Adapting health behavior measurement tools for cross-cultural use

Rachel C. Vreeman1,2,3*, Megan S. McHenry1 and Winstone M. Nyandiko2,4

1Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, USA.
3The Regenstrief Institute, Inc., Indianapolis, IN, USA.
4Department of Child Health and Paediatrics, Moi University School of Medicine, Eldoret, Kenya.

*Correspondence: rvreeman@iupui.edu

Abstract

**Background:** Addressing health disparities in the global community requires awareness of how cultural differences in beliefs, traditions, norms, and values shape health problems and behaviors.

**Review:** This paper reviews methods of assessing health behaviors, how these assessments may be affected by cross-cultural differences, and methods of adaptation of health behaviors across cultures. We describe the methods used in appropriate translation processes and pilot-testing for health behavior assessment tools. We also discuss ways to limit literacy demands and incorporate qualitative interviews.

**Keywords:** Cross-cultural comparison, health behaviors, outcomes assessments, evaluation methodology

Introduction

Many health indicators, such as mortality statistics, heights/weights, and laboratory results, can be measured directly and quantitatively, other aspects of health, such as health behaviors, attitudes and motivations, cannot be measured directly. For example, beliefs about the causes of illness, fears of stigma or discrimination, or a family’s sense of connectedness to the community may all influence the consistency with which they take medications, report depressive symptoms, or seek clinical care [1,2]. Measuring these relevant, but often hidden factors, may require the assembly of a selection of indicators to represent the health outcome. These indicators must reflect problems of social concern or core values of a culture shaping health behaviors [3].

Despite expanding work in cross-cultural health behavior research, few systematic approaches to the cross-cultural adaptation of health behavior measurement strategies and measurement tools exist. In systematic reviews on adherence to HIV therapy, for example, no existing measures had been validated in the resource-limited settings where the majority of the world’s HIV-infected population lives [4,5]. Our objective is to summarize methods of assessing health behaviors, how cross-cultural differences may affect these assessments, and methods for adapting assessments across cultures.

Review of cross-cultural adaptation for health behavior measures

Methods of assessing health behaviors

Health behaviors can include preventive behaviors, illness behaviors taken when an individual perceives him-or herself to be ill, and sick-role behaviors undertaken for the purpose of getting well [6]. In a broad sense, health behavior refers not only to the actions taken related to health, but also to the determinants, correlates, and consequences of those actions [7]. Culturally determined factors, such as cultural diversity in the treatment of illness or in beliefs about the causes of sickness and health, are integral to health behaviors. These factors should also be considered when measuring health behaviors such as adhering to a medication regimen or seeking medical care [8].

Health behaviors can be assessed with either qualitative or quantitative methods. Qualitative methods of measurement strive to document and interpret what is being measured as fully as possible, in order to gather more broad, dynamic information from the participants [9]. For instance, researchers in Sudan used an ethnographic study with in-depth interviews to assess cultural perceptions and health behaviors related to safe motherhood for Sudanese women [10]. Qualitative measures of a health behavior typically involve open-ended questions or immersed observation of a phenomenon, [11] allowing for a more complete description of the measured behavior. In contrast, quantitative methods more narrowly measure the amount of a particular characteristic, emphasizing the measurement of objective, observable, and quantifiable data [9,12]. This method is useful to measure facts and causes of behavior that are assumed to be stable phenomena [13]. Examples of both are seen in (Table 1). Qualitative and quantitative methods have fundamentally different approaches, but they can be used in a complementary manner, also known as “mixed method approach” to give a more comprehensive perspective for a given outcome [4]. As an example, we used a mixed methods approach to evaluate
factors associated with medication adherence and return to clinic for HIV-infected children after a period of conflict and violence in Kenya [14]. Qualitative key informant interviews allowed us to describe prominent barriers in the return to clinical care in the period immediately after the humanitarian crisis, while structured items evaluating missed medication doses and records of clinic attendance provided adherence and attendance rates.

In addition to considering the difference between qualitative and quantitative methods of measuring health behaviors, the language requirements of the measurement method merit consideration. Most methods of measuring a health behavior require the participants to use oral or written language, which becomes a key consideration in cross-cultural measurement. The basic cognitive processes involved in responding to a question—the comprehension of the question, remembering the information needed to answer the question, evaluating that information as shaped by one’s attitudes and beliefs, and responding with an answer—are all influenced by both language and cultural context [15]. For example, when asking Kenyan families about children’s HIV therapy adherence, cognitive interviewing revealed that they had difficulty comprehending key language concepts such as “missed doses” despite seemingly appropriate translation [16]. Among the measurement methods, verbal reports are generally simpler and less costly than formalized questionnaires or observations [3]. Verbal reports could include face-to-face interviews or the verbal reading of formal questionnaires or rating scales. The structure of interview or questionnaire items can vary in scope and formulation; some measurement items may be based on stories or vignettes, others use open-ended questions to elicit qualitative data, and still others employ closed-ended, quantitative questions. Many methods of assessing health behaviors require print literacy, involving reading and writing from the participants [17]. Examples include the use of written questionnaires, rating scales, surveys, or checklists. Sometimes, the literacy demands of these methodologies will be altered by the use of pictures instead of words, which require picture recognition as the form of symbolic representation [18]. A study in Burkina Faso, for example, evaluated self-reports of health through questionnaire items, a wooden representation of a visual analogue scale, and a visual analogue scale drawn on paper and found differences in cultural sensitivity [19].

**Issues in cross-cultural measurement of health behaviors**

The need for cross-cultural measures of health behavior has become increasingly apparent. Health problems have global impact and there are dramatic health disparities both around the world and among subcultures of a given population [20,9]. With this increasing awareness has come more cross-cultural research, as well as a deeper understanding of how cultural differences in beliefs, traditions, norms, and values can shape health problems and behaviors [9]. In the earlier conceptual framework of empiricism, “culture-free” assessments made in a “context-free clinical setting” were considered the most reliable and valid for measuring health behavior [21]. Realizing that culture permeates all aspects of behavior has led to more “culture-common” assessments that try to measure behavior considered common among different cultures [18]. More recent work in cross-cultural measurement has recognized that a single assessment usually cannot be used universally in all cultures [22]. This requires attention to reducing cultural differences in how an assessment performs and adapting assessments for each culture, as the influence of culture will and should be detectable [21]. Some propose treating cultures as a variable under examination, similar to age or sex [23]. Cross-cultural administration of measurement tools often reveals particular cultural variables that need to be considered compared to the design and administration of the original assessment [24,25]. For example, in Stanczak’s 2001 study looking at the development and validation of an Arabic version of the Expanded Trail Making Test, healthy Sudanese subjects had similar scores to American brain-damaged subjects when the adaptation of this neuropsychological test was used in “normal” US and Sudanese citizens, which is an unacceptable false positive diagnostic rate. Although the study was unable to provide a rationale for the significant differences, it did speculate on several hypotheses including differences in cognitive style which are influenced by sociocultural factors, the novelty of psychological testing, and possible developmental factors such as nutrition and educational opportunities [26].

Translation from one language into another is an obvious issue when one wants to measure something in multiple cultures; however, exchanging the words of one language for the words of another is not enough if the conceptual meaning is not equivalent [27]. Focusing solely on translating word-for-word creates poor sentences. Even worse, translations sometimes lack comprehensible meaning—particularly if there are not equivalent words, parts of speech, or concepts within the target language [9]. Even when the same language is used from one country to another, there may be differences in the concepts elicited by a particular phrase [28]. A major issue to consider for measuring health behaviors across cultures is whether there is cultural equivalence for the concept being

---

**Table 1. Methodologies Used to Measure Health Behaviors [9].**

<table>
<thead>
<tr>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating scales</td>
<td>Interviews or focus group discussions with the subjects of interest</td>
</tr>
<tr>
<td>Checklists</td>
<td>Content analysis of documents to identify common themes</td>
</tr>
<tr>
<td>Physiological measures (e.g., plasma drug concentrations to measure medication adherence)</td>
<td>Observation of participants</td>
</tr>
<tr>
<td>Observations or interviews in which structured criteria are evaluated, recorded, and quantified.</td>
<td>Open-ended questionnaires</td>
</tr>
</tbody>
</table>
measured [29]. If conceptual equivalence is not present, a translated question may elicit a very different response from what was originally intended [18]. For example, if the meaning of depression varies from one culture to another, the assessment of health or behaviors related to depression in a given culture must take the meaning into account for both the measurement and for the interpretation of responses. For even better cultural equivalence, one would also have equivalence in how each cultural group operationalizes or defines the construct of interest. For example, good “quality of life” in one culture might include having material possessions such as a house and a car. In another culture, living near extended family or just having basic necessities available, such as food and water, might play an important role. Even with measurement instruments claiming to include items with validity across cultures, careful evaluation can uncover a lack of equivalence for concepts heavily influenced by the culture [30].

Biases affected by or related to one’s culture may undermine the validity of an assessment of health behavior. Cultural biases could be present at the level of the construct, the method, or the individual measurement items being used [29]. To use a previous example, if one failed to understand that the close presence of family was important within a culture, measures of quality of life that never assessed family presence would not appropriately sample this construct. Then, this measure would have a construct bias undermining validity. Biases can also occur based on the methods used to measure in a behavior a particular culture. For example, the subcultures of Hispanic Americans and White Americans have been found to respond differently to measurement methods such as five-point rating scales, with Hispanics more often choosing the extreme responses [31]. Method bias can result if members of different cultural groups respond differently to a specific type of measurement instrument or method, such as different responses because of cultural differences in gender, age, educational level, or concern with maintaining social relationships. Cultural biases can also emerge at the level of an individual item or question used to measure a health behavior. Problems with item content, poor translation, or poorly formulated items can all bias how a particular culture responds [29,32].

These biases and conceptual equivalence issues can create both subtle and significant problems in measuring a health behavior in a given culture. Even within the same culture, subculture differences in ethnicity, dialects or languages, gender, age, and education can impact the experience and results of health behavior measurements [33]. These challenges can be minimized by adapting health behavior assessments appropriately for use across cultures [34].

Adaptation of health behavior assessments across cultures
Investigators seeking to measure health behavior across cultures often try to modify or adapt a particular measurement methodology to meet the challenges of cross-cultural assessment. If the reliability and validity of a strategy can be established in multiple cultures, this strengthens the evidence in favor of using it. However, the process of adaptation must overcome the challenges outlined previously. The majority of research on adapting health behavior assessments across cultures focuses on the adaptation of written questions, whether these questions to be used as part of a questionnaire, rating scale, or structured interview.

A careful approach to translation is crucial to cross-cultural adaptation of any strategy for assessing health behaviors. When developing a questionnaire or writing interview questions, simple steps can be followed to facilitate later translation (Table 2). Clear and objective wording is crucial when questions are translated. Chang’s 2001 study on translation and the equivalence of questionnaires looked at the Menstrual Distress Questionnaire (MDS). Bilingual students were asked to complete Chinese and English versions of MDS and the kappa scores showed overall high equivalence. However, the few low kappa items contained subjective or colloquial Western terms that had been difficult to translate because the literal meanings in Chinese seemed strange [35]. For example, statements such as “feeling sad or blue” have caused difficulty in cross-cultural adaptation [36]. For any written materials, the reading level should be determined, with a goal of keeping it at the grade level of 6 or 7 by criteria such as the Flesch-Kincaid readability index, [17] as reading level adjustment may lessen other differences in cross-cultural performance [37]. Equivalent adaptations and simplifications should be considered for materials in other languages.

<table>
<thead>
<tr>
<th>Table 2. Tested Recommendations for Developing Health Behavior Measures in English.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using short simple sentences</td>
</tr>
<tr>
<td>Using the active voice</td>
</tr>
<tr>
<td>Repeating nouns rather than using pronouns</td>
</tr>
<tr>
<td>Avoiding metaphors and idioms</td>
</tr>
<tr>
<td>Avoiding the conditional mode of verbs</td>
</tr>
<tr>
<td>Avoiding possessive forms</td>
</tr>
<tr>
<td>Using specific terms</td>
</tr>
<tr>
<td>Avoiding words that can generate vague meaning</td>
</tr>
</tbody>
</table>


The process of translation requires careful decisions among a number of processes. Appropriate translation employs methods far beyond simple language conversion; systematic evaluation of questionnaire translations reveal further evaluation is needed to ensure the language conversion maintains
psychometric quality [38]. Translation should include frequency counts of words, aiming to translate words and expressions into words and expressions that occur with approximately the same frequencies in the two languages [39]. One must also decide whether to use unicentered (asymmetrical) translation or decentered (symmetrical) translation (Table 3). Whenever possible, decentered (symmetrical) translation is preferred as it prioritizes meaning equivalence. Decisions must also be made about whether to use a forward-translation design or a back-translation design. The advantages and disadvantages of these design strategies are noted in (Figure 1). The disadvantages mostly involve biases from translation. To minimize these biases, whenever possible, translators should be ethnically and culturally representative of the target population, but fluent in both the source and target language and familiar with both cultures [40]. The source language translators should also know and understand the construct being measured, while the target language translators would not have this familiarity or detailed knowledge [41].

Once translation has been carried out, the next important cross-cultural adaption step is pre-testing the translated instrument. Techniques for pre-testing health measurement instruments include cognitive interviewing, committee reviews, field-testing with bilinguals, and field-testing with monolinguals. Cognitive interviewing is a qualitative research technique to study how targeted audiences understand, mentally process, and respond to the materials presented, particularly for survey questionnaires [42]. Representatives from the target audience are asked to process the measurement items item-by-item, with a trained facilitator [43]. The facilitator uses verbal probing techniques and guided “thinking aloud” to evaluate how well the respondent understands the measurement items and what the questions

| Table 3. Unicentered (Asymmetrical) vs Decentered (Symmetrical) Translation [9]. |
|---------------------------------|---------------------------------|
| **Unicentered (Asymmetrical) Translation** | **Decentered (Asymmetrical) Translation** |
| Target language translation follows or remains loyal to source language. | Both source and target languages considered equally important. |
| Often used when adapting previously developed measure. | Most easily implemented when measurement tools for both cultures are developed at same time. |
| May hinder conceptual equivalence. | Improving meaning and meaning equivalence are given priority for all versions of measurement tool. |

**Figure 1.** Design strategies for translation [39].
They can also complete the source language measurement, which should be piloted for recognition prior to use. Field-testing with bilingual individuals uses a questionnaire or rating scale that exists in two languages to evaluate both language speakers and monolingual speakers. They can also complete the target language measurement. Items that complete both the original and back-translated versions of a measurement instrument to compare how each performs. They can also complete the source language measurement, while simultaneously having target-language monolinguals complete the target language measurement. Items that perform differently in the two groups (and thus do not have item equivalence) can be identified and carefully studied.

Field-testing the health measurement items with either bilingual or monolingual individuals provides another crucial step for evaluating reliability, validity, and equivalence. Field-testing with bilingual individuals uses a questionnaire or rating scale that exists in two languages to evaluate both language versions of the measurement tool. They can rate and compare each measurement item on the equivalence of the versions or the translation quality. Monolingual individuals can complete both the original and back-translated versions of a measurement instrument to compare how each performs. They can also complete the source language measurement, while simultaneously having target-language monolinguals complete the target language measurement. Items that perform differently in the two groups (and thus do not have item equivalence) can be identified and carefully studied.

Recommendations for the adapting health behavior measurement tools for cross-cultural use

In conclusion, assessing health behaviors in a cross-cultural setting requires careful attention to the choices of health behavior measurement methods, awareness of the challenges of cross-cultural health measurement, employment of the processes that can address these challenges, and an understanding of the culture of interest to guide the choices made at every stage. Following these methods to adapt measures to evaluate health behaviors in another culture are highlighted in this manuscript to maximize reliability and validity across cultures.

Key messages

- When possible, assessment tools should be developed simultaneously in both source and target languages, following recommendations to use simple sentences.
- The best possible translation processes should be employed, including back-translation, decentering, committee involvement, and field-testing with bilingual speakers and monolingual speakers.
- Pilot-testing should investigate the cultural equivalence, reliability, and validity of any assessment tools and include representatives of the target population.
- Limiting literacy demands should be a priority. Pictures can be used in the place of written words, but pictures should be piloted for recognition prior to use.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

<table>
<thead>
<tr>
<th>Authors’ contributions</th>
<th>RCV</th>
<th>MSM</th>
<th>WMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research concept and design</td>
<td>✓</td>
<td>--</td>
<td>✓</td>
</tr>
<tr>
<td>Collection and/or assembly of data</td>
<td>✓</td>
<td>✓</td>
<td>--</td>
</tr>
<tr>
<td>Data analysis and interpretation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Writing the article</td>
<td>✓</td>
<td>✓</td>
<td>--</td>
</tr>
<tr>
<td>Critical revision of the article</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Final approval of article</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Statistical analysis</td>
<td>✓</td>
<td>--</td>
<td>✓</td>
</tr>
</tbody>
</table>

Acknowledgement

This work was supported by a grant to Dr. Vreeman from the National Institute of Mental Health (5K23MH087225-03), as well as by the President’s Emergency Plan for AIDS Relief (PEPFAR) through USAID under the terms of Cooperative Agreement No. AID-623-A-12-0001. It is made possible through joint support of the United States Agency for International Development (USAID). The contents of this paper are the sole responsibility of AMPATH and do not necessarily reflect the views of USAID or the United States Government.

Publication history

Editors: Carl Fulwiler, University of Massachusetts Medical School, USA.
Sara A. Jahnke, Institute for Biobehavioral Health Research, New York.
Received: 10-Oct-2013 Revised: 14-Nov-2013 Accepted: 03-Dec-2013 Published: 16-Dec-2013

References

11. Swanson W.C.J. From practice to grounded theory: Qualitative research
49. Matsers GA, Sanderman R, Kim AY and Coyne JC. Problems in cross-