

# STRENGTHS AND LIMITATIONS IN USING PSYCHIATRIC MEASURES ACROSS CULTURES

S. RAJKUMAR AND SHUBA KUMAR

## ABSTRACT

*Issues concerning the application of psychiatric measures across cultures are discussed in the light of the authors experiences with the Geriatric Mental State (GMS) and the WHO Quality of Life Instrument (WHOQOL). The acceptability of the GMS to non - literate populations are examined in the context of culture, age ascertainment, education and cognition and variances between the GMS - psychiatrist diagnosis. The WHOQOL which is in the developmental stage, brings forth issues concerning qualitative techniques in instrument development; conceptual, semantic and technical equivalence; translations and backtranslations and response formats. Potential barriers to using an existing measure in a new setting or developing a cross-culturally comparable one are elaborated.*

**KEY WORDS :** *Psychiatric measures, strengths, limitations, cultures, GMS, WHOQOL*

## INTRODUCTION

An important function of scientific research, is to refine methodology, construct instruments to operationalise diagnoses, evaluate treatments and assess prognosis and outcome in a systematic manner. A vast array of measurement instruments developed and standardised across cultures have come into being, providing a better understanding of psychiatric phenomena the world over. These measures are a contrast to the clinical interview which though sensitive, tend to be individualized and difficult to repeat.

Significant contributions have been made by WHO in terms of the development of several instruments which are widely used. The Present State Examination (PSE) (Wing et al, 1974), a structured clinical interview is one such which assesses the current mental status of psychiatric patients. It has been used in several multi-national projects, the International Pilot Study of Schizophrenia (IPSS) being the most well known. When used with the CATEGO programme, which is a computer derived diagnostic system, it offers a standardised means of comparison between stud-

ies across cultures.

It has gone through nine revisions and been translated into eleven languages, a clear indication of its acceptability and applicability to different cultures. Interesting to note are the three Arabic translations of the PSE (Abdel - Mawgood, 1986; Al-Khani et al, 1986; Okasha and Ashour, 1981) which according to Ghubash et al (1992) differ from each other. One is phrased in colloquial Arabic and the other in classical Arabic. This is an indication of the diversity that exists in languages within a particular culture. The Composite International Diagnostic Interview (CIDI) and the Schedule for Clinical Assessment in Neuropsychiatry (SCAN) are two other examples of WHO instruments which have been used in different cultures, providing good evidence to support their validity and reliability (Janca and Chandrashekar, 1993). Not all instruments, however, are as amenable to cross cultural applications.

In India, documented reports on the methodologies used in developing and standardising instruments are few. Invariably, the practice here and probably in many other countries, has been

to use instruments developed in another culture after it has been suitably translated into the local language. This is an issue that is fraught with problems. "The divergence of cultural norms across the world, differing extents of somatization of psychological distress and the existence of culture-bound syndromes have been held to be important reasons for the inadequacy of psychiatric research instruments originally developed in the west" (Sen and Mari, 1986). Undoubtedly, translating and using an existing standardised instrument has the advantage of being both cost-effective and quick. However, the potential dangers of close adherence to the original language version cannot be minimised. As had been pointed out by Kapur (1992), during the process of trying to maintain similarity of interpretation and response, what is forgotten, is that things may be expressed quite differently in the two cultures, thereby rendering the instrument developed in one culture unsuitable in another. Although some attempts have been made by a few researchers in India to obviate these difficulties by developing culture specific measures (Kapur and Carstairs, 1974; Pai and Kapur, 1981) much more needs to be done in terms of refining methodologies and applying stringent psychometric tests to assess the robustness of measures.

In psychiatry, while it is possible to make a diagnosis, it is difficult to verify it by means of laboratory tests, therefore, classification of a subject's true illness cannot be known with certainty. This is where "gold standards" have a significant role to play. A gold standard has been referred to as any method of evaluation which is based on a known disease pathophysiology that is also known to be a good indicator of true disease status (Faroane and Tsuang 1994). But, in formulating our gold standards, we have to a large extent, been influenced by the norms laid down by other cultures. Consequently, we tend to fit our patients into alien patterns to make them qualify as "patients or "cases" (Chakraborty, 1990). In the process important, qualitatively rich information is lost to us. Faroane and Tsuang (1994) have encour-

aged the use of diagnostic accuracy statistics involving mathematical models that estimate accuracy statistics as a means of overcoming these difficulties. These methods however, supplement but do not replace other validating criteria in making accurate diagnosis. In this context, the gold standard as an ideal, is both useful and necessary and certainly here to stay.

Pre-validated tools thus, have definite advantages and are a pre-requisite for international comparability. But there are some limitations when used in cultures other than the ones they were developed in. Issues concerning its cultural relevance and acceptability; translation methodology; weighting of scores and deciding on cut off points need to be closely scrutinised. Failure to do so could seriously threaten the validity of the instrument. This article examines the strengths and limitations of using instruments across cultures in the light of our experience with (a) the Geriatric Mental State Examination (GMS), a widely used and well validated instrument to assess cognitive decline in the elderly and (b) the WHO Quality of Life Instrument (WHOQOL), a generic measure of QOL which is in the developmental stage. The WHOQOL has been chosen for discussion in view of its bearing on cross-cultural applications.

## COGNITIVE SCREENING

In many developing countries, cognitive screening in elderly is becoming an important issue in view of the major epidemiological and demographic transition. The GMS (Copeland et al, 1976; Gurland et al, 1976) was prepared primarily for use with elderly subjects. In addition to assessing dementia, it also explores symptoms of depression, anxiety, hypochondriasis and impairment in activities of daily living, thereby providing a comprehensive assessment tool. It is a semi-structured screening instrument derived from the PSE and the Psychiatric Status Schedule (PSS), the latter developed by Spitzer et al (1964). The GMS - a third edition, is the community version consisting of 183 items. The

instrument has been translated into several languages and has been widely used in many studies, providing strong evidence of its validity (Henderson et al, 1983). It is fairly easy to administer and couched in simple language to enable ease of comprehension by the subject.

A computer diagnostic system called the Automated Geriatric Examination Computer Assisted Taxonomy (AGECAT) has been developed to be used with the GMS for case finding and diagnosis. In addition to providing a main and alternative diagnosis wherever appropriate, the AGECAT produces 6 levels of confidence of diagnosis on each of the diagnostic syndrome categories. A confidence level of 3 and above represents the psychiatrist's concept of a syndrome case (Copeland and Dewey, 1989).

While the instrument is comprehensive, easy to administer, widely used, well validated and permits cross-cultural comparisons, there are some difficulties, particularly with respect to its applicability to non-literate populations. These issues are discussed under (a) Culture (b) Age ascertainment (c) Literacy and cognitive impairment.

## **CULTURE**

In rural settings in India, symptoms of cognitive impairment among the elderly are often explained away as part of the normal ageing process. Many do not perceive it as a problem until the disease is fairly well advanced. Hence mild cases of dementia tend to get missed. One of the issues that is particularly difficult to measure in rural elderly, is activities of daily living (ADL). For varying reasons, including lack of job opportunities and societal norms of "take rest when old", their personal productivity is low. While many are capable of work, they lack the opportunity to do so and tend to get marginalised.

These dependency fostering cultures also uphold the belief that "care and respect for elderly family members, especially one's parents is expected and is sometimes a religious, legal or certainly a filial duty" (Phillips, 1992). It is consid-

ered culturally appropriate to assist elderly people in routine activities and hence, assessing a person's functional capacity becomes difficult (Rajkumar and Kumar, 1996, in press).

## **AGE ASCERTAINMENT**

For the vast number of India's people, the majority of whom live in villages, knowledge of age, dates and current events is poor, nor do the people perceive a need to know it. Ascertaining accurate age thus proved to be a difficult exercise, particularly during the rural study. This was obviated to some extent through the use of the events calendar technique. By alluding to events in the person's life and linking it with prominent happenings in the community during a set period, an approximate age was ascertained. The events chosen were generally those that were well remembered by individuals, irrespective of socio-economic status or educational background. In addition to personal events as appropriate, like year of attaining menarche, year of marriage, birth of child and year of any significant exit events, important historical events specific to Madras and to the country were also cited (Srinivasan et al, 1993). Further corroborations with family members and friends helped to validate the respondent's age. However, the age estimates thus obtained were at best close approximations. This inability to state age is the norm rather than the exception and is not indicative of the presence of cognitive impairment.

## **LITERACY AND COGNITIVE IMPAIRMENT**

Poverty and poor literacy levels are common in many developing countries. Subjects therefore find it difficult to answer GMS questions on memory and orientation. Studies reporting the relationship between cognitive performance in the context of the MMSE, suggest that the MMSE is biased against the poorly educated (Anthony et al, 1982). Brayne and Calloway (1989) reported that higher education was associated with better scores for all of the more complex questions in

the MMSE. Liu et al (1994) and Mortimer (1988) concluded that, education has a protective effect and helped prevent the onset of dementia.

Whether such changes in educational level reflect true differences in clinical state or are due to bias in test content is as yet unclear. Further, co-morbid states due to nutritional deficiencies and infectious diseases could accentuate the deterioration process. As observed by Folstein et al (1985) it is important to distinguish the meaning of the "objective syndrome" of cognitive impairment from the "diseases which cause the syndrome and from the social conditions such as lack of education, that can limit the expression of cognitive capacity and perhaps cognitive capacity as well."

**GMS-AGECAT AND PSYCHIATRIST DIAGNOSIS**

**TABLE 1. Comparison of the WHODEMS and the Rural Dementia study.**

Area	WHODEMMS Madras city	RURAL STUDY Thiruporur Village
Sample	1,300 >65 yrs	750 > 65 yrs.
Sampling Technique	Multi-stage stratified random sample	Cluster sample
Diagnostic Criteria	ICD-10 Psychiatrists Consensus diagnosis	ICD-10 Psychiatrists consensus diagnosis
Instrument	GMS-A, 3rd ed.	GMS -A, 3rd ed.
Prevalence GMS-AGECAT	4.2%	6.2%
Prevalence Psychiatric interview	2.7%	3.5%

As indicated in Table 1, prevalence estimates as per the GMS-AGECAT were higher in comparison to the consensus diagnosis of the psychi-

atrists. This was all the more in rural settings where knowledge of age, date and current events are not perceived as important by the people. Interestingly, a total of 55 subjects in the WHODEMMS study in Madras and 92 subjects in the rural study, were given a confidence level of 2. While this is not indicative of having reached case level, it suggests that these could become potential cases.

There was greater disparity between the prevalence rates derived by the GMS-AGECAT (6.2) and the psychiatrist (3.5%) in the rural study. Inability to correctly answer questions on age, date of birth and current events were largely responsible for this. With poorer literacy levels in villages, such subjects appear to have been at a disadvantage. Normative data on what constitutes the parameter for cognitive impairment for a particular culture would help provide a means to estimating pathology in the community.

**QUALITY OF LIFE**

Quality of life (QOL) has in recent times, gained momentum as an important health status measure. Apart from providing valuable information on what, according to an individual, is important for a good QOL, it provides a means of evaluating the effectiveness of intervention programmes. It is a broad ranging and complex concept that has been the subject of much debate. Essentially, it incorporates the person's physical health, psychological status; level of independence; social relationships, personal beliefs and their association with salient features in the environment. There are two approaches to its assessment, the objective and the subjective. While the objective refers to dimensions of life that all people value or require like food, shelter, mobility and good health, more important is the subjective evaluation of inner experiences which have both positive and negative dimensions.

QOL instruments generally fall into two broad categories, disease specific and generic. The Arthritis Impact Measurement Scale (AIMS) is an example of a disease specific measure that has

been developed specifically for use among arthritic patients. Though a useful scale, it does not focus on how the symptoms have affected the lives of the patients. The Nottingham Health Profile (NHP) and the Sickness Impact Profile (SIP) are examples of generic measures. While the NHP is simple, sensitive and covers a broad range, it is designed to elicit serious problems, therefore, those with minor problems tend to have low scores. The SIP has proved valuable for assessing impact of illness in the chronically ill, however the length of the measure does cause concern.

### THE WHOQOL INSTRUMENT

The WHOQOL, is an instrument developed by WHO and currently in the process of being standardised in many countries (WHOQOL Group, 1995). During its development in Madras and other centres, certain significant issues emerged. These concern qualitative techniques in instrument development, conceptual, semantic and technical equivalence, translation and backtranslation and response formats.

### QUALITATIVE TECHNIQUES

Threats to questionnaire validity arise when they are used in cultures other than the ones they were developed in. Qualitative methods offer a strong set of validity checks regarding the practices and conditions of people. One such ethnographic method used during the development of the WHOQOL instrument, was the focus group technique. The initial focus groups that were conducted simultaneously in each of the countries, inclusive of Madras, helped in validating the domains and facets of QOL that had been drawn up over several discussions with health professionals. Framing of items, so that they would be culturally appropriate, was also done during the focus groups. This ensured that questions were written in the language of the field centre and thereby retained the cultural flavour. Subsequently they were translated into English, which was the working language, enabling cross-cultur-

al comparisons.

Some of the domains and facets of QOL that emerged were freedom from disease, happiness, working capacity and job opportunities, personal relations, personal dignity, religious and spiritual fulfillment and peace of mind. More specifically, people in Madras stressed the importance of being able to conceive children and being a parent, as being particularly significant for a good QOL.

### CONCEPTUAL, SEMANTIC AND TECHNICAL EQUIVALENCE

Comparison of research findings undertaken in different cultures could be misleading unless common methodologies are adopted. The danger of distortion could be further compounded by differences in languages and cultural background of the subjects. Testing for conceptual, semantic and technical equivalence thus, becomes both important and necessary (Sartorius and Kuyken, 1994).

Conceptual equivalence refers to whether the instrument measures the same theoretical construct in each culture. It refers to the position and significance of the word in the theoretical system to which it belongs. For example, in the Indian culture the concept of "being patronised" had none of the negative connotations that it has in many western cultures.

Semantic equivalence refers to the denotative and connotative sameness of words used in instruments and interviews. Denotation of words refers to their "cognitive meaning" and is based on the definition of words found in dictionaries. Connotation of words on the other hand, refers to the "emotional meaning" of words and can be examined using anthropological analyses. These can be studied by comparison of their semantic spaces i.e an equivalent word in another language occupies a similar semantic space and most synonyms have a direct counterpart in other languages. Use of focus and bilingual groups, as was done during the WHOQOL, helped identify terms for which semantic equivalence needed to be examined. For example, the Tamil word for

depression can be understood as, "worry", "sad", "unhappy" or "extreme unhappiness", depending upon individual perceptions. Spoken Tamil is colloquial in usage and is different from its written form which relies heavily on grammar. These issues could result in both subject and observer bias.

Technical equivalence refers to whether the method of assessment is comparable in each culture. Achieving this becomes difficult in fully standardised instruments which have to be used in exactly the same manner by all interviewers. Thus, ensuring technical equivalence rests on the training imparted to investigators who need to be fully apprised of the nature of rating required.

### TRANSLATION

Development of instruments for each culture is not cost-effective, therefore, the most practical and feasible alternative is the translation of existing measures. While conventional translation strategies are simple and quick to carry out, they have limitations which could affect the applicability of the instrument to the new culture. Of importance to recognise, is that close adherence to the original language questionnaire during translation, does not automatically guarantee the validity of the new language version. Questionnaires are likely to have weaknesses even in their original form. There may be items that do not translate well or else do not make sense in the new language version; some items may be important for the culture for whom it was originally created but trivial for the new one and crucial items for the new culture may be missing in the original questionnaire (Guyatt, 1993).

To obviate potential flaws in translation, WHO used a methodology involving bi and monolingual groups during the development of WHOQOL. For example, in Madras, the WHOQOL which was conceived in Tamil, was initially given to a monolingual group (those who knew only Tamil) to obtain their comments on the language and wording of the instrument. It was then

backtranslated into English by a bilingual group (those who knew both English and Tamil) and the two versions compared for equivalence. In this manner, comparability and cultural applicability were maintained.

### RESPONSE FORMATS

The need for social approval, trait desirability and acquiescence are three response tendencies that need to be considered among cultural groups. As suggested by Carr and Krause (1978), checking for acquiescence is built into many instruments by using paired statements that are logical opposites. In the Indian culture, questions relating to sexual satisfaction and functioning, invariably tend to draw answers which are socially desirable. Responses from unmarried girls to items on heterosexual behaviour are considered offensive as they go against their value system.

In conclusion, instrument development has reached new heights in terms of sophisticated methodologies and statistical calculations to determine their psychometric properties. However, due credence needs to be given to potential barriers to the use of instruments across cultures in the interests of enhancing their acceptability.

### REFERENCES

- Abdel-Mawgood M.** (1986) *The English-Arabic version of PSE : An Arabic translation of the full version of Present State Examination (PSE)*. Psychiatric Hospital, Bahrain.
- Al-Khani M.F., Bebbington P.E., Watson J.P. and House F.** (1986) Life-events and schizophrenia: A Saudi Arabian study. *British Journal of Psychiatry*. 148 : pp.12 - 22.
- Anthony J.C., Niaz U., LeResche L.A., Von Korff M.R. and Folstein M.F.** (1982) Limits of the Mini-Mental State as a screening test for dementia and delirium among hospital patients. *Psychological Medicine*. 12, pp. 397-408.
- Brayne C. and Calloway P.** (1989) An epidemiological study of dementia in a rural population of elderly women. *British Journal of Psychiatry*. 155, pp. 214 - 219.

- Carr L.G. and Krause N.** (1978) Social states, psychiatric symptomatology and response bias. *Journal of Health and Social Behaviour*. 19 : pp. 86-91.
- Chakraborty A.** (1992) Cultural perspectives in Indian psychiatry (editorial). *Indian Journal of Psychiatry*. 34, 1, 1 - 2.
- Copeland J.R.M. and Dewey M.E.** (1989) The computer assisted systems. In : *Classification and diagnosis of Alzheimer's disease*. (eds) T. Hovaguimian, S. Henderson, Z. Khachaturian and J. Orley. pp. 87 - 94, Hogrefe and Huber Publishers.
- Copeland J.R.M., Kelleher M.J., Kellet J.M., Gurland B.J., Fleiss J.L. and Sharpe L.** (1976) A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly. The Geriatric Mental State Schedule : I development and reliability. *Psychological Medicine* 6 : pp. 439 - 449.
- Faraone S.V. and Tsuang M.T.** (1994) Measuring diagnostic accuracy in the absence of a "gold standard". *American Journal of Psychiatry*. 151 : pp 60 - 657.
- Folstein M., Anthony J.C., Parhead I., Duffy B. and Gruenberg E.M.** (1985) The meaning of cognitive impairment in the elderly. *Journal of American Geriatric Society*. 33: pp. 228 - 235.
- Ghubash R., Hamdi E. and Bebbington P.E.** (1992) The Dubai community psychiatric survey : prevalence and sociodemographic correlates. *Social Psychiatry and Epidemiology* 27 : pp. 53 - 61.
- Gurland B.J., Fleiss J.L., Goldberg K., Sharpe L., Copeland J.R.M., Kelleher M.J. and Kellete J.M.** (1976) A semi-structured clinical interview for the assessment of diagnosis and mental state in the elderly. The Geriatric Mental State Schedule : II A factor analysis. *Psychological Medicine* 6 : pp. 451 - 459.
- Guyatt G.H.** (1993) The philosophy of health related quality of life translation. *Quality of Life Research*. 2 : pp. 461 - 465.
- Henderson A.S., Duncan - Jones P. and Finlay - Jones R.** (1983) The reliability of the Geriatric Mental State Examination. *Acta Psychiatrica Scandinavica*. 67 : pp. 281 - 289.
- Janca A. and Chandrashekar** (1993) Catalogue of assessment instruments used in the studies coordinated by the WHO mental Health Programme. Division of Mental Health, WHO, Geneva.
- Kapur R.L.** (1992) The family and schizophrenia : Priority areas for intervention research in India. *Indian Journal of Psychiatry*. 34, 1, 3 - 7.
- Kapur R.L., Kapur M. and Carstairs G.M.** (1974) Indian Psychiatric Interview Schedule : IPIS. *Social Psychiatry*. 9, 61.
- Liu H.C., Chou P., Lin K.N., Wang S.J., Fuh J.L., Lin H.C., Liu C.Y. Wu., G.S., Larson E.B., White L.R., Graves A.B. and Teng E.L.** (1994) Assessing cognitive abilities and dementia in a predominantly illiterate population of older individuals in kinmen. *Psychological Medicine* 24 : pp. 763 - 770.
- Mortimer J.A.** (1988) Do psychosocial risk factors contribute to Alzheimer's disease ? In : *Etiology of dementia of Alzheimer's type*. (eds.) A.S. Henderson and J.H Henderson. pp. 39 - 52. John Wiley and Sons, New York.
- Okasha A. and Ashour A.** (1981) Psychodemographic study of anxiety in Egypt : The PSE in its Arabic version. *British Journal of Psychiatry*. 139 : pp. 70 - 73.
- Pai S. and Kapur R.L.** (1981) The burden on the family of a psychiatric patient : Development of an interview schedule. *British Journal of Psychiatry*. 138, 333 - 335.
- Phillips D.R.** (1992) Ageing in east and south-east Asia. Edward Arnold. London, pp. ix.
- 23.Rao K., Subbakrishna D.K. and Prabhu G.G.** (1989) Development of a coping checklist - A preliminary report. *Indian Journal of Psychiatry*. 31, 2, 128 - 133.
- Rajkumar S. and Kumar S.** (1996 in press) Measuring quality of life among the elderly in developing countries. *International Journal of Geriatric Psychiatry*.
- Sartorius N. and Kuyken W.** (1994) Trans-

lation of health status instruments. In : *Quality of life assessments. International perspectives*. Proceedings of WHO and Foundation IPSEN Paris meeting (eds.) J. Orley and W. Kuyken, pp. 3 - 17, Springer-Verlag Publication.

**Sen B. and Jesus Mari J de.** (1986) Psychiatric research instruments in the transcultural setting: Experiences in India and Brazil. *Social Science and Medicine*. 23, 3, 277 - 281.

**Spitzer R.L., Fleiss J.L., Burdock E.I. and Hardesty A.S** (1964) The Mental Status Schedule : Rationale, reliability and validity. *Comprehensive Psychiatry*. 5 : pp. 384 395.

**Srinivasan T.N., Suresh T.R. and Rajkumar S.** (1993) Age estimation in the elderly. Relevance to geriatric research and developing countries. *Indian Journal of Psychiatry*. Vol. 35, No.1 : pp. 3 - 17.

**WHOQOL Group** (1995) The World health Organisation quality of life assessment (WHOQOL) : Position paper from the World Health Organisation. *Social Science and Medicine*. 41, 10, 1403 - 1409.

**Wing J.K., Cooper J.E. and Sartorius N.** (1974) *Measurement and classification of psychiatric symptoms*. Cambridge University Press, Cambridge.

---

*S. Rajkumar M.D, MRCPsych. Professor of Psychiatry, Shuba Kumar M.S.W\* Social Scientist, Clinical Epidemiology Unit, Madras Medical College & Government General Hospital, Madras - 600 003*

*\* Correspondence*