Behaviour Research*, 24, 445-455.

30. Tait, R. J., 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.
31. Streiner, D.L. (1994). Figuring out factors: the use and misuse of factor analysis. *Can J Psychiatry*, 39,135-140.
32. Kalliath, T. J., O'Driscoll, P. O., & Brough, P. (2004). A confirmatory factor analysis of the General Health Questionnaire-12. *Stress and Health*, 20, 11-20.
33. Worsley, A., & Gribbin, C. C. (1977). A factor analytic study on the twelve item general health questionnaire. *Australian and New Zealand Journal of Psychiatry*, 11, 260-272.
34. Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, NY: The Guilford Press.

Table 1.

Mean and standard deviation of replies to each item in the GHQ-12, percentage of subjects classified as cases, and score distribution for each item.

GHQ-12 items	Mean	SD ¹	Score ²			
			0	1	2	3
			%	%	%	%
Have you...						
1. Been able to concentrate on whatever you are doing?	1.14	0.58	8.1	71.8	17.5	2.7
2. Lost much sleep over worry?	1.02	0.91	34.6	34.4	24.9	6.1
3. Felt that you were playing a useful part in things?	0.99	0.58	14.8	73.4	9.3	2.6
4. Felt capable of making decisions about things?	0.76	0.58	30.7	63.0	5.4	1.0
5. Felt constantly under strain?	1.32	0.93	22.7	31.9	35.3	10.1
6. Felt that you couldn't overcome your difficulties?	0.99	0.93	37.5	32.3	23.9	6.4
7. Been able to enjoy your normal day-to-day activities?	0.98	0.58	16.8	69.3	12.8	1.1
8. Been able to face up to your problems?	0.94	0.6	19.5	67.7	11.4	1.4
9. Been feeling unhappy and depressed?	0.92	0.97	45.2	23.7	24.8	6.3
10. Been losing confidence in yourself?	0.54	0.82	64.2	19.5	13.5	2.7
11. Been thinking of yourself as a worthless person?	0.37	0.75	75.9	13.6	7.5	3.0
12. Been feeling reasonably happy, all things considered?	0.91	0.63	22.5	64.9	10.8	1.8
Mean overall score	10.94	5.31				
Classified as cases (score ≥ 12) %	38.5%					

¹ SD: Standard deviation.

² Higher scores indicate a worst situation in each item.

Table 2.

Correlation between items and overall GHQ-12 scale. Cronbach's alpha when eliminating the corresponding item.

GHQ-12 items	Correlation of item with overall scale	Cronbach's alpha when eliminating each item
Have you...		
1. Been able to concentrate on whatever you are doing?	0.499	0.817
2. Lost much sleep over worry?	0.591	0.814
3. Felt that you were playing a useful part in things?	0.495	0.817
4. Felt capable of making decisions about things?	0.358	0.826
5. Felt constantly under strain?	0.639	0.809
6. Felt that you couldn't overcome your difficulties?	0.701	0.800
7. Been able to enjoy your normal day-to-day activities?	0.507	0.816
8. Been able to face up to your problems?	0.492	0.817
9. Been feeling unhappy and depressed?	0.742	0.795
10. Been losing confidence in yourself?	0.690	0.801
11. Been thinking of yourself as a worthless person?	0.656	0.804
12. Been feeling reasonably happy, all things considered?	0.562	0.812
Cronbach's alpha	0.824	

Table 3.

Exploratory factor analysis of the GHQ-12. Factor loadings for each item and proportion of variance explained.

GHQ-12 items	Factor loadings		
	Factor 1	Factor 2	Factor 3
Have you...			
1. Been able to concentrate on whatever you are doing?	0.498		
2. Lost much sleep over worry?	0.722		
3. Felt that you were playing a useful part in things?		0.554	
4. Felt capable of making decisions about things?			0.718
5. Felt constantly under strain?	0.700		
6. Felt that you couldn't overcome your difficulties?	0.465		
7. Been able to enjoy your normal day-to-day activities?	0.544		
8. Been able to face up to your problems?			0.632
9. Been feeling unhappy and depressed?	0.539		
10. Been losing confidence in yourself?		0.749	
11. Been thinking of yourself as a worthless person?		0.795	
12. Been feeling reasonably happy, all things considered?			0.409
% of variance explained by each factor	34.7%	10.9%	8.2%
% of total variance explained by all three factors: 53.7%			

Table 4.

Confirmatory factor analysis of the GHQ-12. Goodness-of-fit indices for each model.

Models ¹	X ²	DF	RMSEA (90% CI)	GFI	CFI	ECVI	AIC
Unidimensional	1370.9	54	0.077 (0.073-0.080)	0.98	0.96	0.34	1418.9
Two-dimensional (positive and negative items)	1041.6	53	0.067 (0.064-0.071)	0.99	0.97	0.26	1095.6
Three-dimensional (Graetz)	774.61	51	0.057 (0.054-0.061)	0.99	0.98	0.19	798.1
Three-dimensional (EFA)	801.21	51	0.062 (0.058-0.065)	0.99	0.98	0.22	909.2

X²: Chi square; DF: Degrees of freedom; RMSEA: Root Mean Square Error of Approximation; GFI: Goodness of Fit Index; CFI: Comparative Fit Index; ECVI: Expected Cross Validation Index; AIC: Akaike Information Criterion; EFA: Exploratory Factor Analysis

¹ Models:

Unidimensional: Factor 1: 1,2,3,4,5,6,7,8,9,10,11,12

Two-dimensional: Factor 1: 1,3,4,7,8,12; Factor 2: 2,5,6, 9,10,11

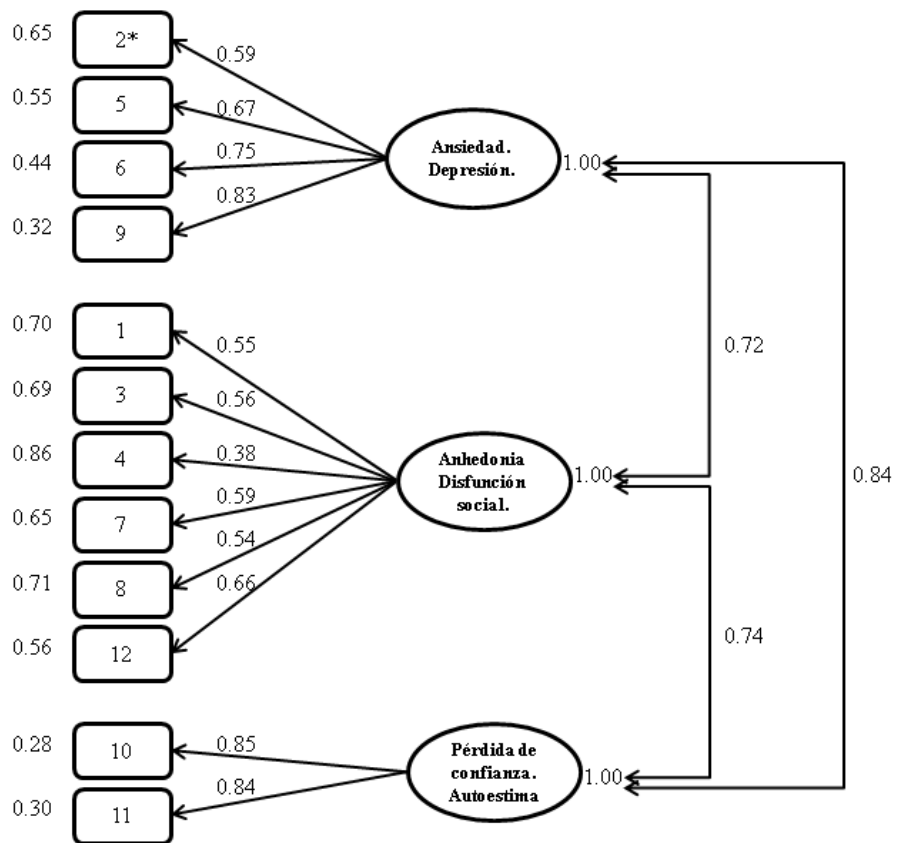
Three-dimensional (Graetz): Factor 1: 2,5,6,9; Factor 2: 1,3,4,7,8,12; Factor 3: 10,11

Three-dimensional (EFA) : Factor 1: 1,2,5,6,7,9; Factor 2: 3,10,11; Factor 3: 4,8,12

Figure. 1

Flow diagram of the three-dimension model proposed by Graetz for the GHQ-12.

Presented from left to right: variance of item not explained by its factor, correlations of items with their corresponding factor, and correlations between factors.



*Número de item.