Translation and Cultural Adaptation of Health Questionnaires

A. Rahman1,Z. Iqbal2,W. Waheed3,N.Hussain4 (School of Psychiatry and Behavioural Sciences, Univerity of Manchester1, U.K., Human Development Research foundation2, Islamabad, Pakistan, Department of Psychiatry, Manchester Royal Infirmary3, Manchester, U.K. and Culture Community and Health Studies Program, Centre for Addiction and Mental Health,University of Toronto4, Toronto, Canada.)

Introduction

Often, health questionnaires have been developed for content, validity, and reliability in one country or language, and are then used in different language and cultural settings. More often than not, little attention is paid to the cross-cultural adaptation that is necessary. While the ideal solution is to develop indigenous instruments and establish their psychometric properties in the local population, this is not always possible because of lack of resources and expertise. Furthermore, most health constructs are universal and can be applied to diverse populations after cultural adaptation. Therefore it is often more feasible to use tried and tested instruments after appropriate adaptation.

The process of translation and adaptation can be broken down into three steps: (a) the translation process; (b) cross-cultural verification and adaptation; and (c) verifying the psychometric properties of the instrument in the target population. The first two steps of the process will be considered in this paper, which will describe the development of a translation protocol, and the use of key informant interviews and focus groups as part of the cultural adaptation process. The third step - verifying the psychometric properties of the instrument - has not been addressed in this paper although an earlier study in Pakistan has examined the issue.1

The Translation Process

The translation process takes place at three levels, and equivalence between the original and translated versions needs to be considered at every level:2

a. The linguistic or semantic level:

Linguistic or semantic equivalence means to retain a similar meaning of a measure in the original and in the translated version. The translated meaning should remain as near as possible to the original meaning.

b. The technical level:

Technical equivalence refers to both the technical features of the language (i.e., grammar, tense, question length, acceptable level of abstraction), and their relationship to the sociocultural context; the feasibility of the nature and mode of questioning used in the original and translated versions; for example self-rating questionnaires developed for educated Western subjects may be problematic to administer to illiterate rural subjects.

c. The conceptual level:

This refers to the need to obtain an identical meaning of concepts, which may have different cultural understandings. For example, the concept of a "miracle" may differ in cultures.

To these three can be added the 'comprehension level'. While it is important for the translation to be equivalent at these three levels, it is equally important to ensure that the target population understands the translated material as easily as the source population for whom the original questionnaire was designed.

Characteristics of the Translator (s)

These language, culture and comprehension related equivalencies are key issues in translations of questionnaires. The translator needs to have an excellent command over technical and colloquial aspects of both the original and the translated language. They also need to have an in-depth insight of the cultures in question so that they can relate this to the terms and concepts used in the questionnaire. Finally, they need to be highly qualified in the area investigated by the questionnaire and have the necessary technical and scientific background in order to understand the concepts and constructs used.
Not many translators will meet these rigorous criteria. When a translator does not have all these characteristics, strategies can be devised to compensate for this, including focused group discussions and holding in-depth interviews with key informants and experts within the community to obtain a better understanding of cultural issues both in the original and target cultures.

Key informant interviews

This entails interviewing individuals whose social positions make them influential and bring them into contact with a large number of people. The key informants need to have varied backgrounds and can be selected on the basis of race, age, sex, socio-economic status and profession.3 These informants can be a rich source of information on cultural and language issues. This information can then form the basis of discussion for focus groups for further refinement and verification of concepts.

Focus groups

The focus group interview is a qualitative research method. It was originally used for market research with the purpose of staying close to consumers, their changing attitudes, emotions, thoughts and buying impulses.4 The method is now gaining credibility in social sciences research. It is a focussed group interview or an arranged meeting among a selected group of people. Its aim is to uncover important dimensions of a given problem, experience or other phenomenon. In translation of questionnaires, it can be useful to give a wide range of information and uncover important understanding of the problem to be addressed. It is possible to address questionnaire problems like the readability of a

measure, the construct of a concept, or the understanding of items, and at the same time address the issues of acceptability of the content of the questions. The size of the group is 8-10 participants. The discussion is led by a moderator who has a list of key questions around which the discussion revolves. Developing the discussion outline requires careful thought and planning. Key informant interviews can provide the basis for the discussion outline.

Questionnaire

The SRQ (Self-Reporting Questionnaire)5 is a brief self-assessment or interviewer administered scale (though methods of administration must not be mixed in a single study) designed to screen for psychiatric disturbances, particularly in developing countries. It can be used as a first research instrument in a two-stage detection procedure. The first stage refers to the use of a screening test so that probable cases will be detected and in the second stage a clinical diagnosis is made using a second stage diagnostic instrument and/or a clinical interview.

The SRQ (Table) consists of twenty items designed to identify 'mental distress' (particularly neurotic disorders including depression, anxiety related disorders and somatoform disorders). Each item has a yes/no answer. The time span refers to the individual's feelings over the past 30 days. Each item is scored 0 or 1. A score of 1 indicates that the symptom was present during the past month; a score of 0 indicates the symptom was absent. The maximum score is therefore 20. SRQ takes only minutes to administer.

As the SRQ is a measure of the mental well being of the respondent, the translation is not a straightforward procedure. Mental symptoms can be variously expressed, depending on the culture and language of the respondent. The subtleties and nuances of language and culture have to be taken into account in the translation procedure.

The translation protocol

A general protocol was developed for the translation of instruments into Urdu, describing the procedure for each step in the translation process.

1. Initial translation:

This is ideally carried out by clinically experienced health professionals whose native language is Urdu and second language English. It is of great benefit if at least one of the translators has good knowledge of Western culture, ideally by having lived and worked there. A useful reference at this stage is Jalibi's English-Urdu dictionary.6 Each item should be carefully considered at the four levels described above. A note should be made of problematic items.

2. Key informant interviews:

The second step involves consultations with a wide range of key persons across the community; these can include teachers, health professionals, and laypersons. If the questionnaire is to be administered to illiterate persons, it is important to get their opinion about the comprehension and cultural relevance of the items. Both genders should be

represented, as should be a range of ages on which the questionnaire is to be used. These key informants can be given the translated questionnaire and asked to comment on each item, especially those that had proven problematic in the first translation.

3. First Revision of Translation:
In light of the information provided by the key informants, the translation should be revised.

4. Back Translation:
The revised translation should then be back translated into the original language by a third translator who has excellent command over both languages. This translator can be a non-technical person.

5. Second Revision of Translation:
The original translators should then compare the back-translation with the original version. Differences should be discussed and seriously disputed items changed. Problematic items should be noted for discussion in the focus group.

6. Focus group:

This phase can also be described as the pilot phase. A small group of people (8-10) representing the study population should be assembled. They can be paired into 4-5 groups and each pair asked to complete the questionnaire and to note down problems of comprehension, language, and cultural relevance. They should be encouraged to give suggestions. One translator should act as a moderator while the other should take notes, and the whole exercise should be carried out in a structured manner.

7. Final Revision of Translation:

The focus groups' remarks should be discussed between the researchers and back-translator. Problematic items should be revised, leading to the final version of the translation. If there are substantial changes from the second revision, step 5 should be repeated.

Application of protocol to the SRQ:

Two authors (A.R and Z.I) undertook the initial translation. As our research objective was to use the questionnaire on mothers in a rural community, the items were discussed with 2 lady teachers, 4 Lady Health Workers, 2 uneducated traditional birth attendants, and a doctor working in the same community. The translation was revised in light of their suggestions. A teacher of English at a local school then back translated the questionnaire. Both versions were compared by A.R and Z.I for any discrepancies, which were removed. A convenience group of 8 mothers from the study area was assembled in the community and administered the questionnaire in pairs. Non-literate mothers were read out the questions. Each item was discussed and difficulties noted. The final revision was then carried out.

Results

Presented below is a summary of the conclusions derived from the above process applied to the translation procedure. Items 3, 5, 7, 9, 12 and 20 were straightforward, and no major changes were made after the first step. The remaining items were modified at later stages of the procedure.

Item 1:
This item explores headaches as a somatic manifestation of mental distress. These headaches are usually non-specific in nature, hence the use of the plural term "dardein" instead of the singular "dard".

Item 2:
The term "bhook kharab" has been used for "poor appetite". Although the nearest translation for poor appetite is "bhook mein kami", the former is a more colloquial term that indicates appetite loss associated with mental distress.

Item 4:
This measures the heightened state of arousal that is a manifestation of anxiety disorders. The literal translation of "frightened" is "darr jana", but in order to make it a more semantically appropriate translation, two words,"darr ya seham jana", were used to depict fright related to heightened arousal.

Item 6:
There are many words that depict worry, tension, and nervousness associated with anxiety and mood disorders. The first two stages of the procedure identified generic terms such as "fikermand", "pareshani", "waham", for worry; "tashweesh", "baichaini", "baiarami", "aasabi or zahni tanao" and even the word "tension" is used among many non-English speaking people to mean "tension" used in the questionnaire; "Ghabrana", for "nervousness". Key informant interviews helped us select the words that appear in the final version, and these were endorsed by the focus group.

Item 8:
This item presented considerable problems in translation. It measures the disturbance in concentration and cognition associated with depressive disorders. We could not find an exact substitute for the term, "clear thinking" in colloquial Urdu, and the nearest semantic and technically equivalent term that was acceptable in back translation was "wazay soch bichar". This was acceptable to the focus group.

Item 10:
The literal translation for this item was "Kya aap ko maamool sey zyada rona aata hai?" However after consultation with key informants and focus group, this was changed to "kya aap ko abb baat baat par rona aata hai?" The back translation for this was, "Do you now cry frequently?" which was thought to be acceptable. It was felt that the latter

translation was more sensitive to the tearfulness associated with mood disturbance than the literal translation.

Item 11:
Again, the original literal translation was changed to a more colloquial sentence that was semantically equivalent but easier to understand.

Item 12:
The translation was straightforward. However, it was noted that there could be important gender-based differences in the interpretation of this item. Women, especially in rural areas, are often not as involved in decision-making, and some women participants of the focus group suggested that this might be an irrelevant item for this group. It was suggested that a better option might be to ask them if they found it difficult to make routine everyday decisions. This deviates considerably from the original item, and we decided to stick to the original version.

Item 13:
There is no equivalent term for the phrase' "work suffering" in Urdu, and semantically equivalent terms include "kaam mutassir" or "kaam mein harj". The latter is more suitable in terms of comprehension.

Item 14:
This item measures negative cognitions about ones sense of purposeful participation in life, and is a symptom associated with depression. In many agrarian communities, work is often a joint endeavour of many family members rather than just an individual's, and purposeful participation often simply means ones physical presence at the place of work. Many subtle changes had to be made to this item to increase the sensitivity, comprehension and relevance to the local population.

Item 15:
This item measures anhedonia or loss of pleasure in everyday activities, associated with depression. The translation is straightforward.

Item 16:
A sense of worthlessness is part of the syndrome of depression. The phrase "beyqadar-o-keemat" was most appropriate for this symptom.

Item 17:
In Pakistani culture, suicide is a taboo subject and this item can be problematic in administering. The translation process did not pose any particular problems, but it was felt that in subjects who do not score positive on other features of depression, the item could be skipped.

Item 18 :
We used the term "thaki thaki" instead of "thakawat", as it was felt that the former conveyed more affect associated with depression than the latter.

Item 19:
This item measures gastric somatisation that often accompanies anxiety and depression. The more literal word for stomach is "maida", but "pait mein bey arami" is a more colloquial term that is used to describe these uncomfortable abdominal feelings.

Discussion

The above process illustrates that translation of health questionnaires should not be seen as a straightforward process. An earlier validation study of the SRQ by Minhas and colleagues1 showed that at their best cut-off point of 4/5, the SRQ had a sensitivity of 63% and specificity of 77%. A low sensitivity means that the SRQ failed to detect a number of possible cases. Minhas concluded that the low sensitivity of the SRQ found in their study was because many items were not culture specific. The authors, however, gave no details of how they translated and adapted the questionnaire to the target culture. Our preliminary data on a postnatal sample of about 500 women indicates that at our best cut-off point of 8/9, the SRQ has a sensitivity of about 80% and specificity of 75%. The improved psychometric properties may have arisen because of the translation and adaptation process.

However, it needs to be pointed out that the cultural adaptation procedure carried out in this study was focused to women of childbearing age. This was in line with our need to develop a screening instrument for this particular population. Therefore this particular translation may not be applicable to other groups, and would need to be reviewed prior to generalised use. However, the whole procedure need not apply in revising the translation. This will depend on the degree of cultural overlap between the groups, and is elaborated below.

Unlike the SRQ (which was developed as an international instrument), many health questionnaires for Western populations may contain items that are not culture specific. For example, questionnaire screening health problems in women of childbearing age might enquire about alcohol intake during pregnancy. This item would be culturally irrelevant for a conservatively Muslim rural female population in Pakistan, and ought to be omitted. This issue is especially important for questionnaires measuring psychological, attitudinal or life-style issues related to health.

A useful framework to guide the approach adopted for translation is based on the degree to which the original and target cultures overlap for the item in question:3

a) If there is complete overlap, a simple approach to translation can be adopted (steps 1,4 and 5 only);

b) If there is considerable overlap, some consultation with the target community is required (steps 1-5);

c) If there is less conceptual overlap, the whole process needs to be applied (steps 1-7);

d) Finally, there can be items for which there is no conceptual overlap. Translation of such items will not be possible (or relevant) and they should be omitted altogether.

Refrences

1. Minhas FA, Iqbal K, Mubbashar MH. Validation of self-reporting questionnaire in

primary care setting of Pakistan. J Pak Med Assoc 199545:60-4.
2. Sartorius N, Kuyken W. Translation of health status instruments. In: Orley J, Kuyken W,(eds). Quality of life assessment in health settings. Vol I. Berlin, Springer 1994.
3. Tremblay MA. The key informant technique: a non-ethnographic application. In: Burgess R.D, (ed). Field research: a sourcebook and manual. London: Routledge 1989, pp. 98-104.
4. Greenbaum TL (ed). The handbook of focus group research 2nd ed. Thousand Oaks, Sage 1998, pp. 1-15.
5. World Health Organization. A user's guide to the self reporting questionnaire (SRQ). WHO/MNH/PSF/94.8. Division of Mental Health, Geneva: WHO, 1994.
6. Jalibi J (ed). Qaumi English-Urdu dictionary. Islamabad: Muqtadra Qaumi Zaban, 1979.