

Tilburg University

Developing methods for assessing quality of life in different cultural settings

Skevington, S.M.; Sartorius, N.; Amir, M.; WHOQOL, G.; van Heck, G.L.; de Vries, J.

Published in:
Social Psychiatry and Psychiatric Epidemiology

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Skevington, S. M., Sartorius, N., Amir, M., WHOQOL, G., van Heck, G. L., & de Vries, J. (2004). Developing methods for assessing quality of life in different cultural settings. *Social Psychiatry and Psychiatric Epidemiology*, 39(1), 1-8.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

ORIGINAL PAPER

Suzanne M. Skevington · Norman Sartorius · Marianne Amir and The WHOQOL-Group¹

Developing methods for assessing quality of life in different cultural settings

The history of the WHOQOL instruments

Accepted: 10 June 2003

Abstract *Background* This review traces the 12-year history of an international collaboration of researchers – the WHOQOL Group – who were brought together by the World Health Organisation to develop and produce a cross-cultural measure of quality of life for use in health and health care. *Discussion* The theoretical and philosophical basis of the WHOQOL instrument is outlined as it evolved throughout the design and adjustment of a reflexive methodology that places an assessment of the users' views at the centre of health care. The stages of research are further expanded and explained in recounting the scientific experience of this unique collaboration. The WHOQOL is available in 40 countries and most majority languages. Adaptations exist for assessing particular conditions, e.g. spirituality, religion and personal beliefs.

Key words quality of life – health – the WHOQOL – assessment – development

Introduction

The term quality of life (QoL) has become a by-word for many politicians, but its measurement has, until re-

cently, remained elusive. Here, we outline the history of a unique collaborative project that was set up through the World Health Organisation (WHO) over 10 years ago, and chart milestones in its progress to the present. The objectives of the WHO programme put forward at the time were health-orientated, political and scientific. The first had to do with the continuous deterioration of the doctor-patient relationship; it was hoped that the widespread utilisation of an instrument measuring QoL would make physicians more aware of the need to listen to their patients and to take their feelings into account during treatment. The second objective was scientific;

¹ The paper is based on data and experience obtained as part of the WHO study to develop a QoL measure (WHOQOL). The collaborators in this study have been at WHO Geneva: Dr. Norman Sartorius, Dr. J. Orley, Dr. Willem Kuyken and Dr. Mick Power. In the Field Research Centres collaborating investigators are Prof. Helen Herrman, Dr. H. Schofield and Ms B. Murphy, Univ. of Melbourne, Australia, Prof. Z. Metelko, Prof. S. Szabo and Mrs. M. Pibernik-Okanovic, Institute of Diabetes, Endocrinology and Metabolic Diseases and Dept. of Psychology, Faculty of Philosophy, Univ. of Zagreb, Croatia, Dr. N. Quemada and Dr. A. Caria, INSERM, Paris, France, Dr. S. Rajkumar and Mrs. Shuba Kumar, Madras Medical College, India, Dr. S. Saxena, All India Institute of Medical Sciences, Delhi, India, Dr. D. Baron and Dr. M. Amir, Ben Gurion Univ., Beer Sheeva Israel, Dr. Miyako Tazaki, Dept. of Science, Science Univ. of Tokyo, Japan and Dr. Ariko Noji, Dept. of Community Health Nursing, St. Luke's College of Nursing, Japan, Dr. G. van Heck and Mrs. J. de Vries, Tilburg Univ., The Netherlands, Prof. J. Arroyo-Sucre and Prof. Pichard-Ami, Univ. of Panama, Panama, Prof. M. Kabanov, Dr. A. Lomachenkov, and Dr. G. Burkovsky, Bekhterev Psychoneurological Institute, St. Petersburg, Russia, Dr. R. Lucas Carrasco, Barcelona, Spain, Dr. Yooth Bodharamik and Mr. Kitikorn Meesapya, Institute of Mental Health, Bangkok, Thailand, Dr. S. Skevington, Dept. of Psychology, Univ. of Bath, Bath, UK, Dr. D. Patrick, Ms M. Martin and Ms D. Wild, Univ. of Washington, Seattle, USA and Prof. W. Acuda and Dr. J. Mutambirwa, Univ. of Zimbabwe, Harare, Zimbabwe. An international panel of consultants includes: Dr. N. K. Aaronson, Dr. P. Bech, Dr. M. Bullinger, Dr. He-Nian Chen, Dr. J. Fox-Rushby, Dr. C. Moinpur and Dr. R. Rosser. Consultants who have advised WHO at various stages of the development of the project have included: Dr. D. Buesching, Dr. D. Bucquet, Dr. L. W. Chambers, Dr. B. Jambon, Dr. C. D. Jenkinson, Dr. D. De Leo, Dr. L. Fallowfield, Dr. P. Gerin, Dr. P. Graham, Dr. O. Gureje, Dr. K. Kalumba, Dr. Kerr-Corea, Dr. C. Mercier, Mr. J. Oliver, Dr. Y. H. Poortinga, Dr. R. Trotter and Dr. F. van Dam

Prof. S. M. Skevington (✉)
WHO Field Centre for the Study of Quality of life
Dept. of Psychology
University of Bath
Bath, BA2 7AY, UK
Tel.: +44-1225/386-830
Fax: +44-1225/386-752
E-Mail: s. m.skevington@bath.ac.uk

N. Sartorius
Dept. of Psychiatry
University of Geneva
Geneva, Switzerland

M. Amir
Dept. of Psychology
Ben Gurion University of the Negev
Beer Sheeva, Israel

advances in treatments increased the need to complement the assessment of outcomes of medical interventions by a measure of the effects that interventions had on QoL.

The instrument that the WHO set out to develop had to satisfy the usual psychometric requirements. In addition, however, this instrument, that was intended for worldwide use, had to be simple so as to be easily used, and demonstrably applicable in different cultural settings. Any such instrument would be useful in monitoring multinational clinical trials, epidemiological studies of sick and well populations and in comparing the impact of different conditions on human health and satisfaction. Furthermore, at a theoretical level, it would assist in answering questions about the universality of the QoL concept. The WHOQOL group defined QoL as “an individual’s perceptions of their position in life, in the context of the culture and value systems in which they live, and in relation to their goals, expectations, standards and concerns” (WHOQOL Group 1995).

With these aims in mind, the Division of Mental Health at WHO convened an international meeting in 1991, bringing together anthropologists, health psychologists, medical sociologists, psychometricians, policy makers, cross-cultural researchers and clinicians with expertise in the major disease groups. The WHO instrument was to be conceptually novel in that it was to give equal weight to the person’s level of functioning – often measured in some methods as a substitute for QoL – as well as the individual’s assessment of the importance of that level of functioning for their QoL. Several priority areas would need to be tackled including the assessment of people suffering from severe illnesses or disability, groups like caregivers of the elderly, disabled and distressed, whose QoL was rarely assessed, also people living in highly stressful situations, like migrants and refugees. It was recognised that, while a personal and subjective assessment seemed the most desirable outcome, there was a need to produce the instrument in a manner that would assess QoL in those who were unable to communicate, like stroke victims and those with advanced dementia. This paper reviews the historical development of the philosophy and methodology that has underpinned the WHOQOL project at a time when the body of psychometric work on the main assessments is becoming known, giving attention to some areas of design and conceptualisation that have not been elaborated previously.

■ Early development and conceptualisation of the WHOQOL

A literature review revealed that unidimensional models of QoL were less useful and satisfying to patients and clinicians than multidimensional models. Clinicians often needed to know more about QoL than whether it had simply improved or deteriorated, and this could not be determined from a unidimensional measure. Many want

to know exactly in what areas QoL has deteriorated so that they can better target their therapeutic action. On the other hand, patients want to be able to tell health professionals in what areas their QoL has changed. For instance, chronic arthritis patients retain good QoL on some dimensions even though their illness detracts from it on others, so that unidimensional QoL ratings hide more than they disclose.

In formulating dimensions of QoL that would be salient and important worldwide, the group that met in Geneva embarked on a creative open-ended exercise to list numerous dimensions and ideas from their professional experience, scientific knowledge and cultural backgrounds. In distillation, 134 facets² were clustered into seven domains³ as the profile was to be scored by domain and facet. A further 21 moderating variables, like race and climate, that might be expected to affect the expression of QoL on all other dimensions were later removed as their status as outcomes *per se* was ambiguous. Focus groups were then held in the field centres to discuss the cultural relevance of the listed domains and facets. This was important because it enabled us to anticipate from a very early stage some of the problems of gaining semantic, conceptual and technical equivalence between different language versions of the new questionnaire, by finding out which facets were difficult or awkward to discuss or rate in that cultural setting (Sartorius and Kuyken 1994; Skevington 2002).

Two meetings in 1992 indicated widespread agreement about constructing a generic measure of QoL as perceived by the individuals concerned, and investigating whether these constructs and measures were equivalent in different cultures. There would be a generic core of questions common to all people, as well as additional modules of extra questions that could be included when assessing the QoL of people with a particular disease, e.g. cancer, or condition, e.g. chronic back pain. It was agreed to produce common and consensually derived methods using a protocol that all centres would follow throughout instrument development. Because all centres were involved in designing the protocol at each stage, this afforded the opportunity to carry out necessary “fine tuning” to accommodate cultural variations within the instrument and address difficulties arising from these. National versions of the instrument would be supervised through national centres and the data owned jointly by that centre and WHO. National items tapping the specifics of different cultures would round out the concept of QoL for that cultural group, so improving conceptual equivalence and, hence, any comparisons between the different language versions. They would be highly recommended by focus groups as ne-

2 Facets describe behaviours (e.g. activities as provider), states of being (e.g. fatigue), capacities (e.g. the ability to move around) or subjective perceptions of experiences (e.g. pain) (WHOQOL Group 1994).

3 Domains describe core aspects of QoL cross-culturally (WHOQOL Group 1994).

cessary to the full assessment of the concept of QoL in that culture, e.g. skin colour in India, security in Israel and feeling 'fed up' in the UK. National items would also need to be as satisfactory in psychometric terms as the international core items. The collaborating centres were selected to represent different levels of industrialisation and cultural traditions; some already had connections with WHO, while others were invited to join because of their experience of working on QoL.

At this stage, several styles of questioning about QoL were identified and discussed. It was possible to write questions at five different levels, but was it desirable to include them at all these levels of assessment? Objective observed measures enable health professionals to judge functioning and prognosis on the basis of EEG readings, biomechanical activities, etc. Although sometimes used as a proxy for QoL and important, it was argued that they do not measure QoL as a subjective or perceived experience. At a second level, individuals' 'objective' judgements about their own state provide 'objective' self-reports on, for example, how many hours they sleep. Three other self-assessments of internal state are the perceived satisfaction with a given level of functioning, perceived interference with the achievement of goals, and the perceived importance of a particular function (or disturbance of it) to the perception of their QoL.

Information on face validity from field centre discussions enabled facet descriptions and their organisation into parent domains to be reworked and refined during meetings. Discussion focused on such issues as whether the availability of nutritious food and sufficient clean water are indicators of good QoL and, therefore, part of an environment domain. To what extent does an impaired appetite and enjoyment of food affect QoL universally? Is sex predominantly physical, psychological or social? Should relationships with family, friends and spouse be assessed as separate facets or together as personal relationships? Collectively, facet definitions were drafted that were perceived to be important and universal, and this in turn honed the definition of QoL that the group would eventually adopt. Brief scenarios were written to enable lay people who read the facet definitions to consider some concrete examples. Following this iterative and reflexive process, a manual was assembled providing information that could be used by patients of all types, health professionals and community members, including informal caregivers, to generate core questions in focus groups that would be considered for inclusion in the international questionnaire, or by centres as additional national items.

Facet definitions drawn from several sources were couched in operational terms (see Footnote 2) and a literature review, guides and dictionaries and existing QoL questions also informed this work. Questions were positively framed wherever possible to emphasise positive well-being, so avoiding the more usual problem-centred focus of many other QoL instruments. Labels chosen for facets and domains underscored this positive orientation, for example, a 'levels of independence' not 'depend-

dence' domain, and an energy facet that also dealt with fatigue. This ensured that each dimension evaluated both ends of the positive-negative continuum.

This was one of the strengths of the WHOQOL project because it made it possible to write items in the vernacular, reflecting the culture and language from which they arise. This procedure had not been carried out before in the cross-cultural QoL field and, therefore, constituted a changed orientation towards a more user-centred and, therefore, user-friendly approach. This item-writing procedure serves the dual objectives of, firstly, providing a meaningful questionnaire to the users and, secondly, approaching concept validation and question-writing as a participatory and non-patronising activity. The WHOQOL is an instrument that has been designed by the users for the users. For this reason, it avoids some of the problems of misinterpretation and misunderstandings experienced by former generations of instrument makers. The decision to create an instrument relying on patient directives and concerns rather than on conceptualisations of medical professionals has remained important in later work and in developing new modules for particular diseases or conditions.

From these preliminary investigations, the WHOQOL group agreed the definition of QoL, which was based on the individuals' perceptions of their position in life, and was viewed as being intrinsically influenced by their culture and value systems (WHOQOL Group 1995). This orientation means that the instrument is very well suited to the cross-cultural assessment of healthy as well as sick populations. But, more importantly, this is probably the first definition of QoL to explicitly build in culture as a central and integral part of the concept, rather than as an additional, incidental variable, to be partialled or synthesised out, with the ostensible aim of creating a 'culture free' assessment (Fox-Rushby 1994). So, here, culture and values are seen as quintessential and integral elements of QoL and not as nuisance variables that hamper a 'clean' assessment.

■ Qualitative phase and designing the instrument

During early 1993, centres used the finalised manual to run focus groups in all the participating field centres. Groups began with an open-ended discussion about QoL and participants rated the most important aspects of QoL to them. These pooled ratings and transcribed discussions were later used to refine the facet and domain list and facet definitions. Transcriptions provided context and meaning in affirming parent domains for each facet. Two moderators guided the discussion of facet definitions; they encouraged participants to suggest questions that might be asked to summarise QoL in that facet. As there would not be many questions for each facet, it was important to tap into the overall impact of, say, negative feelings on QoL, rather than disaggregate it into separate components.

The focus group investigations resulted in a total of

2,500 questions. To reduce these to a manageable number, several criteria were applied. Poorly formulated items and those with great similarity to others were removed. Items were assigned to facet clusters as there were distinctive subgroups within many facets and to ensure that at least one item from each cluster was included in the pilot instrument. A total of 235 items were derived from the slimmed-down global pool and attached to one of 29 facets – all facets were attributed to one of five domains. Spirituality, religion and personal beliefs (SRPB) were spontaneously discussed by focus groups as vital areas pertaining to QoL and health that were absent from the existing structure. This was confirmed as important to QoL in many centres. Consequently, suitable SRPB items were derived from the transcripts that could assess a new sixth SRPB domain (and simultaneously as a single facet) of the WHOQOL.

Items were attached to one of five WHOQOL rating scales used in the pilot test. In 1993, work was carried out in each national centre to develop rating scales with appropriate language and interval scaling, that would also have international comparability. Agreement was reached about the international labels that would be attached to the poles of each of five different 5-point scales. Translatable poles were needed in every language to enable respondents to answer questions that asked ‘how much...’ (intensity), ‘how completely...’ (capacity), ‘how satisfied’, ‘happy’ or ‘good’ (evaluation) and ‘how often...’ (frequency) (Szabo for the WHOQOL Group 1996). A fifth response scale on importance had been piloted in a previous WHO international project and was added as acceptable because it had been selected to optimise the discriminative power of the scale.

A range of words was assembled that could be used to answer each of these questions in each language. Participants placed each description on separate 100 mm lines representing the appropriate response scale, and means and standard deviations calculated were used to select the descriptor that would be inserted at the 25%, 50% and 75% points of the interval scale for that centre. Following translation and back-translation, an empirical check was made on the ranked order of descriptors by a new group of lay people. There was also some translation and, hence, cross-checking of these points by other centres, e.g. the Dutch worked with Tamil speakers in Madras. Thirdly, bilingual speakers answered a randomly chosen half of the questions in one language, and the second half in their other language. Where significant differences were found between scores for the two language versions, non-equivalent responses were revised.

At this stage, two types of items were being considered: ‘perceived objective’ items such as ‘How well can you walk?’ and ‘self-report subjective’ items such as ‘How satisfied are you with your ability to walk?’. Making full use of the transcripts, panels of question-writers in the centres submitted these two types of question for inclusion. ‘Perceived objective’ questions assess global evaluations of behaviour, states or capacities, and ‘self-report subjective’ questions address perceived satisfaction or

dissatisfaction with these behaviours, states and capacities (WHOQOL Group 1994). However, for some facets it was found that this distinction could not be applied. Through an exercise where independent raters assigned 36 items (randomly drawn from the pool) to either ‘perceived objective’ or ‘self-report subjective’ categories, there was only 81% successful assignment. Only subjective self-report items could be identified for self-esteem, body image, work satisfaction, positive and negative feelings because of their inherently subjective nature, so it was impossible to create suitable items to reflect a perceived objective perspective. Consequently, the group decided to create an instrument developed entirely of statements at the subjective self-report level. Few other QoL instruments had focused exclusively on this aspect, although others had addressed issues from a perceived objective viewpoint and, therefore, a need was identified in the literature to design such an instrument.

■ Constructing the WHOQOL-100

Several important conceptual decisions were made about the instrument and its nature around this time. Firstly, it was decided to focus items on the perception of the effects or outcomes of the disease and the impact of health care interventions. This contrasted with a view of QoL assessment as the aggregation of symptoms, diseases or conditions. By counting symptoms or assessing their intensity, it was argued that it would not be possible to deduce a person’s QoL. Quality of life is about the *meaning* of these symptoms for individuals, and this meaning cannot be calculated from a simple numerical aggregate. Secondly, discussion about the time frame of the instrument raised theoretical questions about the nature of QoL as a psychological entity. Is QoL a mood, attitude, belief or a personality trait? This answer would indicate how long the time frame would need to be. It was agreed that QoL did not have the transience of mood or the permanence and traditional stability of personality states, but a better framework for explaining QoL is the person-environment fit model (French 1973) where individual differences are interpreted in relation to their specific environment or context. Judgements about QoL are affected not only by internal conditions like mood and personality, but also by important environmental features like the actions of other people, chance happenings of nature and other contextual factors. Time frame decisions were tempered by results from cognitive research showing decrements in remembering health states after 2 weeks, and the deteriorating reliability of information recalled over longer periods (e.g. Reville et al. 1976). A 2-week time frame was chosen for the WHOQOL core items, but none was applied to the importance items, as they appeared less volatile, e.g. ‘How important to you is it to be free from pain?’.

From the beginning, the WHOQOL instrument was conceived as having a long form – the WHOQOL-100 – that could be used for comprehensive assessment and

research, and a shorter version – the WHOQOL-Bref – that would be valuable in large surveys, where repeated measures were necessary and where the state of the patient precluded long assessment. The WHOQOL was meant to be self-administered, but was also constructed to be interviewer-assisted or administered where patients were unable to read or too ill or infirm to complete it themselves. The instrument was also meant to have clear instructions about use and, because of the way in which it was developed, it was expected that the questions (and response scales) would have a high level of acceptability to the users.

Translation of items and response scales has been carried out using a new procedure throughout this project (Sartorius et al. 1994). Several groups and individuals were required to complete a successful translation; a bilingual panel (2–4 people) and monolingual panel (4 people) commented on the translated instrument, with a professional translator. The measure was first translated into the local language by the bilingual group, then the monolingual group identified any incomprehensible or ambiguous language. Both commented on response scale descriptors and instructions. The monolingual group operated solely in the language of the field centre and were not highly educated. Thirdly, the bilingual group reread, commented on and incorporated suggestions from the monolingual group, seeking out further inconsistencies and making amendments. A second professional translator then back-translated the revised instrument into English. The original and back-translated versions were reviewed by bilingual experts at WHO for accuracy and equivalence. Significant discrepancies resulted in further translation and back-translation work that continued until there was a successful transfer of meaning. Lastly, the back-translators reported the summarised procedure and findings.

A comparison of centres that were ostensibly using the same language, e.g. Spanish in Barcelona and Panama City, English in Melbourne, Seattle and Bath, indicated differences in the words chosen by each centre for the three intermediate points of the response scales (Szabo et al. 1996). This points to cross-cultural differences in the psychological intervals and labels used by people in scaling, even where they share the same language. Because the anchor points at the ends of each response scale were internationally agreed as important and translatable before the scaling exercise began, even with these cultural differences in intermediate labelling, there is still equivalence between centres which enables sound comparisons to be made.

The pilot version of the WHOQOL was organised in response scale blocks to speed completion, and this has continued in subsequent versions. However, certain facets like work, SRPB and mobility were completed *en bloc* because each required a special rubric to clarify the types of answers required. For example, for working capacity, ‘work’ was stated to include not only paid work, but all forms of voluntary, unpaid and study work (Table 1).

Table 1 Structure of the WHOQOL

Domains and facets (F) of the WHOQOL	
Domain 1	Physical Health
F1	Pain and Discomfort
F2	Energy and Fatigue
F3	Sleep and Rest
Domain 2	Psychological
F4	Positive Feelings
F5	Thinking, memory, learning and concentration
F6	Self-esteem
F7	Bodily Image and Appearance
F8	Negative Feelings
Domain 3	Levels of Independence
F9	Mobility
F10	Activities of Daily Living
F11	Dependence on Medication and Treatment
F12	Work Capacity
Domain 4	Social Relationships
F13	Personal Relationships
F14	Practical Social Support
F15	Sex
Domain 5	Environmental
F16	Physical Safety and Security
F17	Home Environment
F18	Financial Resources
F19	Health and Social Care: availability and quality
F20	Opportunities for Acquiring New Information and Skills
F21	Participation in, and new opportunities for Recreation and Leisure
F22	Physical Environment
F23	Transport
Domain 6/F24	Spirituality, Religion and Personal Beliefs
General QOL/F25	Overall QoL and general health perceptions

It was necessary to know how important every facet of QoL was perceived to be. Initially, focus group participants gave their 5-point ratings, then results from 41 importance ratings of the pilot survey instrument assisted in selecting facets for the final questionnaire. Only facets obtaining a mean rating of 3.0 or more in at least eight of the 15 countries of the pilot test were considered for retention, so that issues of minority interest were avoided in this international instrument (despite their potential importance in some focused studies) because their redundancy for many respondents would add burden and irritation. Importance ratings are still collected as an appendix to the WHOQOL international core items by many field centres. They reflect the priorities of different countries that were expected to be different from the start. Also, they have enabled us to assess the theoretical basis of this work and to provide a means of identifying who has the very poorest QoL (Skevington et al. 2003).

Item reduction was carried out by first selecting around 1,000 items from a pool of over 2,000 questions using 11 selection criteria. These referred to conceptual issues, e.g. the relationship of questions to their parent facet and domain, to technical issues of administration

like choosing shorter items, and to semantic issues like avoiding double negatives. The reduced pool was ranked and rated by every centre to achieve consensus about the best items to include in the international pilot version (325 items). Key conceptual clusters of items were identified within facets; for example, pain and discomfort contained five clusters describing intensity, coping, control, etc. of pain. Ranked clusters showed the importance of the concept to QoL in each culture, then questions were ranked within clusters. The pilot survey data were used to distinguish four items for each of 25 facets of the WHOQOL-100 (Table 2). Centre data were examined for non-significant and negative correlations, for non-significant discrimination between sick and well samples, and for overlapping items. Regressions identified which items best explained overall QoL. Structural equation modelling (the WHOQOL Group 1998a) and more recently Rasch analysis (Leplege et al. 2000) have supported the construct validity of the WHOQOL. Concep-

tual equivalence was addressed by investigating the psychometric performance of the national items, and the best items are now included in their respective language versions (Skevington et al. 1999). Validation of the WHOQOL-100 has been carried out on depressed patients in the UK (Skevington and Wright 2001) and Argentina (Bonicatto et al. 2001), pregnant women in the US (Bonomi et al. 2000) and chronic pain in the UK (Skevington et al. 2001).

■ The WHOQOL-Bref

In the development of the WHOQOL-Bref from the WHOQOL-100, 26 items (one from each of the 24 facets plus two general items) were selected. It was necessary that they explained a substantial proportion of the total variance within the WHOQOL-100, also for particular domains and the general QoL facet. The final questionnaire was to be structurally sound in terms of confirmatory factor analysis, and was expected to discriminate between identified or 'known' groups (ill vs. well). The psychometric properties of the WHOQOL-Bref items have been tested using the original pilot data ($n = 4,802$), field data ($n = 4,104$) and new data ($n = 2,369$) from more recent centres, and the results are similar and very good. Analysis of the WHOQOL-Bref items shows that domain scores were very similar to those found for the WHOQOL-100; around 95 % of the total facet score variance was explained for the four domains (range 77 % – 91 %). Cronbach's alpha indicated acceptable internal consistency for domains with a marginal result for the small three-item, social relationships domain (physical = 0.83, psychological = 0.75, social = 0.66, environment = 0.80). Domain scores discriminate significantly between sick and well groups ($p < 0.001$). Confirmatory factor analysis showed an acceptable fit of facets within a four-domain solution, and these four domains loaded onto a second order factor representing global QoL. In the multivariate model, the Comparative Fit Index was moderately high, indicating that the parameter estimates were relatively equivalent across all datasets. Regressions showed that all domains significantly contribute to explaining overall QoL and general health (see WHOQOL Group 1998a, 1998b). Test-retest reliability data (over 2–8 weeks) from four centres showed generally high correlations between answers obtained on two occasions.

The WHOQOL-Bref was field-tested in 23 countries ($n = 11,830$) and good to excellent psychometric qualities confirm a high quality instrument in almost all languages (WHOQOL Group 2003). Recently, it has been validated internationally within a study of the natural history of depression – Longitudinal Investigation in Depression Outcomes (LIDO) – where it was administered to 2,359 depressed primary care patients in Australia, Brazil, Israel, Spain, Russia and the USA. Reliability, validity, test-retest and sensitivity to change analyses show that the WHOQOL-Bref performs according to in-

Table 2 A sample page from the UK version of the WHOQOL-100. The following questions ask about *how much* you have experienced certain things in the last two weeks, for example, positive feelings such as happiness or contentment. If you have experienced these things an extreme amount, circle the number next to "An extreme amount". If you have not experienced these things at all, circle the number next to "Not at all". You should circle one of the numbers in between if you wish to show that your answer lies somewhere between "Not at all" and "Extremely". Questions refer to the *last two weeks*

1. How much do you worry about pain or discomfort? (F1.2)

Not at all	Not much	A moderate amount	Very much	An extreme amount
1	2	3	4	5

2. How difficult is it for you to handle pain or discomfort? (F1.3)

Not at all	Not much	Moderately	Very much	Extremely
1	2	3	4	5

3. How much do you feel that pain prevents you from doing what you need to do? (F1.4)

Not at all	Not much	A moderate amount	Very much	An extreme amount
1	2	3	4	5

4. How easily do you get tired? (F2.2)

Not at all	Not much	Moderately	Very much	Extremely
1	2	3	4	5

5. How much are you bothered by fatigue? (F2.4)

Not at all	Not much	A moderate amount	Very much	An extreme amount
1	2	3	4	5

ternational standards (Amir et al. 2003) and is, therefore, ready for use (Table 3).

While the WHOQOL instruments cover all major domains of assessment, they do not provide a detailed evaluation of issues important to the QoL of those with particular diseases and their treatment, e.g. nausea and vomiting during cancer treatment. The procedure for designing adaptations of the WHOQOL through the production of disease-specific modules of items, and the creation of a pilot instrument is identical to that used in developing the core instrument. Starting with focus groups – this time in a minimum of three diverse centres from around the world – it would reassess the existing facet structure to find out whether included facets would be sufficient to assess the QoL of people with the target disease or condition. If the qualitative work indicated that the project was viable, the next step would be a survey and psychometric work.

This programme of work has been completed on several occasions recently and the most advanced of these is about the QoL of people living with HIV/Aids (Skevington and O'Connell 2003). A pressing public health case has been made for the availability of a good quality cross-cultural instrument that would facilitate assessment globally (Lamboray and Skevington 2001; O'Connell et al. 2003). With increasing awareness of the need to better measure the spiritual dimensions in relation to health and QoL, the WHOQOL Group has expanded the assessment of spirituality, religion and personal beliefs within the WHOQOL (Table 4).

Table 3 Historical changes to the WHOQOL

Time	1991/2	1993/4	1995/7	1998/9
Centres	10	15	15	30+
Methods	Qualitative	Quantitative
Versions	Pre-pilot	Pilot	100	Bref
Domains	7	6	6	4
Facets	134	29	25	25
Items	2,500	235	100	26

Table 4 Progress in the psychometric assessment of the WHOQOL global instruments

Scale/qualities	WHOQOL-100	WHOQOL-Bref	WHOQOL-HIV	WHOQOL spirituality
Reliability				
Internal consistency	√	√	√	√
Test-retest reliability	√	√		
Validity				
Content validity	√	√	√	√
Construct validity	√	√	√	√
Concurrent validity	√	√	√	√
Divergent validity	√	√	√	√
Sensitivity	√*	(√)		

Key: √ Assessed; (√) In progress; √* Assessed for some language versions

■ Where do we go from here?

We began with the WHO definition of health and took its positive orientation towards health as something that is 'not merely the absence of disease or infirmity' as the main principle in the creation of an instrument that transcends the problem-centred boundaries necessitated by the clinical consultation and provides a holistic, more balanced view of QoL. The WHOQOL Group continues to expand, and new centres are joining the network using the internationally agreed, standard common protocol that is the 'blueprint' for generating equivalent language versions.

However, there are still important populations whose QoL cannot be assessed, and we may need different methods to adequately obtain reliable judgements of QoL from people challenged by cognitive or communication difficulties, e.g. stroke, dementia, learning disabilities, schizophrenia. Aware of the limitations of our adult age range (up to 65 years), we are developing a WHOQOL-Old for the over-60s that takes account of their special needs and concerns. A children's WHOQOL would be valuable, and this work has started in Thailand (Jirojanakul and Skevington 2000; Jirojanakul et al. 2003). Although the WHOQOL-Bref is now available in nearly 50 language versions, others have yet to be developed. The 'national' language versions have to be studied to establish whether they are linked to the dominant or majority culture, neglecting the idioms or specific problems of minorities.

The WHOQOL is suitable for measuring QoL in healthy and ill populations and is, therefore, neither a disability assessment nor only a measure of distress. It is the most widely used QoL measure in the world. The broad network of collaboration, and the fact that centres are working with valued populations, promises that in the years to come it will be possible to report on QoL worldwide with a common, validated measure. Subjected to structural equation modelling, our extensive data show a high level of agreement among peoples from the most diverse backgrounds about what constitutes a good QoL and well-being, providing good empirical support for a universalist view (Power et al. 1999) and for the hope that the results of QoL investigations will be suitable for use in the development of civic society.

References

- Amir M, Herrman H, Fleck M, Lomashenkov A, Lucas R, Patrick DL (2003) Reliability, validity, reproducibility and responsiveness in the WHOQOL-Bref in six study sites. (Submitted)
- Bonicatto SC, Dew MA, Zaratiegui R, Lorenzo L, Pecina P (2001) Adult outpatients with depression: worse quality of life than other chronic medical diseases in Argentina. *Soc Sci Med* 52(6): 911–919
- Bonomi A, Patrick DL, Bushnell DM, Martin M (2000) Validation of the United States version of the World Health Organisation Quality of Life (WHOQOL) instrument. *J Clin Epidemiol* 53(1): 1–12

4. Fox-Rushby J (1994) The relationship between health economics and health-related quality of life. In: Orley J, Kuyken W (eds) *Quality of Life Assessment: international perspectives*. Springer-Verlag, Berlin, pp 61–74
5. French JR (1973) Person-role fit. *Occupational Mental Health* 3:15–20
6. Jirojanakul P, Skevington SM (2000) Developing a quality of life measure for children aged 5 to 8 years. *Br J Health Psychol* 5(3): 299–322
7. Jirojanakul P, Skevington SM, Hudson JR (2003) Predicting young children's quality of life. *Soc Sci Med* 57(7):1277–1288
8. Lamboray J-L, Skevington SM (2001) Defining AIDS competence: a working model for practical purposes. *J Intern Develop* 13(4):513–521
9. Leplege A, Ecosse E, the WHOQOL Rasch Project Scientific Committee (2000) Methodological issues in using the Rasch model to develop a quality of life index: the analysis of four WHOQOL-100 data sets (Argentina, France, Hong Kong, UK). *J Applied Measurement* 1(4):389–418
10. O'Connell K, Skevington SM, Saxena S on behalf of the WHOQOL-HIV Group (2003) Preliminary development of the World Health Organisation's Quality of Life HIV Instrument (WHOQOL-HIV): analysis of the pilot version. *Soc Sci Med* 57(7): 1259–1275
11. Power M, Bullinger M, Harper A on behalf of the WHOQOL Group (1999) The World Health Organisation WHOQOL-100: tests of the universality of quality of life in fifteen different cultural groups worldwide. *Health Psychol* 18(5):495–505
12. Reville SF, Robinson JO, Rosen M, Hogg MIJ (1976) The reliability of the linear analogue scale for evaluating pain. *Anaesthesia* 31:1191–1198
13. Sartorius N, Kuyken W (1994) Translation of health status instruments. In: Orley J, Kuyken W (eds) *Quality of life assessment: international perspectives*. Springer-Verlag, Berlin, pp 3–19
14. Skevington SM (2002) Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. *Qual Life Res* 11:135–144
15. Skevington SM, Wright A (2001) Changes in the quality of life of patients receiving anti-depressant medication in primary care: validating the WHOQOL-100. *Br J Psychiatry* 178:261–267
16. Skevington SM, Carse M, Williams AC de C (2001) Validation of the WHOQOL-100: pain management improves quality of life for chronic pain patients. *Clin J Pain* 17:264–275
17. Skevington SM, Bradshaw J, Saxena S (1999) Selecting national items for the WHOQOL: conceptual and psychometric considerations. *Soc Sci Med* 48:473–487
18. Skevington SM, O'Connell KA (2003) Measuring Quality of life in HIV and AIDS: a review of the recent literature. *Psychol Health* 18(3):331–350
19. Skevington SM, O'Connell K and the WHOQOL Group (2003) Can we identify the poorest quality of life? Assessing the importance of quality of life using the WHOQOL-100. *Quality of Life Research* (13(1):in press)
20. Szabo S on behalf of the WHOQOL Group (1996) World Health Organisation Quality of Life Assessment Instrument. In: Spilker B (ed) *Quality of life and pharmacoeconomics in clinical trials* (2nd edn). Lippincott-Raven, Philadelphia, pp 355–362
21. The WHOQOL Group (1994) Development of the WHOQOL: rationale and current status. *Intern J Ment Health* 23(3):24–56
22. The WHOQOL Group (1995) The World Health Organisation Quality of Life Assessment (WHOQOL): position paper from the World Health Organisation. *Soc Sci Med* 41(10):1403–1409
23. The WHOQOL Group (1998a) The World Health Organisation Quality of Life Assessment (WHOQOL): development and general psychometric properties. *Soc Sci Med* 46(12):1569–1585
24. The WHOQOL Group (1998b) Development of the WHOQOL-Bref Quality of Life Assessment. *Psychol Med* 28:551–558
25. Skevington SM, Lotfy M, O'Connell K for the WHOQOL Group (2003) The World Health Organisation's WHOQOL-Bref Quality of Life Assessment: psychometric properties and results of the international field trial. *Quality of Life Research* (in press, online)