

The General Health Questionnaire (GHQ)

Comparison of the English version and a translated Indian version

T. G. Sriram¹, C. R. Chandrashekar¹, M. K. Isaac¹, and V. Shanmugham²

¹ Department of Psychiatry and ² Department of Biostatistics, National Institute of Mental Health and Neurosciences, Bangalore, India

Summary. The English version of the General Health Questionnaire (GHQ) and a translated Indian version were administered to a sample of hundred bilingual college students. Both the versions of the GHQ showed adequate internal consistency and reliability. There was a high concordance between the two versions on high scorers and low scorers. These results were also applicable to GHQ-30 and GHQ-12. Item analysis revealed certain differences between the two versions attributable to semantic and technical problems. Validity of the GHQ in the Indian setting are discussed.

The General Health Questionnaire (GHQ) devised by Goldberg (1972) has been widely used to detect nonpsychotic psychiatric disturbances in a variety of settings. In its full form and in its shorter versions, the GHQ has been used in community surveys (Duncan-Jones and Henderson, 1978; Tarnapolsky et al. 1979; Benjamin et al. 1982; Banks 1983), general practice (Goldberg 1972; Fontanesi et al. 1985; Shamasundar et al. 1986; Boardman 1987), hospital outpatients (Ballinger 1977; Byrne 1984; Vazquez-Barquero et al. 1985; Lobo et al. 1986) and hospital in-patients (De Paulo et al. 1980; Bridges and Goldberg 1986). A number of studies have also examined the factor structure of the GHQ which have been reviewed by Vieweg and Hedlund (1983). The questionnaire has been translated to not less than 36 languages (Goldberg and Williams 1988).

In India the GHQ has been translated into Hindi (Gautam et al. 1987) and Bengali (Bandopadhyay et al. 1988). The 12-item version has also been used successfully as the first stage screening tool in an investigation of psychiatric morbidity in Bangalore City, India (Shamasundar et al. 1986).

Studies examining the validity of GHQ in languages other than English have shown that the validity coefficients are comparable to those of the

English version (Goldberg and Williams 1988). Attempts have also been made to examine the comparability of the GHQ in its original and translated version at the item level. In a recent study, Chan (1985), from China examined the effect of language on GHQ responses. In this study, it was found that the English and Chinese versions of the GHQ were comparable at the scale and dimension level, but at the item level certain differences emerged between the two versions, attributable to linguistic and cultural factors. No such investigation has been carried out in an Indian setting. Additionally the psychometric adequacy of the English version of the GHQ has not so far been investigated in India. This study, therefore, aimed to examine the performance of the English version and a translated Indian version of the GHQ, using a methodology similar to the one employed by Chan (1985). It was hypothesized that discrepancies would emerge between the two versions at the item level similar to that noted in the study of Chan (1985).

Method

The GHQ was first translated to Kannada, the regional language of the state of Karnataka in South India, by the process of translation-back translation. One of the authors (C. R. C) who is proficient in both the languages and who was also involved in the translation of the Crown-Crisp Experiential Index (CCEI) into Kannada (Sriram et al. 1987) first translated the GHQ into Kannada. Emphasis was given to conceptual equivalence rather than to literal translation of the items. Another bilingual school teacher back-translated the questionnaire into English. Based on discrepancies arising in the back translated version the translated version was re-examined and the final Kannada version of the GHQ prepared.

The study sample comprised of 100 female bilingual graduate students in the age range of

18–19 years from a women's college in Bangalore city. After an explanation of the purpose of the study and obtaining informed consent, the GHQ was administered on two occasions separated by a period of one week. On the first occasion, 50 of the students received the English version and the other 50 received the Kannada version. On the second occasion, the order of administration was reversed. The circumstances at the time of administration was similar on both occasions. Thus there were 100 English and 100 Kannada GHQ responses for the same group of subjects.

Results

Using the GHQ scoring method (0-0-1-1), endorsement frequencies (i.e. the proportion of subjects endorsing an item in the pathological direction) were computed for both the versions. The endorsement frequencies ranged from 0.03 to 0.23 for the English version, and 0.02–0.27 for the Kannada version. Five of the 60 items had discrepancies in endorsement frequencies greater than 0.10 (Table 1). For 37 items the endorsement frequencies were higher for the English version, for 19 items they were higher for the Kannada version, while for 4 items, the endorsement frequencies were similar for both the versions.

Table 1. Items with discrepant endorsement frequencies between the Kannada and English versions of the GHQ

Item	Endorsement frequency	
	English	Kannada
3. Been feeling run down and out of sorts	0.07	0.20
12. Been getting up feeling your sleep has not refreshed you	0.16	0.04
30. Been satisfied with the way you have carried out your task	0.18	0.08
41. Been finding like a struggle all the time	0.13	0.27
42. Been able to enjoy your normal day to day activities	0.07	0.18

Table 2. GHQ items with non significant item-item correlations between the English and the Kannada versions

Item	Correlation coefficient
21. Been managing to keep yourself busy and occupied	0.09
27. Been managing as well as most people would in your shoes	0.05
33. Spent much time chatting with people	0.18
46. Been able to face up to your problems	0.18

Table 3. Internal consistency and split-half reliability of the English and Kannada versions of the GHQ

Internal consistency	Alpha coefficient	
	English	Kannada
GHQ-60	0.96	0.96
GHQ-30	0.95	0.92
GHQ-12	0.86	0.84
Split-half reliability	Reliability coefficient	
	English	Kannada
GHQ-60	0.92	0.93
GHQ-30	0.90	0.86
GHQ-12	0.76	0.72

All the coefficients shown are significant at $P < 0.001$

Using the Likert scoring method (0-1-2-3) item-item correlations were computed between the two versions for all the hundred subjects. The correlations ranged from 0.05–0.69. For fifty two items the correlations were significant at $P < 0.01$, while for four, they were significant at $P < 0.05$. Table 2 shows the items with non-significant ($P > 0.05$) item-item correlations.

Of the nine items with discrepant endorsement frequencies and non-significant item-item correlations shown in Table 1 and Table 2, seven items also appear in GHQ-30 and two items in GHQ-12. The correlations between the total scores of GHQ-60, GHQ-30 and GHQ-12 were examined for the English and Kannada versions. The resulting correlation coefficients were 0.84, 0.81 and 0.76 (all significant at $P < 0.001$) respectively. These results also reflect the test-retest reliability of the GHQ since the two versions were administered at different time periods. Reliability coefficients were also computed separately for the English-Kannada ($n = 50$) and Kannada-English ($n = 50$) test-retest administrations. The correlations were 0.88 ($P < 0.001$) for both the administrations. The internal consistency (alpha coefficient) and split half reliability of the English and Kannada versions of the GHQ-60, GHQ-30 and GHQ-12 were examined separately using Likert scoring method. For computing the split-half reliability, alternate items were chosen, rather than splitting the questionnaire into first half and second half since the GHQ items are arranged such that the initial questions mainly refer to general physical symptoms while the items pertaining to the psychological sphere, especially depression, appear in the latter part. As can be noted from Table 3, both the internal consistency and reliability coefficients were comparable for the two language versions.

The mean GHQ scores were similar for both the versions regardless of whether the Likert scoring method or the GHQ scoring method was used (Table 4).

Table 4. Mean scores (\pm SD) of English and Kannada versions of the GHQ-60 ($n = 100$)

	Mean	SD	Difference
Likert scoring method			
English	41.93	19.72	$t = 1.33$ $df = 98$
Kannada	38.30	18.80	NS
GHQ scoring method			
English	6.59	8.37	$t = 0.37$ $df = 98$
Kannada	6.14	8.33	NS

NS = not significant

Table 5. Two-way classification of high scorers and low scorers by the English and Kannada versions of the GHQ

		GHQ-60 English					
		Cut-off score 10/11		Cut-off score 11/12		Cut-off score 12/13	
		LS	HS	LS	HS	LS	HS
GHQ-60	LS	68	10	73	8	77	7
Kannada	HS	10	12	8	11	6	10
Concordance rate (%)		80%		84%		87%	
		GHQ-30 English					
		Cut-off score 3/4		Cut-off score 4/5		Cut-off score 5/6	
		LS	HS	LS	HS	LS	HS
GHQ-30	LS	57	13	67	7	68	7
Kannada	HS	11	19	8	18	10	15
Concordance rate (%)		76%		85%		83%	
		GHQ-12 English					
		Cut-off score 1/2		Cut-off score 2/3		Cut-off score 3/4	
		LS	HS	LS	HS	LS	HS
GHQ-12	LS	61	15	71	10	83	4
Kannada	HS	8	16	6	13	2	11
Concordance rate (%)		77%		84%		94%	

Using different cut-off scores, the high scorers and low scorers were identified for the English and Kannada versions (Table 5). Concordance rates between the two versions were high and tended to increase with higher cut-off scores.

Discussion

The findings of the present study show results which are very similar to those of Chan (1985). At the scale level, the GHQ has demonstrated adequate psychometric characteristics in both its original English version and the translated Kannada version. It has good internal consistency, the alpha coefficients being similar to those reported by Chan (1985). The

split-half reliability coefficients are comparable to those reported by Goldberg (1972). Version equivalence at the scale level is demonstrated by the high positive-correlations between the two language versions. The concordance rates on high scorers and low scorers obtained in the present study are similar to those reported in the Chinese study.

These results also apply to the shorter 30-item and 12-item versions as well. Item analysis of the GHQ, however, reveals certain discrepancies between the two versions as evident from the nine items with non significant item-item correlations and discrepant endorsement frequencies. The results merit a detailed discussion of the issues related to the validity of GHQ in different cultural settings.

Kleinman (1987) in his recent article has discussed a number of issues related to translation of instruments to assess psychopathology in cross-cultural research. Quoting a study conducted in Peru by Gaviria et al. (1984) in which the Diagnostic Interview Schedule (DIS) was translated into the local language, Kleinman (1987) highlights the aspects of validity involved in such investigations. These include semantic, technical, content, conceptual and criterion validity. Though GHQ is only a screening instrument, a number of these issues are relevant to it.

Semantic validity of an instrument depends chiefly on how meticulously the process of translation was carried out. However it may simply not be possible to obtain precise semantic equivalence to certain words and phrases in a different language. In the present investigation, it was only possible to obtain approximate equivalence (shown in brackets) to the following words: nervous (anxious), spell (short period), task (work), panic (intense anxiety). Similar difficulties were encountered while translating phrases like 'out of sorts', 'personal appearance', 'strung up' etc. These findings support the hypothesis put forth by Leff (1977) that language is an important factor influencing emotional expression. It is however interesting to note that though there were several items in which difficulties were encountered with semantic equivalence, only some of these items showed poor item-item correlations and discrepant endorsement frequencies. The problem appears to have been mitigated by the presence of synonymous expressions in many of the GHQ items like 'nervous and strung up', 'edgy and bad tempered', 'warmth and affection' etc. This highlights the suggestions of Sartorius and Brooke (1973) and Kleinman (1987) on the need to include culture-appropriate synonymous expressions so that the meaning of the items can be properly conveyed.

Problems related to difficulty in understanding the items (technical validity) might have further contributed to discrepancies between the two versions at

the item level. For example items like 'been able to manage as well as most people would in your shoes' may not be readily understood. The sample in the present investigation was a literate one, and this does not appear to have been a major problem. However, it remains to be seen to what extent the GHQ items would be adequately understood by illiterate rural respondents. In such situations it is preferable to use a questionnaire like Self-Reporting Questionnaire (SRQ) (Sen 1987) which has a simple response style, or alternatively, the shorter version of GHQ like the GHQ-12 which is relatively easily understood.

While criterion validity cannot be commented upon on the basis of this investigation, content and conceptual validity merit some discussion. In so far as non-psychotic psychiatric disturbances are universal experiences, the GHQ can be said to possess relevance across cultures. However, certain forms of neurotic disturbances like hysterical conversion reactions, which are still common in the developing countries may not be adequately picked up by the GHQ, as the items relevant to such disorders have not been covered adequately in the questionnaire. Similarly, patients with isolated somatic symptoms as a manifestation of their psychological distress may also be missed since they are unlikely to score above the threshold levels. While these issues need to be examined in future systematic research, investigators using GHQ in developing countries need to be aware of these possible limitations.

In conclusion, the results of the present study establish the psychometric adequacy of the English and Kannada versions of the GHQ. However, because of the problems related to technical, content and conceptual validity, the use of GHQ, especially in its long version, in countries like India may be limited to certain populations like urban literate respondents.

Acknowledgements. The authors wish to express their grateful thanks to the following persons: Smt. M. C. Pankaja, for doing the back translation; Prof. R. S. Murthy, Head of the Department of Psychiatry for support and assistance; Prof. S. M. Channabasavanna, Dean of NIMHANS and Prof. G. N. Narayana Reddy, Director of NIMHANS, for their kind permission to carry out the study; Smt. M. P. Lakshmi for typing the manuscript.

References

- Ballinger B (1977) Psychiatric morbidity and the menopause: survey of a gynaecological out-patient clinic. *Br J Psychiatry* 131: 83-89
- Bandopadhyay G, Sinha S, Sen B, Sen G (1988) Validity of the General Health Questionnaire (GHQ-36/GHQ-12) in the psychiatric OPD of a general hospital - a pilot study. *Int J Soc Psychiatry* 34: 130-134
- Banks MH (1983) Validation of the General Health Questionnaire in a young community sample. *Psychol Med* 13: 349-353
- Benjamin S, Decalmer P, Haren D (1982) Community screening for mental illness. A validity study of the General Health Questionnaire. *Br J Psychiatry* 140: 174-180
- Boardman AP (1987) The General Health Questionnaire and the director or emotional disorder by general practitioners: a replicated study. *Br J Psychiatry* 151: 373-381
- Bridges KW, Goldberg DP (1986) The validation of the GHQ-28 and the use of the MMSE in neurological in-patients. *Br J Psychiatry* 148: 548-553
- Byrne P (1984) Psychiatric morbidity in a gynaecological survey. *Br J Psychiatry* 144: 28-34
- Chan DW (1985) The Chinese version of the General Health Questionnaire: does language make a difference? *Psychol Med* 15: 147-155
- Duncan-Jones P, Henderson S (1978) The use of a two-phase design in a prevalence survey. *Soc Psychiatry* 13: 231-237
- Fontanesi F, Gobetti C, Zimmermann-Tansella CH, Tansella M (1985) Validation of the Italian version of the GHQ in general practice. *Psychol Med* 15: 411-415
- Gautam S, Nijhawan M, Preet Kamal (1987) Standardisation of the Hindi version of the Goldberg's General Health Questionnaire. *Indian J Psychiatry* 29: 63-66
- Gaviria M, Pathak D, Flaherty J, Garcia-Pacheco C, Martinez H, Wintrob R, Mitchell T (1984) Designing and adapting instruments for a cross-cultural study of immigration and mental health in Peru. Paper presented at the American Psychiatric Association Meeting
- Goldberg DP (1972) The detection of psychiatric illness by questionnaire. Mansley Monograph No. 21, University Press, London Oxford
- Goldberg D, Williams P (1988) A user's guide to the General Health Questionnaire. NFER, Windsor
- Kleinman A (1987) Anthropology and psychiatry. The role of culture in cross-cultural research on illness. *Br J Psychiatry* 151: 447-457
- Leff JP (1977) The cross-cultural study of emotions. *Cult Med Psychiatry* 1: 317-350
- Lobo A, Perez-Echeverria MJ, Artal J (1986) Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. *Psychol Med* 16: 135-140
- Paulo JR de, Folskin MF, Gordon B (1980) Psychiatric screening on a neurological ward. *Psychol Med* 10: 125-132
- Sartorius N, Brook EM (1973) Translation. In: The international pilot study of schizophrenia, Geneva, WHO
- Sen B (1987) Psychiatric phenomena in primary health care - their extent and nature. *Indian J Psychiatry* 29: 33-40
- Shamasundar C, Krishna Murthy S, Om Prakash, Prabhakar N, Subbakrishna DK (1986) Psychiatric morbidity in general practice in an Indian city. *Br Med J* 292 (6537): 1713-1714
- Sriram TG, Radhika MR, Michael A, Chandrashekar CR (1987) Standardisation of the Kannada version of the Crown-Crisp Experiential Index. *Int J Soc Psychiatry* 33: 203-209
- Tornapolsky A, Hand DJ, McLean EK, Roberts H, Wiggins RD (1979) Validity and uses of screening questionnaire (GHQ) in the community. *Br J Psychiatry* 134: 508-515
- Vasquez-Barquero JL, Padierna Acero IA, Marton CP, Ochoteco A (1985) The psychiatric correlates of coronary pathology: validity of the GHQ-60 as a screening instrument. *Psychol Med* 15: 589-596
- Vieweg BW, Hedlund JL (1983) The General Health Questionnaire (GHQ) - a comprehensive review. *J Operat Psychiatry* 14: 74-81

Accepted May 18, 1989

Dr. T. G. Sriram
Department of Psychiatry
National Institute of Mental Health
and Neurosciences
Bangalore 560029
India