

## Factorial validity and reliability of the General Health Questionnaire (GHQ-12) in the Brazilian physician population

Questionário de Saúde Geral (QSG-12) na população médica brasileira: evidências de validade fatorial e consistência interna

Valdiney V. Gouveia <sup>1</sup>  
Genário Alves Barbosa <sup>1</sup>  
Edson de Oliveira Andrade <sup>2</sup>  
Mauro Brandão Carneiro <sup>3</sup>

### Abstract

*The 12-item General Health Questionnaire (GHQ-12) is a widely used screening instrument. One- and two-factor structures have been identified in some countries. In Brazil, the best factor structure is still unclear. This study aimed at knowing its factorial validity and reliability, and testing the one-factor and two-factor models. The participants were 7,512 Brazilian physicians. They answered the GHQ-12 and demographic questions. Unrotated (one-factor) and rotated (two-factor) structures of the GHQ-12 were extracted by principal component analysis. Confirmatory factor analyses (ML) were used to compare the one- and two-factor solutions. The two-factor model fitted the data better than the one-factor one. Those two factors were depression and social dysfunction, and they showed themselves to be directly correlated to one another. They also showed adequate reliability coefficients. The two-factor model is remarkably adequate, showing better fit indices, although it is acceptable to admit a common factor, which could be defined as psychological distress.*

*Mental Health; Physicians; Questionnaire*

### Introduction

There has been a considerable increase in the number of people reporting mental symptoms (such as depression or anxiety) that could be confused with organic problems, and therefore be treated erroneously. Although there are diagnostic tools based on the *Diagnostic and Statistical Manual of Mental Disorders* <sup>1</sup> for instance, the defined criteria for such diagnoses are the presence of symptoms, its continued duration, and the corresponding deficit to psychical functioning. Despite the heuristic and practical nature of these classifications, they take into account the patient's behaviors and his or her complaints. Such information is often unreliable and diffuse. There is thus a clear need for counting with available objective measures to assess exclusively current mental health symptoms <sup>2,3</sup>. The *General Health Questionnaire* (GHQ) was developed by Goldberg in the 1970s to achieve this goal <sup>4</sup>.

The original GHQ is composed of 60 items. However, different shortened versions of this instrument are currently available, according to the number of items (e.g., 30, 28, and 12). The GHQ-12, i.e. the 12-item version, due to its brevity, has probably been the most popular. Searching in *Google Scholar* (<http://scholar.google.com.br/scholar?q=ghq-12&hl-pt-BR&lr=>, accessed on 05/Jul/2008), and introducing GHQ-12 as a keyword, 4,410 papers were identified. This version is used in many countries and languages <sup>3,5,6,7,8,9</sup>.

<sup>1</sup> Universidade Federal da Paraíba, João Pessoa, Brasil.  
<sup>2</sup> Universidade do Estado do Amazonas, Manaus, Brasil.  
<sup>3</sup> Instituto de Pesquisa Clínica Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brasil.

**Correspondence**  
V. V. Gouveia  
Universidade Federal da Paraíba,  
Rua Hortêncio Osterne  
Carneiro 598, João Pessoa,  
PB 58035-120, Brasil.  
vvgouveia@pesquisador.  
cnpq.br

This instrument asks whether the respondent has experienced a particular symptom or behavior recently. Each item is rated on a four-point scale (less than usual, no more than usual, rather more than usual, or much more than usual), using one of two most common scoring methods: dichotomous (0-0-1-1) or Likert type (0-1-2-3).

Considering the GHQ-12 to be a brief, simple and easy to complete instrument, and the fact that its application in research settings as a screening tool is well documented<sup>10,11</sup>, we decided to check its psychometric properties in a sample of Brazilian physicians. In spite of evidences of validity and reliability of this measure in this cultural milieu<sup>12,13,14</sup>, most of them are of an exploratory nature, considering only one state. Moreover, there is not a consensus about the number of factors to extract in the GHQ-12 in Brazil. For instance, Sarriera et al.<sup>13</sup> identified three factors in Rio Grande do Sul, using Principal Components analysis (varimax rotation), and Borges & Argolo<sup>12</sup>, in Rio Grande do Norte, found two factors when using Principal Axis Factoring (oblimin rotation). On the other hand, at least in other countries, some researchers often have compared two-factor and one-factor models by confirmatory factor analysis, concluding that the former has a more adequate fit<sup>2,6,15</sup>. Nevertheless, one and three-factor models are also compared<sup>2,16,17</sup>. Usually, the two-factor solution (depression and social dysfunction) accounts for between 45.3 and 56.5 per cent of the total variance<sup>3,7</sup>, presenting an internal consistency that is close to 0.80<sup>12,13</sup>.

Gouveia et al.<sup>18</sup> tested three factor models (1-, 2-, and 3-factors) by confirmatory factor analyses. They concluded that the most adequate model was the two-factor one, measuring depression and anxiety (social dysfunction), which showed reasonable Cronbach's alpha coefficients (0.81 and 0.66, respectively). However, their sample was specific to João Pessoa, and considered a medium sized city in the Northeast of Brazil. Oliveira<sup>19</sup>, in this same city, took into account a sample of 246 health professionals, including 98 psychologists, 81 physicians and 67 nurses, all of whom answered the GHQ-12. She performed only exploratory factor analysis, without comparing different factor models for this measure. Two factors were observed, explaining 51.5% of the total variance, with alphas of 0.83 (psychological distress) and 0.76 (lack of self-efficacy). There is no information about the fitness of this model to data.

These considerations motivated the current study. Its objectives were, therefore, three-fold. Responding to the recommendation to consider a large sample<sup>2</sup>, it aimed at (1) knowing the fac-

tor structure underlying the 12-item of the GHQ by performing an exploratory factor analysis; (2) testing the two most common factor models to explain the data obtained by this measure, as discussed in the literature (one- and two-factor models); and (3) knowing evidences of its homogeneity and reliability. In sum, this study searched evidences of factorial validity and reliability of the GHQ-12 in a large physician sample from all 26 Brazilian states and Federal District. Physicians demand mental health attention because they are a professional group often working in stressing labor context, and experiencing many mental illness symptoms<sup>19,20</sup>.

## Method

### Participants

A national mail survey was carried from December 2005 to August 2006. Taking into account the Brazilian physician population at that time (281,939), we randomly selected 67,468 of them to whom were sent the questionnaires. The response rate was of 11.8% (7,700), which was consistent with previous studies in this cultural milieu<sup>21</sup>. Participating effectively in this study were 7,512 physicians, who answered all 12 items of the GHQ. Most of them were male (63.1%), married (75.7%), and had children (78.1%). Their mean age was 47.2 years old (standard deviations – SD = 11.28, ranging from 24 to 93; 95.1% under 65 years-old). The detailed method is described elsewhere<sup>22</sup>.

### Instrument

All participants answered a questionnaire comprised of different psychological measures (e.g., fatigue, suicidal ideation), and demographic questions. The Brazilian-Portuguese version of the GHQ-12<sup>18</sup> was also included. The Likert type answer scale was adopted, and is described above. Psychometric properties of this measure in Brazil and other cultures were previously detailed. It is only this measure which receives attention in this article.

### Procedure

The dataset of Brazilian physicians was requested from the Federal Council of Medicine, Brazil. Taking into account this dataset, potential participants for this study were selected. To each selected physician, a questionnaire was sent on one double-sided sheet, and his/her voluntary and anonymous participation in the study was

requested. All participants were informed that filling and returning the questionnaire was considered as acceptance of the term of free and informed consent.

### **Statistical analyses**

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 0.40 range were deemed to indicate good reliability<sup>23</sup>. Exploratory factor analysis was performed using principal components, and confirmatory factor analysis was carried using Analysis of Moments Structures 16<sup>th</sup> revision (AMOS) and maximum-likelihood estimation procedures, taking the observed covariance matrix as the input. This procedure has been used in previous studies<sup>2</sup>. The degree to which the data fit the confirmatory models was assessed using the adjusted goodness-of-fit-index (AGFI), the comparative fit index (CFI), and the root mean square error of approximation (RMSEA). Models with AGFI and CFI values close to .90 or higher, and RMSEA of .08 or lower indicate acceptable fit<sup>24,25</sup>.

The alternative factor models of the GHQ-12 were assessed with respect to three fit indices. Specifically, the  $\chi^2$  difference test ( $\Delta\chi^2$ ), the expected cross-validation index (ECVI) and the consistent Akaike information criterion (CAIC) were used to calculate improvements over competing models. Significant results for the  $\chi^2$  difference test in favor of lower value, and lower ECVI and CAIC values reflect better fit<sup>26</sup>.

## **Results**

### **Descriptive statistics**

Table 1 presents the means (*m*), SD, and correlations between items of the GHQ-12. The items' means ranging from 1.23 (Item 11: *Been thinking of yourself as a worthless person?*) to 2.38 (Item 7: *Been able to enjoy your normal day-to-day activities?*). The mean inter-item correlation of the total set of items was 0.42 ( $p < 0.001$ ), ranging from 0.22 ( $p < 0.001$ ; items 2 and 4) to 0.65 ( $p < 0.001$ ; items 5 and 9).

### **Factor structure**

Initially, before performing the factor analysis, the adequacy of the correlation matrix of the 12-item GHQ was checked (Table 1). The observed

values supported this type of statistical analysis:  $KMO = 0.93$  and Bartlett's Sphericity Test,  $\chi^2 (66) = 38,705.09$ ,  $p < 0.001$ . Thus, it sounds appropriate to conduct the exploratory factor analysis. In this case, two steps followed: first, the unrotated solutions were produced, and then the varimax rotation, admitting orthogonal factors. Independently, it was adopted a strict cutoff of factor loading of  $> 0.50$ , used by other researchers<sup>27</sup>. Results are described below.

### **Unrotated factor structure**

Factor analysis was carried using principal component (PC) analysis, in line with previous research<sup>3</sup>. This analysis, without fixing the number of factors to extract, allowed identifying two factors with eigenvalue (Kaiser's criterion) greater than 1 (5.67 and 1.18), conjointly accounting for 57.1% of the total variance. This solution clearly produced a general unipolar factor, all items with positive loadings  $> 0.50$ , ranging from 0.57 (Item 3: *Felt that you are playing a useful part in things?*) to 0.81 (Item 9: *Been feeling unhappy and depressed?*). The second factor aggregated items with lower factor loadings; only item 3 attained the factor loading cutoff. In this way, this factor was discarded at this moment.

The number of factors to extract was based on three criteria: the eigenvalue greater than 1 (Kaiser), the scree test (Cattell), and the parallel analysis (Horn). According to the first two criteria, the extraction of two factors seems evident. The final decision of extracting these factors was obtained by parallel analysis, where only two first observed eigenvalues were higher than those that would be obtained from 1,000 replications of random data with the same number of items and the same sample size<sup>28</sup>.

### **Varimax rotation structure**

Fixing the criterion to extract two factors with varimax rotation, trying to identify a simple structure, the PC analysis reveals a clear solution. According to the first column of Table 2, the 12 items were equally distributed into two principal factors. The eigenvalues after the rotation were 3.82 and 3.03, explaining conjointly 57% of the total variance. The cutoff to define the item as representing the factor was factor loading  $> 0.50$ . The items loadings on the first factor (e.g., constantly under strain, lost sleep over worry, and unhappy or depressed) seem to evince the construct *depression*, meanwhile those loadings on the second factor (e.g., play useful part in things, capable of making decisions, and thinking of self as worthless) express the *social dysfunction*

Table 1

Means (*m*), standard deviations (SD), and intercorrelations for the 12-item *General Health Questionnaire* (GHQ).

Item	01	02	03	04	05	06	07	08	09	10	11	12
01												
02	0.41											
03	0.30	0.23										
04	0.33	0.22	0.46									
05	0.48	0.54	0.28	0.28								
06	0.42	0.46	0.35	0.36	0.60							
07	0.45	0.40	0.26	0.28	0.51	0.46						
08	0.41	0.36	0.37	0.47	0.45	0.50	0.50					
09	0.46	0.51	0.36	0.34	0.65	0.61	0.50	0.51				
10	0.39	0.39	0.39	0.39	0.48	0.57	0.38	0.49	0.61			
11	0.28	0.27	0.39	0.32	0.34	0.42	0.28	0.39	0.42	0.57		
12	0.38	0.33	0.46	0.42	0.41	0.44	0.43	0.50	0.54	0.45	0.42	
<i>m</i> (SD)	2.30 (0.623)	2.11 (0.947)	1.84 (0.644)	1.91 (0.585)	2.36 (0.937)	1.91 (0.880)	2.38 (0.711)	2.13 (0.618)	2.05 (0.971)	1.56 (0.806)	1.23 (0.594)	1.98 (0.678)

Note: all Pearson's *r* are significant ( $p < 0.001$ ).

Table 2

Exploratory and confirmatory factorial structure of the 12-item *General Health Questionnaire* (GHQ).

Item	Exploratory factor analysis			Confirmatory factor analysis		
	I	II	Uniqueness (1 - $h^2$ )	I	II	Error variance
05	0.82	0.18	0.30	0.78		0.35
02	0.75	0.07	0.43	0.62		0.55
09	0.73	0.38	0.32	0.82		0.31
07	0.68	0.22	0.38	0.64		0.30
06	0.68	0.38	0.40	0.76		0.33
01	0.62	0.27	0.35	0.61		0.25
03	0.10	0.77	0.40		0.57	0.28
04	0.13	0.74	0.43		0.58	0.23
11	0.26	0.64	0.52		0.62	0.22
12	0.39	0.63	0.45		0.69	0.24
10	0.50	0.57	0.43		0.74	0.29
08	0.49	0.55	0.46		0.71	0.19
$\Phi = 0.15$						

construct. These two factors were positively correlated to each other ( $r = 0.69$ ,  $p < 0.001$ ).

#### Testing uni- and two-factor models

Previously, the uni- and multivariate distribution of items was checked. Taking into account the absolute values of the univariate skewness

and kurtosis, all items do not deviate from the normality (absolute values below 1.30)<sup>29</sup>, with the exception of item 11 (values of 2.88 and 8.05, respectively). The multivariate kurtosis was 46.69, indicating non-normality. Therefore, the data show moderate non-normality, which suggests adopting the ADF (asymptotic distribution free) estimation. However, there is evidence that

for large samples the ML (maximum likelihood) estimation is more adequate<sup>30</sup>, and that on this condition the sampling error's impact is minimized<sup>31</sup>. In line with these recommendations, the ML estimation was adopted.

Considering the possibility of two factor solutions, i.e. one-factor and two-factor models, according to the literature, we decided to know their fit to data and compare them to each other. In the first model ( $M_1$ ), all 12 items were defined to load on only one factor; in the second model ( $M_2$ ), the items were established loading on two factors ( $\Phi = 0.15$ ), according to the extracted solution in previous PC analysis. In both models, all items loadings were statistically different from zero ( $z > 1.96$ ,  $p < 0.05$ ). Fit indices were AGFI = 0.84, CFI = 0.88, and RMSEA = 0.11 (confidence interval 90% – CI90%: 0.103-0.108) for  $M_1$ , and AGFI = 0.90, CFI = 0.92, and RMSEA = 0.088 (CI90%: 0.086-0.091) for  $M_2$ . Comparing these nested models, the latter revealed to be more adequate than the former [ $\Delta\chi^2(1) = 1,431.93$ ,  $p < 0.001$ ]. The respective values of CAIC and ECVI for  $M_1$  (4,818.69 and 0.616) and  $M_2$  (3,396.68 and 0.426) support this finding.

### **Homogeneity and reliability**

As previously observed, the 12 items of the GHQ showed homogeneity. Its mean inter-item correlation was above 0.40. This solution, with all items loading on only one factor, showed Cronbach's alpha of 0.89. With respect to the two-factor model, the homogeneities (mean inter-item correlation) were 0.50 ( $r = 0.40-0.65$ ,  $p < 0.001$ ) and .43 ( $r = 0.32-0.57$ ,  $p < 0.001$ ) for first (*depression*) and second (*social dysfunction*) factors, respectively (Table 1). Their reliabilities (Cronbach's alpha) were 0.85 and 0.82, respectively.

### **Discussion**

The current study aimed at knowing evidences of factorial validity and reliability of the GHQ-12 in the Brazilian physician population. It considered a large sample, corresponding to approximately 3% of these professionals in Brazil, including participants of all its 26 states and the Federal District. Despite this effort, it is important to observe that the aim was not to generalize the findings to the whole country. This study had a psychometric nature, assessing measure parameters of a specific instrument. Perhaps its main limitation was not testing the impact of different demographic variables (state, gender) on the factorial structure of the GHQ-12. However, this was not its objective. One methodological aspect

might demand attention in the future: the use of ML estimation with non-normality item distribution. In line with Ory & Mokhtarian<sup>30</sup>, it could be interesting to explore different procedures of estimation (e.g., ML, ADF, bootstrapping).

As previously mentioned, only one study was found in Brazil in which the factor structure of the GHQ-12 was assessed among physician participants<sup>19</sup>. However, that study takes into account a small sample, performing an exploratory factor analysis. Therefore, our study improves it, running exploratory and confirmatory factor analyses, testing two common factor models for this measure (i.e., uni- and bi-factor<sup>5,7,15,18</sup>), and presenting information about homogeneity and reliability of the corresponding factors.

The findings reported in the current study support the psychometric appropriateness of the GHQ-12. For instance, our findings were consistent with those described by Werneke et al.<sup>3</sup> When an unrotated solution was defined, a one-factor model seemed most adequate, which could be named as psychological distress. However, fixing the varimax rotation, it was possible to find two factors, which explained close to 60% of the total variance. Although these factors were named in a different way in previous studies in Brazil<sup>12,18</sup>, most items comprising each factor observed in this study clearly reproduce the observed most common two-factor model (*depression* and *social dysfunction*) identified in other studies<sup>2,3</sup>. In line with previous findings, in this study the two-factor model was better than the one-factor model<sup>15,18</sup>. Overall, the goodness-of-fit indexes for the two-factor one were acceptable<sup>24,25</sup>, and coherent with those observed in the literature<sup>2</sup>. Finally, the homogeneity and reliability were higher than the cutoff recommended<sup>23</sup>. For instance, Cronbach's alpha was always higher than 0.80.

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, regional culture) 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. In this case, it could take into account as gold standard the classification of formal psychiatric diagnoses of depression or anxiety experienced by physicians.

## Resumo

O Questionário de Saúde Geral de 12 Itens (QSG-12) é um instrumento de triagem amplamente usado. As estruturas com um e dois fatores têm sido observadas em alguns países. No Brasil não é ainda clara a melhor estrutura fatorial. Este estudo objetivou conhecer evidências de sua validade fatorial e consistência interna. Os participantes deste estudo foram 7.512 médicos brasileiros, que responderam o QSG-12 e perguntas demográficas. Foram extraídas estruturas fatoriais rotadas (unifatorial) e não-rotadas (bifatorial) por meio de análise de componentes principais. Realizaram-se análises fatoriais confirmatórias (ML) para comparar as soluções uni e bi-fatorial. O modelo com dois fatores se ajustou melhor aos dados do que o unifatorial. Os dois fatores foram depressão e disfunção social, sendo diretamente correlacionados entre si; ambos apresentaram coeficientes de confiabilidade aceitáveis. O modelo bifatorial foi claramente adequado, apresentando os melhores indicadores de ajuste, embora possa ser aceito um fator comum, concebido como desconforto psicológico.

Saúde Mental; Médicos; Questionário

## Contributors

V.V. Gouveia was responsible for designing the research, data analysis and writing the article. G. A. Barbosa participated in the discussion about the data used in the article. E. O. Andrade made a significant contribution towards the conception and scope of the study. M. B. Carneiro participated in the analysis and interpretation of the data.

## Acknowledgements

This study was funded by the CFM (Federal Council of Medicine). The first author was also supported in part by a grant from the CNPq (National Research Council). We are grateful to these two Brazilian institutions, and to Valeschka M. Guerra for her helpful comments on an earlier draft of this article.

## References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV). 4<sup>th</sup> Ed. Arlington: American Psychiatric Association; 1994.
2. Gao F, Luo N, Thumboo J, Fones C, Li S-C, Cheung Y-B. Does the 12-item General Health Questionnaire contain multiple factors and do we need them? Health Qual Life Outcomes 2004; 2:63.
3. Werneke U, Goldberg DP, Yalcin I, Üstün BT. The stability of the factor structure of the General Health Questionnaire. Psychol Med 2000; 30:823-9.
4. Goldberg DP. The detection of psychiatric illness by questionnaire. London: Oxford University Press; 1972.
5. González-Romá V, Lloret S, Espejo B. Comparación de los modelos de medida del Cuestionario de Salud General (GHQ-12). Psicológica 1993; 14: 259-68.
6. Goldberg DP, Gater R, Satorius N, Üstün TB, Piccinelli M, Gureje O, et al. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol Med 1997; 27:191-7.
7. Doi Y, Minowa M. Factor structure of the 12-item General Health Questionnaire in the Japanese general adult population. Psychiatry Clin Neurosci 2003; 57:379-83.

8. Montazeri A, Harirchi AM, Shariati M, Garmaroudi G, Ebadi M, Fateh A. The 12-item General Health Questionnaire (GHQ-12): translation and validation study of the Iranian version. *Health Qual Life Outcomes* 2003; 1:66.
9. Hu Y, Stewart-Brown S, Liz T, Weich S. Can the 12-item General Health Questionnaire be used to measure positive mental health? *Psychol Med* 2007; 37:1005-13.
10. Patel V, Pereira J, Mann AH. Somatic and psychological models of common mental disorder in primary care in India. *Psychol Med* 1998; 28:135-43.
11. Piccinelli M, Simon G. Gender and cross-cultural differences in somatic symptoms associated with emotional distress. An international study in primary care. *Psychol Med* 1997; 27:433-44.
12. Borges LO, Argolo JCT. Adaptação e validação de uma escala de bem-estar psicológico para uso em estudos ocupacionais. *Aval Psicol* 2002; 1:17-27.
13. Sarriera JC, Schwarcz C, Câmara SG. Bem-estar psicológico: análise fatorial da escala de Goldberg (GHQ-12) numa amostra de jovens. *Psicol Reflex Crít* 1996; 9:293-306.
14. Wagner A, Ribeiro LS, Arteché AX, Bornholdt EA. Configuração familiar e o bem-estar psicológico dos adolescentes. *Psicol Reflex Crít* 1999; 12: 147-56.
15. Kalliath T, O'Driscoll MP, Brough P. A confirmatory factor analysis of the General Health Questionnaire-12. *Stress Health* 2004; 20:11-20.
16. Daradkeh TK, Ghubash R, El-Rufaie OE. Reliability, validity, and factor structure of the Arabic version of the 12-item General Health Questionnaire. *Psychol Rep* 2001; 89:85-94.
17. Makikangas A, Feldt T, Kinnunen U, Tolvanen A, Kinnunen M-A, Pulkkinen L. The factor structure and factor invariance of the 12-item General Health Questionnaire (GHQ-12) across time: evidence from two community-based samples. *Psychol Assess* 2006; 18:444-51.
18. Gouveia VV, Chaves SSS, Oliveira ICP, Dias MR, Gouveia RSV, Andrade PR. A utilização do QSG-12 na população geral: estudo de sua validade de construto. *Psicol Teor Pesqui* 2003; 19:241-8.
19. Oliveira GF. Trabalho e bem-estar subjetivo: compreendendo a situação laboral dos médicos [Tese de Doutorado]. João Pessoa: Universidade Federal da Paraíba; 2008.
20. Miller N, McGowen R. The painful truth: physicians are not invincible. *South Med J* 2000; 93: 966-73.
21. Gouveia VV, Günther H. Taxa de resposta em levantamento de dados pelo correio: o efeito de quatro variáveis. *Psicol Teor Pesqui* 1995; 11:163-8.
22. Barbosa GA, Andrade EO, Carneiro MB, Gouveia VV. A saúde dos médicos no Brasil. Brasília: Conselho Federal de Medicina; 2007.
23. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychol Assess* 1995; 7:309-19.
24. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, editors. *Testing structural equation models*. Newbury Park: Sage Publications; 1993. p. 136-62.
25. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 1999; 6:1-55.
26. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods* 1996; 1:130-49.
27. Peterson RA. A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters* 2000; 11:261-75.
28. Hayton JC, Allen DG, Scarpello V. Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organizational Research Methods* 2004; 7:191-205.
29. Brown TA. *Confirmatory factor analysis for applied research*. New York: The Guilford Press; 2006.
30. Ory DT, Mokhtarian P. The impact of non-normality, sample size and estimation technique on goodness-of-fit measures in structural equation modeling: evidence from ten empirical models of travel behavior. *Qual Quant* 2009; 44:427-45.
31. Hair Jr. JF, Anderson RE, Tatham RL, Black WC. *Multivariate data analysis*. 5<sup>th</sup> Ed. Delhi: Pearson Education; 1998.

---

Submitted on 28/May/2009

Final version resubmitted on 19/Feb/2010

Approved on 20/Apr/2010