Translation and cross-cultural adaptation of a mental health battery in an African setting

J Smit¹,², CE van den Berg¹, L-G Bekker², S Seedat¹, DJ Stein¹,³

¹MRC Unit on Anxiety and Stress Disorders, University of Stellenbosch, Cape Town
²Desmond Tutu HIV Centre, Institute for Infectious Diseases and Molecular Medicine, University of Cape Town, South Africa
³Department of Psychiatry and Mental Health, University of Cape Town

Abstract

Background: Standardised measuring instruments are increasingly used in psychiatric research cross-culturally. These instruments are considered to be culturally equivalent when all forms of biases, or social norms specific to the culture of origin, have been removed.

Objectives: To describe the qualitative process of selection, translation and cultural adaptation of a mental health battery for use in a Xhosa-speaking community that is, as far as possible, ‘culture-free’ or equivalent.

Method: Informal discussions were held with key members in the community to determine what would be considered as appropriate for the community in respect of psychiatric screening instruments. Existing rating-scales for depression, alcohol abuse and posttraumatic stress disorder that would meet these criteria were identified and then translated from English into Xhosa. Cultural equivalence was achieved by combining methods of back-translation, committee consensus approach and decentering. Discussions during the committee consensus meetings were recorded and categorized into themes. Two themes emerged: (1) issues related to the attainment of semantic equivalence and (2) broader problems inherent in cross-cultural research.

Results: Issues related to individual questionnaires included the use of terms to describe emotional distress cross-culturally. Broader issues related to the translation process itself included the form of language to be used, time-frames, and the use of Likert-scales. It also demonstrated the problems inherent in the categorization of emotions.

Conclusion: A method of combining a group approach, back-translation, and decentering was effective and efficient in this context for establishing content and semantic equivalence. Cross-cultural adaptation can never completely remove all forms of bias from a research instrument, but such limitations should be acknowledged and openly discussed, rather than hidden or ignored.

Key words: Cross-cultural; rating scales; research; psychiatry; South Africa

Introduction

Large-scale epidemiological studies have shown that psychiatric disorders occur across different societies and cultures¹ suggesting that some common underlying biological mechanisms may be at work. However, discrepancies in global prevalence rates of the major psychiatric disorders suggest that variations exist in how these disorders are expressed¹;². This variability may, in part, be explained by the fact that emotional distress is experienced and communicated differently in different social contexts³.

Examining psychiatric disorders in different cultures is controversial and best explained by the two positions within cross-cultural psychiatry, namely the universalistic and relativistic positions. The universalistic position argues that emotions are biological phenomena; the result of neurophysiologic processes in the limbic system and are a limited repertoire of universal emotional experiences.⁴ This position is based within biomedicine and focuses on the categorization and labeling of syndromes.⁵

The relativist position on the other hand asserts that emotional expression is socially constructed and therefore unique to a particular historical, social and cultural system.⁶ It argues that tools developed in one setting cannot capture the idiosyncratic ways that other cultures express emotional distress as it ignores the context within which a person lives and experiences the world, a position held by ethnographic and anthropological studies.⁵

Both positions have their limitations: Ethnographic studies can provide in-depth knowledge about emotions, but are limited to small samples and do not facilitate large-scale comparisons.⁷ While they can give insights into the differences between groups, they tell us little about similarities.⁸ This runs the risk of reifying difference by ignoring the influences of acculturation and cultural assimilation.⁹ Universalistic arguments, on

---

Correspondence:
S Seedat
MRC Unit on Anxiety and Stress Disorders
PO Box 19063, Tygerberg, 7505
Cape Town, South Africa
Ph: +27-21-938-3974, Fax: +27-21-933-5790
E-mail: sseedat@sun.ac.za

African Health Sciences Vol 6 No 4 December 2006
the other hand, have been criticized as imperialistic for ignoring cultural difference and their insistence on using concepts evolved within a Western context as a blueprint for understanding other cultures.\(^7\)

These two positions overlap with the emic and etic orientations, a concept from linguistics that deals with the origin of concepts under investigation.\(^2,7\) "Emic" refers to the meaning a particular group attaches to a specific notion, and is comparable to the relativistic position. "Etic" refers to a description of phenomena independent of meaning, and is comparable to the universalistic position.\(^7\) Recently, theorists have attempted to integrate these two positions by combining etic and emic orientations. Concepts and descriptions derived from studies in anthropology (an emic orientation) are incorporated into measuring instruments, a predominantly etic activity.\(^3\)

Cross-cultural adaptation of psychiatric research instruments

Standardised measuring instruments are increasingly being used in psychiatric research cross-culturally to diagnose and measure psychiatric disorders.\(^1,8\) These instruments have the advantage of being easy to use and interpret.\(^7\)

Cultural adaptation of research instruments aims to achieve, as far as possible, research tools that are ‘culture-free’ or culturally equivalent. An instrument can be considered culturally equivalent when all forms of biases, or social norms specific to the culture of origin, have been removed.\(^10\) The process whereby equivalence is established can be seen as an example where an emic approach penetrates a predominantly etic research activity.\(^11\)

Van de Vijver & Poortinga\(^10\) list three types of biases that can impact on cross-cultural research, namely construct bias, method bias and item bias. Construct bias occurs when the concept under investigation differs substantially across cultural groups. Method bias occurs when the methods used to examine a construct are culturally unfamiliar or inappropriate. This occurs when the type of enquiry is unfamiliar, such as paper-and-pencil tests (in cultures where oral traditions predominate) or interviews on topics that are considered taboo. A second example of method bias may be idiosyncratic response styles, i.e. the tendency to use extreme scores or high non-response rates, and can occur when self-disclosure is uncomfortable. Item bias refers to discrepancies in certain items of an instrument. It may occur when a specific item does not fit the description of a concept under investigation in the target culture.

Flaherty\(^1\) described a model for the selection, adaptation and validation of cross-cultural research ins-

truments. This model has also been used in other studies where it provided a guide for cross-cultural instrument development.\(^12\) The model consists of two stages. Instrument selection is the first stage and refers to the process whereby the literature is searched for an appropriate instrument. The following criteria have been suggested as a benchmark for decision-making: (1) an instrument that has already been used in different cultures, (2) an instrument that has been developed to examine the construct under investigation, has known psychometric properties, but has not been used cross-culturally and, (3) an instrument where the psychometric properties are unknown, but with high face validity that requires further field-testing.

Once an instrument has been selected, the authors describe the process an instrument needs to go through to ensure cultural equivalence. This process has five dimensions:

1. **Content equivalence** - ensures that each item of an instrument is relevant to that particular culture and will remove item biases from the instrument;
2. **Semantic equivalence** - ensures that the meaning of each item is maintained after translation into the other culture;
3. **Technical equivalence** - removes method biases from an instrument;
4. **Criterion equivalence** - refers to the ability of an instrument to accurately distinguish between those with the construct from those without the construct. It is also termed predictive validity and is measured by statistical procedures to establish instrument sensitivity and specificity;
5. **Conceptual equivalence** — is obtained when all other forms of bias have been removed and an instrument is considered a valid tool to capture the specific construct regardless of culture. Each dimension is mutually exclusive and any instrument may be equivalent in some, but not all, of these dimensions.

**Methods to establish cross-cultural equivalence in research instruments**

A number of methods have been described to establish cross-cultural equivalence. These methods are outlined below and are particularly useful where an instrument already exists in one culture and needs to be adapted for use in another.\(^11\)

(i) **Back-translation**

Brislin\(^13\) recommended back-translation as the method of choice, as it gives the researcher control over the translation process especially when the researcher does not understand the target language. The process involves two bilingual translators. The first translates the original instrument into the target language, and the second translates that version back into the original language. The researcher now has two original language versions and a
comparison can highlight translation difficulties. The first (original) instrument can now be changed in the light of these discrepancies and the process is repeated until the two versions are the same or deemed equivalent.

(ii) Decentering
Decentering is often used together with back-translation. It refers to the process whereby both the source and target languages are deemed equally important in the research endeavor. If the two instruments (the original and the back-translated version) are not identical, bilinguals can make changes to both the original and the translated instrument to obtain equivalence. However, too many changes to the original version may impact on its psychometric properties. This may suggest that the instrument contains construct biases and is not appropriate as a cross-cultural instrument.

(iii) The bilingual approach
In this approach bilingual participants complete the instrument in both languages. Discrepancies in responses between the two versions provide information about specific items that have not translated adequately.

(iv) The committee approach
A group of bilinguals translates from the source to the target language and discusses any discrepancies that may arise during the translation process. However, conflict during translation may not be seen as an opportunity for debate, but as a tension to be overcome, resulting in an instrument where conflicting items are glossed over or avoided.

This paper describes the process of instrument selection and establishment of cross-cultural equivalence in a Xhosa-speaking, peri-urban community in South Africa. We discuss issues related to instrument translation as well as issues specific to each instrument. The method described earlier by Flaherty and colleagues was utilized. The aim was to ensure content and semantic equivalence. This was achieved by combining the methods of back-translation, decision-making by a committee and decentering, a comprehensive method for examining content and semantic equivalence. It also provides an opportunity to explore issues that may impact on method and construct equivalence.

Methods
Study Setting:
This qualitative and descriptive study was conducted in a peri-urban settlement outside Cape Town, South Africa. The community is home to an estimated 10,500 Xhosa-speaking residents of low socio-economic status and high unemployment. The site was recently chosen by the South African AIDS Vaccine Initiative (SAAVI) as a phase III HIV vaccine trial-site. Permission to conduct this research was obtained from the ethics committee of the University of Cape Town.

Study context
This study formed part of a much larger phase III HIV vaccine preparedness project. South Africa’s history of political violence in informal settlements has created a context where violence is endemic. This includes high levels of sexual, physical and domestic violence. Violence has been shown to have a negative impact on physical and mental health and has been associated with psychopathology in victims.

Depression, substance abuse and post-traumatic stress disorder were selected for assessment, as these disorders have all been associated with increased sexual risk-taking and are considered risk-factors for contracting HIV. High prevalence rates for depression, PTSD and hazardous alcohol-use have also been observed in community and clinic studies in South Africa.

A series of brief rating-scales were selected to be used as self-report measures. The aim of this battery was to develop a series of questionnaires that would assess levels of depression, alcohol abuse and PTSD in a Xhosa speaking community. It is important to stress that the battery was not intended to provide accurate prevalences of psychiatric morbidity, but to answer preliminary questions about mental health in this community.

Measuring instruments
Process of instrument selection:
The process of instrument selection followed the method outlined by Flaherty. The process involved two steps. First, informal discussions held with key members in the community revealed that mental health conditions were considered to be a major problem, with depression, alcohol and marijuana abuse most prominent. The community was also familiar with pen-and-paper tests as a number of research studies have taken place. Secondly, key requirements for the proposed instruments were identified by the research team consisting of two psychiatrists and a psychologist working on the project. Six requirements were identified: The battery (1) should be appropriate for use in a community setting, (2) it should be appropriate for a Xhosa-speaking population, (3) it should be easy to understand and use, (4) it should be brief with few items, (5) it should be easy to score by an independent rater, (6) scoring should be simple, with cut-off scores indicating clinical significance.
Existing screening instruments that would meet these criteria were identified. Inclusion was based on: (1) whether the measure was available (quickly and cheaply), (2) whether it had been validated for use in community settings, (3) whether it had been validated as a self-report measure in non-western or African settings, (4) whether it had been validated or previously used in South African communities, and if not validated, (5) whether it had been used in South Africa or other non-western sample with adequate results.

The final mental health battery consisted of three rating scales:

- The Center for Epidemiology Depression Scale (CES-D) to measure depression.\(^{28}\)
- The Alcohol Use Disorders Identification Test (AUDIT) to measure alcohol use disorders.\(^{29}\)
- Harvard Trauma Questionnaire (HTQ) to measure Post-traumatic Stress Disorder (PTSD).\(^{30}\)

The battery also included a Life Event Checklist (LEC) from the PTSD Checklist (PCL)\(^{31}\) and two questions on drug use adapted from the Mini International Psychiatric Interview (MINI).\(^{32}\)

Instrument translation

Questionnaires were translated from English into Xhosa by a clinical psychologist fluent in both languages. Back-translation and a ‘committee consensus approach’ were combined. From July-September 2003, six Xhosa-speaking HIV educators-in-training who were proficient in English and resident in the community assisted the primary author (JS), a clinical psychologist, in adapting the battery into culturally attuned and semantically equivalent screening instruments. Their gender, age and educational level were reflective of the target group considered at risk for HIV. Focus group sessions took place once a week and lasted about 3 hours each. The educators were given both the original English questionnaire and the translated Xhosa version. Each question was read in English and the meaning and intent explained and compared with the translated item. The team then discussed whether they agreed with the translation. If they disagreed, this was further interrogated until the best translation (the version everyone agreed on) was established.

Decentering was also used. Changes were made to the original English questionnaire in cases where a) translation into Xhosa proved difficult, b) the English was not easily understood or c) where semantic equivalence could only be achieved if the English phraseology was adapted to fit in with the Xhosa version. Detailed field notes of all group discussions were taken by the first author (JS). Only conversations in English were recorded. Most of the heated debates took place in Xhosa and these were summarized to the first author in English. Most of the nuances of these debates, unfortunately, remained unrecorded. A descriptive analysis of the recorded discussions is presented.

Results

Several challenges emerged relating to both the process of translation and the attainment of semantic equivalence.

The process of translation

a) Appropriate form of Xhosa to be used

One of the main concerns during the focus group discussions revolved around the right to speak for all Xhosa people and the form of Xhosa best-suited for the battery. The group considered research a high-status activity and was concerned that they were speaking for all Xhosa people when they themselves did not have status within the community. They were also concerned that they were not able to provide the form of Xhosa (formal Xhosa mostly spoken by the elderly or rural people) which was considered representative of high-status activities. None however were comfortable with formal Xhosa and mostly used conversational Xhosa which consists of a mixture of dialects and other African languages. This form was considered a ‘lower’ or less pure form of the language and they had many reservations about using this form despite our assurances that the language should be informal and representative of the community.

b) Problems with time-frame

There is no immediate past tense in Xhosa. Both the CES-D and the HTS ask about emotional events that have occurred within the past 7 days. The term “ndizive” (I had) refers to the past tense, but includes a time-frame longer than 7 days and excludes current symptoms. The term “ndiziva” (I am) on the other hand only refers to symptoms experienced on the day of questioning. It was decided to stay with the past tense “ndizive” and to highlight in bold “during the past week” to frame the time-period under investigation.

c) Increasing the amount of visual display

While paper-and-pencil tests were familiar to this community, too many choice options were considered confusing, especially in the context of providing a subjective measure of distress. Likert-scales were also considered difficult to interpret as it was hard to distinguish between categories such as “most of the time”
and "some of the time". Visual analogies, to describe levels of distress in terms of numbers were suggested. For the HTS, we ranked the Likert categories from 1 to 4 to show the increasing order of distress indicated by 'most', 'often' and 'sometimes'.

Specific issues pertaining to semantic equivalence
Substitutions for words and phrases on the four instruments, as suggested by the group participants, are highlighted below:

**CES-D:**

a) The concept of sadness: Question 2 of the CES-D probes for the presence of sadness or 'the blues'. The term 'blues' was not familiar to the group, so 'sadness' was used for both language versions. "Sadness" was initially translated into Xhosa as "ingxaki" which means "to worry about problems" or "to have problems" a term which may be confused with anxiety rather than depressed mood. The wording was changed to refer to an internal state, namely "ukukhathazekile" which means "to feel bad inside". There was concern that this term could be confused with somatic complaints, or other medical conditions such as the flu. However, some members in the group felt that in the context of emotions it referred to "matters of the heart" rather than physical conditions.

b) Self-esteem and worthlessness: Question 4 of the CES-D probes feelings or worthlessness and self-reproach with the statement "I felt that I was not as good as other people". In Xhosa a "good" person refers to someone who leads a moral life rather than self-esteem. However, others felt that the two concepts were linked – a good person was one with positive self-esteem. The two concepts, self-esteem and worthlessness created confusion within the group and needed careful explanation. The team concluded that these feelings were best described in Xhosa by the term "Ndizive ndingalunganga nje ngabanye abantu" and literally translated means "to undermine yourself".

c) The term "concentration": Question 5 examines concentration difficulties with the statement: "I had trouble keeping my mind on what I was doing". The team felt that the initial Xhosa version translated to "losing your mind" and meant going mad, rather than concentration. This resulted in considerable debate and both Xhosa and English versions were felt to be inadequate. The English version was amended to include the word "concentration" in brackets, while the Xhosa version was amended to "Ndinengxaki yonkungakwazi ukuzikisa ingqondo yam kwinto endiyenzayo" (I have a problem of not being able to apply my mind to whatever I am doing).

d) The term "crying spells": The term "crying spells" created much confusion and both English and Xhosa version had to be changed to "I cried frequently" for both language versions.

**Audit**

This questionnaire was easy to understand and very few difficulties were encountered. The only exception was item 10 of the questionnaire which asks about other people commenting on one’s drinking - comments on other’s drinking behaviour were very common and was not considered to be an adequate indicator of pathological drinking behaviour.

The drug-use questions of the MINI created some cross-cultural difficulty. Apart from marijuana, none of the other terms were familiar and they did not have Xhosa equivalents. Ecstasy tablets were known as "8-hour-pills" to those who frequented nightclubs and discotheques in the city. This was considered a rare occurrence, as the city centre is more than an hour’s drive by car, with no public transport available in the evenings. The group suggested that the English terms be used for both versions as those who were familiar with the drugs would also be familiar with the English terms.

**Life Event Checklist (LEC)**

a) The concept of trauma: The group was unfamiliar with the English term "trauma" when used in the context of emotions, as they thought that it referred to physical damage. The word was thus changed to "terrible things that can happen to you" in both English and Xhosa versions.

b) The term "natural disaster": Question 1 on the LEC asks whether a person has experienced, witnessed or heard about natural disasters (including floods, drought, and hurricanes). The Xhosa language has terms for hurricanes, lightning, floods and storms, but no term that encompasses all of these into one, such as "natural disaster" in English. After much discussion a descriptive term "inkanyamba" was used. This means "the great destroyer" and refers to a snake in mythology which moves through and ravages villages.

c) The phrase "human suffering": Question 12 includes the term "severe human suffering" as a life event that may constitute a diagnosis of PTSD. However, the group felt that the term may not have validity for this community, where severe socio-economic deprivation and hardship were part of daily life. The term was considered too broad.
to have any real meaning and therefore removed from the questionnaire.

Harvard Trauma Questionnaire (HTQ)
a) The phrase “feeling on guard”. The phrase “feeling on guard” was not understood. This was changed to “mistrusting others” in the English version so as to translate into Xhosa. However, the former may capture anxiety while the latter may capture paranoia.

b) The phrase “feeling hopeless”. In English it refers to a psychic state of despair. However, in Xhosa it may refer to an actual skills deficit (e.g. being hopeless in doing a certain task), namely, “you cannot do this job, you are hopeless”. This was changed to “I felt there was no hope” in the English version, and translated into Xhosa as “Usive ungenathemba”.

c) The concept of detachment: Detachment is probed in the HTQ with the phrase “Feeling as if you are split into two people and the one of you is watching what the other is doing”. There was no Xhosa word that captured the psychiatric condition of detachment. Eventually it was described by the term “ukususa” which literally means “to be taken away or ignored” and used within a context of emotions was considered the closest approximation to such a psychic state within the Xhosa language.

Discussion
This paper outlined the process of selection and adaptation of a mental health battery for use in a cross-cultural setting combined with back-translation, consensus by committee and decentering approaches to establish content and semantic equivalence. To our knowledge this is the first time that this combination of methods has been used to ensure cross-cultural equivalence for mental health measures in an African setting.

Overall, this study highlighted a number of issues inherent in the translation process. The most prominent issue related to language and the power and status associated with research. The question about who was qualified to make decisions ‘and speak for all Xhosa people’ and the tension between high status language and perceived low status forms was a prominent concern throughout the translation process. This was also found by Drennen who understood this in terms of the power-relationship inherent in research and the use of language to negotiate this. Notwithstanding, this may also refer to the process of acculturation, and this community’s struggles to maintain their identity whilst simultaneously adapting to an urbanized landscape. The degree to which western psychological concepts are understood may also be related to acculturation and could account for some of the translation difficulties. Moreover, it could explain the finding that those who had a closer connection with the city and its nightlife were more familiar with terminology used to describe drug-use.

In terms of instrument design, confusion with Likert-scale categories was overcome by using numerics to indicate increased levels of distress. Cross-cultural use of Likert-scale categories is known to be problematic and demonstrates the difficulty in quantifying distress.

This study similarly demonstrated the difficulty in using local idioms such as ‘blues’ and ‘feeling on guard’ in cross-cultural research, in particular the translation of the term ‘blues’ has also been noted in other studies. Blues, within Western psychiatry, is akin to sadness and is understood as a symptom of depression. However, when translated into Xhosa it refers to having worries, and reflects anxiety. Similarly, ‘feeling on guard’ refers to an internal state of hyperarousal, but when translated it refers to mistrust of others, or paranoid states. Each of these concepts are understood and treated very differently within Western psychiatry. Leff criticizes this differentiation of emotions into discrete categories within biomedicine. He argues that such differentiation is artificial and part of a mechanistic world-view, where emotions can be taken apart and analysed in isolation.

This mechanistic world-view is also apparent when exploring the difficulty in separating feeling states from external circumstances. Negative emotional states, such as worthlessness, hopelessness and lack of self-esteem, in this study, were explained as the result of external hardships. This was also found in other cross-cultural studies. The idea that emotions can exist as psychic phenomena in the absence of actual external causes was unthinkable. Swartz provides an in depth discussion of the relationship between emotions and deprived social circumstances in southern Africa. The tendency to separate emotions from its context within Western biomedicine runs the risk of denying the very real economic deprivations and serves to pathologise individual responses.

This tendency of biomedicine to classify, isolate and analyse emotions as observable and discrete objects is inherent in positivism, and its implication within cross-cultural research has been much debated. Lutz explained the tension between emotions and context in terms of the disease versus illness model. Disease refers to the signs and symptoms, which may, or may not correspond to psychiatric models, while illness refers to the individual experience of those symptoms, and is
fully understand and intervene within different cultures illness models needs to be incorporated into the design and validation of research-instruments. Decentering via a committee approach can be seen as a useful method towards such a goal.

Lastly, the AUDIT questionnaire presented the least amount of debate, possibly because it focuses on behaviour rather than complex emotional states. Of interest is that guilt in the AUDIT was not considered a problematic term in this community and was translated as “to blame yourself, or to have remorse”. This is in contrast to other studies which found that guilt, in relation to depression was less often experienced in Africa, compared to Britain and North America and may be related to how the self is experienced in relation to others. However, in this study, guilt was explored within the context of alcohol abuse and relates to actual behavior. In township communities such as this one, alcohol abuse is closely related to violence and guilt or remorse, may be related to actual events of harm, rather than a diffuse concern for others.

A major limitation of this study is that cross-cultural equivalence was explored with a group of participants who may not be representative of the community. However, within the remit of the larger HIV-study, those involved in the translation process represented the group most at risk for HIV.

This study has also highlighted the importance of acknowledging the limitations inherent in the use of an instrument designed in one culture for use in another. It will never be possible to remove all forms of bias when adapting Western rating-scales for use cross-culturally. The aim is to remove as much bias as possible by using a combination of methods. Even then there are limitations as to how much can be changed, without impacting on the validity and reliability of both the target and source measures. It is our view that researchers should be mindful of these limitations. Further field-testing procedures are also required to examine construct and technical equivalence.

Conclusion

This study describes the process of cross-cultural adaptation of a mental health battery. A method of combining back-translation, consensus by committee and decentering was used to establish content and semantic equivalence. We also report on the difficulties that arose during this process and therefore highlights the limitations inherent in cross-cultural research. Biases can never be completely removed from the research endeavor but should be acknowledged and discussed openly, rather than hidden or ignored. In terms of the instruments discussed here (CES-D, AUDIT and the HTQ), more work is needed to establish their construct and criterion equivalence. This will entail further quantitative validation and factor analysis. Nevertheless, the model described here provides a format for the ‘fine-tuning’ of existing mental health screening instruments for use in communities of different cultural, language and ethnic backgrounds.

Acknowledgements:
The authors wish to thank the following persons for their assistance; Prudence Mthimunye, Lububalo Vellem and Patricia Mhlambiso. This research was funded by the South African Aids Vaccine Initiative (SAAVI).

References


