

Validation of the structural coherency of the General Health Questionnaire

Validação da coerência estrutural do Questionário de Saúde Geral

Hudson Wander de Carvalho,¹ Christopher J. Patrick,² Miguel Roberto Jorge,¹ Sérgio Baxter Andreoli^{1,3}

¹ Department of Psychiatry, Escola Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil

² Department of Clinical Psychology and Psychopathology, The Florida State University, USA

³ Universidade Católica de Santos, Santos, SP, Brazil

Abstract

Objective: Investigate the structural coherency of the 60-item version of the General Health Questionnaire via exploratory and confirmatory factor analyses. **Method:** The study design is a cross-sectional survey. A random sample of 146 individuals from the city of Divinópolis-MG volunteered to participate in the present study and responded to the 60-item version of the General Health Questionnaire adapted and validated for use in Brazil. Statistics consisted of exploratory and confirmatory factor analysis. Reliability was estimated using Cronbach's alpha method. **Results:** Alpha coefficients for all five content scales of the General Health Questionnaire were high ($\alpha > 0.8$). For four of the five scales, a unifactorial model of constituent items provided a good fit to the data. Items comprising the fifth scale, Psychic Stress, exhibited a two-correlated factor structure. A factor analysis of scores for the five scales yielded strong evidence of coherency, with all scales loading substantially on a single common factor. **Conclusion:** The General Health Questionnaire shows good psychometric coherency as evidenced by high internal consistency and unidimensionality of all but one of its constituent scales, and uniformly high loadings of all scales on a single overarching factor. These results are consistent with prior findings from the General Health Questionnaire developmental study and Brazilian adaptation studies.

Descriptors: Mass screening; Psychometrics; Factor analysis; Validity of tests; Questionnaires

Resumo

Objetivo: Investigar a coerência estrutural da versão de 60 itens do Questionário de Saúde Geral de Goldberg por meio da aplicação de análises fatoriais exploratórias e confirmatórias. **Método:** Trata-se de um estudo de levantamento de desenho transversal. Uma amostra aleatória de 146 indivíduos voluntários da cidade de Divinópolis-MG participou do presente estudo respondendo à versão de 60 itens do Questionário de Saúde Geral. As estatísticas consistiram de análises fatoriais exploratórias e confirmatórias. Fidedignidade foi estimada pelo método alpha de Cronbach. **Resultados:** Os coeficientes alpha para todas as cinco escalas do Questionário de Saúde Geral foram elevados ($\alpha > 0,8$). Para quatro das cinco escalas, um modelo unifatorial com base em seus respectivos itens apresentou bom ajuste aos dados. Itens que compunham a quinta escala, Estresse Psíquico, exibiram uma estrutura de dois fatores correlacionados. Uma análise fatorial dos escores das cinco escalas produziu fortes evidências de coerência, com todas as escalas apresentando cargas fatoriais robustas em um único fator. **Conclusão:** O Questionário de Saúde Geral apresenta bons indicadores psicométricos de coerência deflagrados pela elevada consistência interna e unidimensionalidade de todas menos uma de suas escalas constituintes e pelas cargas fatoriais, univocamente elevadas, em uma única e inclusiva dimensão. Os resultados apresentados são congruentes tanto com os do estudo de elaboração do Questionário de Saúde Geral quanto com os de sua adaptação para o Brasil.

Descritores: Programas rastreamento; Psicometria; Análise fatorial; Validade dos testes; Questionários

Introduction

The General Health Questionnaire (GHQ), originally developed in the United Kingdom,¹ is one of the most frequently used self-rating screening tests for the detection of current non-psychotic mental disorders in primary health care and community settings internationally, including Brazil and other developing countries.²⁻⁵

The origin of the GHQ is associated more with practical than theoretical concerns.^{1,6} The main goal in developing the GHQ was to create an objective, easy to administer, and informative questionnaire regarding the current mental state of a person. As screening test, it was designed to discriminate healthy individuals

from non-psychotic psychiatric patients in general, rather than to diagnose specific disorders. Unlike other self-rating instruments, the GHQ was not created to measure long-standing attributes or the liability of an individual to meet criteria for a diagnostic condition in the future: its focus is, instead, on the present state of mental functioning and its fluctuations.

In conceptual terms, the GHQ is concerned with the "hinterland between psychological sickness and psychological health".⁶ This means that it aims to assess a general factor reflecting overall severity of non-psychotic mental disorder symptoms. This factor

Correspondence

Hudson W. de Carvalho
Núcleo de Estatística e Metodologia Aplicada (NEMAP-UNIFESP)
Rua Dr. Barcelar, 368, conjunto 142 - Vila Clementino
04026-001 São Paulo, SP, Brazil
Email: hdsncarvalho@yahoo.com

Submitted: August 11, 2009
Accepted: July 20, 2010

can be viewed as a continuous dimension along which varying degrees of psychological health versus psychological illness are represented in quantitative terms.

Most factor analytic studies of the 60 item version of the GHQ are based on exploratory factor analysis (EFA) or principal component analysis.⁶ Results are usually consistent with the original proposition,¹ producing one general factor along with four or five specific subfactors.⁵⁻⁷ More sophisticated structural modeling studies utilizing confirmatory factor analysis have been applied to abbreviated versions of the GHQ, most frequently its 12-item version.⁸⁻¹⁰ The structural coherency of the 60-item version remains to be evaluated using confirmatory factor analytic methods.

The goal of the present investigation was to evaluate the unidimensionality and structural coherency of the subscales of the full 60-item GHQ using exploratory and confirmatory factor analysis (CFA) along with a traditional index of internal consistency (coefficient alpha).

Method

1. Study design

The design of this study is a cross-sectional survey. The protocol for the study was approved by the ethics committee of the IPPEX-FACED (Process ETIC n° 001/08) and informed consent was obtained from all respondents prior to their participation.

2. Sample

Required sample size was estimated based on a criterion of at least five respondents per item.^{11,12} Performance diffidence was the largest subscale in terms of items (17 items), and therefore we calculated a minimum *N* of 85 participants (17 x 5). Data were in fact collected from a total of 146 participants. Based on estimations of statistical power,¹¹ and considering this number of participants, we considered factor loadings interpretable only if they exceeded a value of 0.45. This calculation was based on a targeted statistical power of 80%, a type II error (alpha) criterion of 0.05, and observed mean standard errors for factor loadings.

3. Sampling strategy

We sought to maximize the generalizability of our findings by collecting data from a demographically and socioeconomically diverse set of participants recruited from six separate districts in the city of Divinópolis-MG.¹³ Two districts were considered upper class, two others middle class, and the last two were poor communities. In the upper and middle class neighborhoods, participants were selected using a list in which random numbers coded different houses. Individuals from targeted households were contacted by telephone in order to invite their participation, clarify research goals and procedures, and to schedule the questionnaire administration session. A more direct approach was undertaken in the poor communities, in which phone contact information was unlisted for many households. Participants in the two poor communities were approached in-person in their areas of residence by research team members during the day time. Two additional

downtown locations were included: the central bus station and the main city square. These areas were targeted based on their ability to provide representation of individuals from different social and cultural backgrounds.

The 146 individuals who participated in the study consisted of 66 males and 70 females, with a family monthly income ranging from one to 18 minimum salaries ($M = 3.4$, $SD = 2.27$) and age ranging from 18 to 70 years ($M = 28.77$ and $SD = 11.27$). With respect to formal education, 28.3% of the sample had earned a middle school degree, 44.2% a high school degree, and 27% an university or similar degree.

4. Instrument

The GHQ⁵ is a self-rating screening test for the detection of current non-psychotic mental disorders, with primary emphasis on depressive and anxiety disorders. The inventory contains 60 items and yields a general scale score along with scores on five lower-order subscales. Prior research has demonstrated high internal consistency for the five subscales and the general scale as indexed by high Cronbach's alphas and strong loadings of individual items on the dominant factor emerging from an item-level factor analysis of each of the following scales: Psychic stress (13 items; $\alpha = 0.89$, factor loadings of individual items from 0.30 to 0.68), Death ideation (8 items; $\alpha = 0.89$, factor loadings from 0.34 to 0.75), Performance diffidence (17 items; $\alpha = 0.89$, factor loadings from 0.30 to 0.73), Sleep disorders (6 items; $\alpha = 0.80$, factor loadings from 0.38 to 0.75), Psychosomatic disorders (10 items; $\alpha = 0.83$, factor loadings from 0.32 to 0.6), and the general score, named Severity of Mental Disorder (60 items; $\alpha = 0.95$, factor loadings from 0.15 to 0.79).

5. Statistical analysis

Structural analyses were carried out using the Mplus computer package¹⁴ and reliability analyses were performed using the 14th version of SPSS for Windows.

To evaluate the suitability of the dataset for implementation of factor analysis, we examined the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and the Bartlett's Test of Sphericity. The KMO tests whether the partial correlations among variables are robust enough to accommodate factor analysis. It offers an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. A KMO value of 0.6 or higher is considered acceptable for a satisfactory factor analysis to proceed.^{11,12} The Bartlett's test of sphericity tests the null hypothesis that the correlation matrix for the dataset derives from a population in which constituent variables are noncollinear, that is, it evaluates whether correlations among variables are equal to the identity matrix. A p-value of 0.05 or smaller serves as the criterion for rejection of the null hypothesis, indicating that the correlation matrix is not an identity matrix and implementation of factor analysis is possible.^{11,12}

Factor analytic evaluation of the dataset was conducted in two stages. First, to verify the unidimensionality of subscales of the Brazilian GHQ,⁵ internal consistency (coefficient alpha) was

evaluated for each subscale. In addition, separate unweighted least squares EFA models were conducted utilizing the polychoric correlation matrices of items for each subscale. Scree plots and root mean square residual (RMR) values from these models were used as criteria. Second, a scale-level confirmatory factor analysis (CFA) was conducted utilizing the Pearson correlation matrix for the GHQ subscales. More specifically, to evaluate the coherency of the differing subscales as common indicators of a broad psychological health versus maladjustment continuum, a one-factor model was parameterized using maximum likelihood CFA. The Standardized Root Mean Square Residual (SRMR) was used to evaluate absolute model fit. Smaller values of SRMR reflect better fit, with SMRs below a value of 0.08 considered to be indicative of adequate fit, and values below 0.05 indicative of close fit.¹⁵

Results

1. Internal consistency analysis

Cronbach's alpha coefficients indicated adequate levels of homogeneity for all individual scales of the GHQ: Psychic stress, 0.87; Death ideation, 0.87; Performance diffidence; Sleep disorders, 0.83; Psychosomatic disorders = 0.81. The alpha coefficient for the full GHQ item set (60 items), reflecting general severity of mental health problems, was very high (0.95).

2. Exploratory factor analysis

Sample adequacy indicators attested to the suitability of the dataset for implementation of factor analysis. The KMO index achieved adequate levels ($KMO > 0.8$) for all GHQ scales, and the Bartlett's Test of Sphericity was significant in all cases ($p < 0.001$).

EFA and CFA results for four the five GHQ subscales were clearly indicative of unidimensionality, with items corresponding to each factor fully accounted for by one factor when freely estimated ($RMR < 0.05$). For each of these scales, all constituent items exhibited moderate to high loadings on the single extracted factor: Death ideation (8 items), item-factor loadings ranged from 0.58 to 0.77; Performance diffidence, (17 items), 0.41 to 0.71; Sleep disorders (6 items), 0.54 to 0.78; Psychosomatic disorders (10 items), 0.43 to 0.69; In contrast, the EFA for the Psychic stress scale yielded a correlated two-factor structure ($r = 0.63$), with the first factor characterized by items ($n = 10$) involving stress symptomatology content and the second factor by items ($n = 3$) involving (low) self-efficacy content (e.g. *Have you been feeling incapable of overcoming difficulties?*). Robust weighted least squares confirmatory factor analysis of the two-factor model of the Psychic stress scale indicated superior fit ($CFI = 0.943$; $TLI = 0.968$; $WRMR = 0.97$, $SRMR = 0.068$) compared with the unifactorial solution ($CFI = 0.89$; $TLI = 0.94$; $WRMR = 1.19$; $SRMR = 0.081$). To ensure unidimensionality of the Psychic Stress scale in the scale-level model of the GHQ general factor, the three items related to self-efficacy were omitted in the subsequent analyses.

3. Confirmatory factor analysis

Fit index indicated close fit for a one-factor model of the five GHQ scales ($SRMR = 0.038$). Figure 1 depicts this model. The general factor explained 62.35% of the total variance in constituent scales. Psychic stress displayed the strongest loading on the general factor (0.925) and Sleep disorders the weakest (0.64).

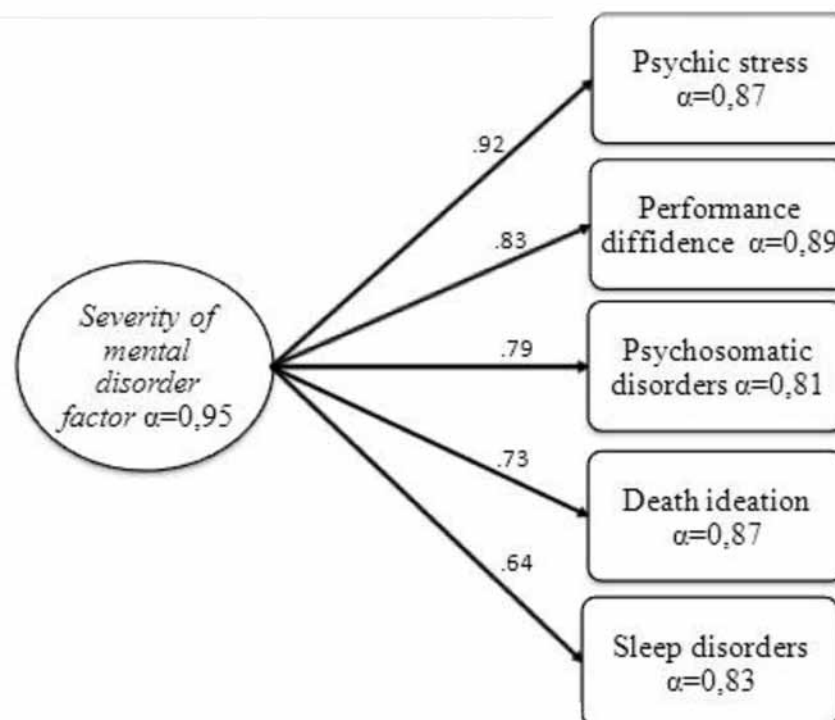


Figure 1 - One-factor confirmatory model of the GHQ depicting factor loadings for each specific subscale of the inventory.

Discussion

The GHQ is a screening test used worldwide for non-psychotic mental illness detection in primary health care and community settings, and in psychological and medical research. Though prior studies of content and structure of the 60-item-GHQ have consistently revealed a general factor, most studies have utilized EFA rather than CFA to evaluate structure. Our investigation sought to fill this gap in the evaluation of the Brazilian version of the GHQ unidimensionality by using both EFA and CFA along with the coefficient alpha index of item homogeneity.

Internal consistency analysis showed similar or identical values to those available in the Brazilian manual⁵ of the GHQ, ranging from good ($\alpha \geq 0.8$) to excellent ($\alpha \geq 0.9$). Item-level EFAs indicated that the subscales of the Brazilian GHQ are unidimensional, with the exception of the Psychic stress subscale, which evidenced better fit when parameterized in terms of two factors, one reflecting distress/depression symptomatology and the other low self-efficacy content. In modeling the scale-level data, the items reflecting self-efficacy were omitted from the Psychic stress subscale to render this scale unidimensional. A CFA model incorporating this modified Psychic stress subscale along with the other four GHQ subscales as indicators of a single general factor provided an excellent fit to the data, confirming the coherency of the GHQ subscales as indicators of a common health/maladjustment dimension.

The present findings have three major implications, two of them practical and the third theoretical in nature. First, according to the International Test Commission Guidelines,¹⁶ all psychological tests should be reviewed periodically to guarantee high quality services to the public. Since the adaptation of the GHQ inventory to Brazil⁵ took place over 10 years ago, it is clearly time for follow-up validation studies, and the current study addresses this need. Second, the current results highlight the need

for re-formulation or exclusion of certain items of the Psychic stress scale to enhance its unidimensionality. Third, the GHQ^{1,6} assumes that the organization of common mental disorders is dimensional rather than categorical, and the current findings for the GHQ provide empirical support for the dimensional approach to psychopathological conceptualization.

Our methodology has strengths and weaknesses. The use of CFA to evaluate the unidimensionality of the GHQ and the use of a random sample of participants recruited from the general community are regarded as strengths. With regard to weaknesses, the sample size (N = 146) was just adequate for estimating factor structure in an item-level procedure and somewhat small for purposes of generalizability. Given this, additional studies employing larger samples sizes are recommended in the future.

Conclusion

The results of the present investigation, when taken together, provide support for the hypothesis that the items and subscales of the GHQ measure a general overarching factor reflecting variations in mental health versus psychological maladjustment. Our results are in agreement with the original GHQ¹ and Brazilian adaptation⁵ studies.

Acknowledgements

We would like to thank Cíntia Teixeira, chairman of the IPPEX-FACED, for the institutional support and the undergraduate students associated with the NuPPSi for the dedication.

Disclosures

Writing group member	Employment	Research grant ¹	Other research grant or medical continuous education ²	Speaker's honoraria	Ownership interest	Consultant/ Advisory board	Other ³
Hudson Wander de Carvalho	UNIFESP	IPPEX-FACED*	-	-	-	-	CAPES*
Christopher J. Patrick	The Florida State University	National Institute of Mental Health**	-	-	-	-	-
Miguel Roberto Jorge	UNIFESP	-	-	-	-	-	-
Sérgio Baxter Andreoli	UNIFESP	CNPq* FAPESP**	-	-	-	-	-

* Modest

** Significant

*** Significant: Amounts given to the author's institution or to a colleague for research in which the author has participation, not directly to the author.

Note: UNIFESP = Universidade Federal de São Paulo; IPPEX-FACED = Instituto de Pós-graduação Pesquisa e Extensão da Faced; CNPq = Conselho nacional de Desenvolvimento Científico e Tecnológico; FAPESP = Fundação de Amparo à Pesquisa do Estado de São Paulo.

For more information, see Instructions for Authors.

References

1. Goldberg D. *The detection of psychiatric illness by questionnaire: a technique for the identification and assessment of non-psychotic psychiatric illness*. London: Oxford University Press; 1972.
2. Lee DT, Yip WC, Chen Y, Meng Q, Kleinman A. Ethno-psychometric evaluation of the General Health Questionnaire in rural China. *Psychol Med*. 2006;36(2):249-55.
3. Patel V, Araya R, Chowdhary N, King M, Kirkwood B, Nayak S, Simon G, Weiss HA. Detecting common mental disorders in primary care in India: a comparison of five screening questionnaires. *Psychol Med*. 2008;38(2):221-8.
4. Shamasundar C, Murthy SK, Prakash OM, Prabhakar N, Krishna DK. Psychiatric morbidity in a general practice in an Indian city. *Br Med J (Clin Res Ed)*. 1986;292(6537):1713-5.
5. Pasquali L, Gouveia VV, Andriola WB, Miranda FJ, Ramos AL. *Questionário de saúde geral de Goldberg: manual técnico*. São Paulo: Casa do Psicólogo; 1996.
6. Goldberg D, Williams P. *A user's guide to the general health questionnaire*. Windsor: NFER-Nelson; 1988.
7. Vazquez-Barquero JL, Williams P, Diez-Manrique JF, Lequerica J, Arenal A. The factor structure of the GHQ-60 in a community sample. *Psychol Med*. 1988;18(1):211-8.
8. Salama-Younes M, Montazeri A, Ismail A, Roncin C. Factor structure and internal consistency of the 12-item General Health Questionnaire (GHQ-12) and the Subjective Vitality Scale (VS), and the relationship between them: a study from France. *Health Qual Life Outcomes*. 2009;5(7):22-8.
9. Mäkikangas A, Feldt T, Kinnunen U, Tolvanen A, Kinnunen ML, Pulkkinen L. The factor structure and factorial invariance of the 12-item General Health Questionnaire (GHQ-12) across time: evidence from two community-based samples. *Psychol Assess*. 2006;18(4):444-51.
10. Gouveia VV, Chaves SS, Oliveira SC, Dias MR, Gouveia RS, Andrade PR. A utilização do QSG-12 na população geral: estudo de sua validade de construto. *Psic Teor Pesq*. 2003;19(3):241-8.
11. Hair JF, Thatam RL, Anderson RE, Black W. *Análise multivariada de dados*. Porto Alegre; 2005.
12. Fabringer LR, Wegener DT, MacCallum RC, Stahan EJ. Evaluating factor analysis in psychological research. *Psychol Methods*. 1999;4(3):279-99.
13. Fernandes LM. Mapa socioeconômico de Divinópolis. Instituto de Pós-Graduação, Extensão e Pesquisa, Faculdade Administração, Ciências Contábeis e Econômicas de Divinópolis [research report]. Divinópolis: Faculdade de Ciências Econômicas Administrativas e Contábeis de Divinópolis; 2008.
14. Muthén B, Muthén L. *Mplus user's guide*. 3rd ed. Los Angeles: Muthén & Muthén; 1998-2004.
15. Pilati R, Laros JA. Modelo de equações estruturais em Psicologia. *Psic Teor Pesq*. 2007;23(2):205-16.
16. International Test Commission. *International Guidelines for Test Use*. Downloaded electronically [cited 2010 Jan 04]. Available from: http://www.intestcom.org/itc_projects.htm; 2000.