

Developing a low-cost technique for parallel cross-cultural instrument development: the Question Appraisal System (QAS-04)

Dean, Elizabeth; Caspar, Rachel; McAvinchey, Georgina; Reed, Leticia; Quiroz, Rosanna

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Dean, E., Caspar, R., McAvinchey, G., Reed, L., & Quiroz, R. (2005). Developing a low-cost technique for parallel cross-cultural instrument development: the Question Appraisal System (QAS-04). In J. H. P. Hoffmeyer-Zlotnik, & J. Harkness (Eds.), *Methodological aspects in cross-national research* (pp. 31-46). Mannheim: GESIS-ZUMA. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49145-7>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

DEVELOPING A LOW-COST TECHNIQUE FOR PARALLEL CROSS-CULTURAL INSTRUMENT DEVELOPMENT:

The Question Appraisal System (QAS-04)

*ELIZABETH DEAN, RACHEL CASPAR, GEORGINA MCAVINCHHEY,
LETICIA REED & ROSANNA QUIROZ*

Many approaches are used to prepare instruments for multicultural administration, depending on the scope, schedule and budget of the study. Sequential questionnaire development, the most common approach to developing cross-cultural instruments, is also the most affordable. Designers formulate and pretest an instrument in the source language, then translate it into the target language(s) using culture-specific tailoring. In contrast, parallel development incorporates target cultures throughout the design and pretesting process. The disadvantages to parallel development are that it is expensive, time-consuming, and subject to version control problems. The Question Appraisal System (QAS) is a coding tool for pretesting instruments (Willis & Lessler, 1999). The QAS is supported by an item taxonomy of the cognitive demands of a question and documents the features that may lead to response error. Results of the appraisal are used to revise question wording, response wording, questionnaire format, and question ordering. This paper describes research conducted to update the QAS to identify problems due to cross-cultural and cross-linguistic application of questions.

1 Introduction

Globalization tendencies and demographic shifts demand understanding between cultures as well as the technical ability to communicate effectively in multiple languages (Cronin, 2003). The need for understanding other cultures and languages is such that institutions of higher education now require diversity training in most fields of study. Within the United States, demographic changes in minority populations have brought about cultural diversity that is often accompanied by linguistic gaps. Latin America and Asia are the points of

origin for over 75 percent of the foreign born U.S. population, and meeting linguistic demands is a challenge for health and social service agencies, educators, policy planners and researchers. As the needs for information from diverse populations rise, survey research is charged with finding ways to bridge linguistic and cultural gaps and ensuring accurate representation in research studies (Flores et al., 2002; Li et al., 2001).

A compelling need for sound methodological practices in instrument translation and adaptation has been present in survey research for over half a century. Early studies on cross-cultural congruence revealed that wording and translation were considered ‘the weakest link’ in the process of attaining comparable tools for research purposes (Kumata & Schramm, 1956). During the late 1960s and early 1970s, multinational survey projects appeared, and most cross-cultural researchers dealing with human behaviour favoured investigations across the globe. Benefits derived from cross-cultural research, such as increasing the range of analytic variables and increasing sensitivity to context, would have been impossible to attain in mono-cultural research (Brislin, 1993). The needs of cross-cultural survey research, coupled with the desire to reap its benefits, have pushed the field towards establishing sound practices in instrument adaptation and translation.

In the U.S., standard practices for conducting Spanish interviews have evolved along with our understanding of cross-cultural methodology. Attending to cultural and linguistic differences among populations when developing a new instrument is costly and time consuming, therefore, using existing questionnaires of mono-cultural context is the most frequently used approach in questionnaire development. Although relatively inexpensive, the use of an existing instrument often fails to address semantic, conceptual and normative equivalence. Neglecting to address the comparability of questions can decrease the validity and reliability of measures. Finding sound but inexpensive methods of addressing cultural and linguistic issues during instrument design is a goal that survey research has not yet met. The many approaches developed so far require trade offs between cost (both in terms of money and time) and instrument quality.

A possible solution to incorporating language awareness into the developing stage of an instrument at a relatively small cost is to address potential linguistic and cultural issues in the pre-testing stage. The Question Appraisal System, or QAS-99, (Willis & Lessler, 1999), a coding system for identifying question characteristics likely to result in response errors, is capable of accommodating steps that assess potential language and cross-cultural problems. This research examines a variety of problems that cross-cultural survey designers have encountered throughout the years and proposes an enhanced version of the QAS-99, the QAS-04, as a practical and relatively inexpensive way to improve methodological practices in the cross-cultural research field.

2 Theoretical Framework

To fully understand the implications of developing survey instruments for translation and application across multiple cultures, it is important to place these activities in the wider context of communication between cultures. Too often, survey researchers develop measures without understanding how behaviours are understood differently across cultures. To understand the context of a survey response, researchers must know how concepts, values, and linguistic constructions vary across languages and cultures. Furthermore, researchers must understand how cultures vary simply in their demographic characteristics. This section provides a brief summary of these cross-cultural variations.

2.1 Understanding the link between culture and response

A hypothesis or research question often originates within a specific cultural context. What may be an important question in one culture may not be important or even relevant to people of a different culture. Only when the population studied understands, is able, and is willing to answer a question should other methodological issues be addressed (Fowler, 1995; Peterson, 2000). In adapting an instrument, researchers must aim to establish comparability of concepts, norms and semantics, none of which can be achieved without understanding the culture of the respondents.

Concept comparability. Brislin (1993) notes that understanding behaviour within context can provide insights on responses that otherwise might be attributed to the wrong cause. To establish a common understanding of concepts when studying different cultures, researchers have called the culture-common and culture-specific perspectives ‘emic’ and ‘etic’. The emic perspective involves the evaluation of a studied phenomenon from within the culture and its context and it aims to understand its significance and its interrelations with cultural elements. The etic perspective, in contrast, involves the evaluation of phenomena from outside the culture, aiming to identify and compare similar phenomena across different cultures (Berry, 1969; Brislin, 1993).

Failing to become acquainted with emics of populations being targeted in survey research can lead to misattribution of response. For example, items designed to gather demographic information often fail to acknowledge emic differences in populations. Asking for the age of a participant may appear to be a straightforward and simple question that anyone could answer, but this is not always the case. In some cultures date of birth is not considered a relevant piece of information, and the approximate age of a child is more often calculated by references to agricultural times of the year or other events. Educational attainment, income, employment and marital status are all demographic questions for which special awareness of within-culture perspectives are needed (Braun & Mohler, 2003).

Before addressing the particulars of best practices in translating an instrument, it is necessary to determine whether concept equivalence exists. It is possible that concepts that are emic – that is, unique to the source culture – render the line of research invalid in another culture. A study conducted in New Zealand using the SF-36, a widely used instrument that measures health related quality of life, found that emic views of health among a Maori sample did not discriminate between physical and mental health even though the items were written to differentiate between the two concepts (Scott et al., 1999).

Normative Comparability. Rules about disclosure of information between in-groups vary greatly. What can be said about particular subjects can be greatly influenced by what is perceived to be appropriate within an in-group. Religious preferences, political views and information related to personal matters are usually topics for which willingness to respond varies by culture. When interviewers and respondents share the same ethnicity, the willingness of the respondents to disclose, and hence the validity and the reliability of the data provided, increases (Marin & Marin, 1991).

Semantic Comparability. Semantic equivalence is related to the degree to which terms in the translation connote the same meaning in translation as in the source language. Achieving semantic comparability in questionnaire translation has been considered by some to be the most difficult step in the translation and adaptation of questionnaires. And, perhaps the greatest challenge to functionally equivalent questionnaires is the lack of equivalent markers or terms for words or concepts that do not exist in the target language or culture.

Research describing troublesome areas in test or questionnaire translations often focuses on this type of comparability issue (Arias et al., 1999; Lange, 2002). In the U.S., for example, the foster system trains and pays parents to care for children on a temporary basis. Collecting information about foster children is quite challenging for cultural, linguistic, and semantic reasons, however, as the term ‘foster child’ has no comparable translation in many other countries.

Trying to keep language simple has been long been advocated in writing survey questions (Peterson, 2000) but due to the absence of equivalent terms simplicity does not always equate to brevity. Terms that need to be translated often have either a more specific or a broader meaning in the target language. As a result, additional information needs to be presented. For example, the word ‘youth’ in English is often translated as ‘niño (child)’ in Spanish. However, the word ‘niño’ denotes someone twelve years old or younger, and is therefore not the optimal choice. To arrive at the closest meaning to the word it is necessary to use the equivalent of ‘young person’ and specify a range of ages included within the term. Doing this requires either adding more items to the instrument (Smith, 2003) or more words to an item.

In cross-cultural research, translating a word to its precise meaning is further complicated by regional variations in the particular language. Marin and Marin (1991) suggest that to avoid offending any given subgroup within a culture the standard or most neutral version of the language should be used. If regional variations are needed to accommodate a subgroup, both the standard term and the specific variation should be presented.

2.2 The impact of demographic variation across cultures

Educational level is often overlooked when designing cross-cultural questionnaires. U.S. literacy data reveal that a large number of adults can only read at elementary levels (Weech-Maldonado et al., 1999), with recent immigrants from Guatemala and El Salvador reporting less than a secondary education.

Flores et al. (2002) report that translations are written ‘at an inappropriately high reading level for the target populations’ and suggest that research instruments be validated by Spanish speaking families with ‘poor and low-literacy populations’. In addition, when addressing different generations within one-culture, accommodations should to be made to meet their particular needs. During the 2000 Census, Chinese translations used traditional Chinese. Pan (2003) reported that during that census younger generations of immigrants were not able to read the traditional Chinese characters, reducing the pool of participants in this event.

Although the relationship between literacy levels and the ability to follow skip patterns in cross-cultural research has not been studied, familiarity with questionnaires is likely to influence the ability to follow instructions and respond to items. Forms literacy is an important element to consider when constructing a questionnaire. Navigating through an instrument presents a challenge to many foreign born respondents. Lack of exposure to ordinal scales and to multiple choice answers may require training for the participant (Lange, 2002).

2.3 Approaches to translation and cultural adaptation of instruments

The last twenty years have seen great changes in the methodology of instrument adaptation. There are many approaches to translation and cultural adaptation of research instruments. Popular criteria for evaluating translation techniques include the level of comparability achieved when translating an instrument and the cost involved in the task. Besides these criteria, research has focused on the number of individuals involved in completing a particular translation task or on the steps that should be taken to ensure the translation work is done accurately. The most frequently used techniques are:

- *Simple direct translation*, the translation of a document by a single individual from its original version to another language, is the easiest and cheapest method used. The most obvious drawback of this approach is that it does not offer ways to verify that the translation is true to the original and that different aspects of the language used are the most appropriate.
- *Translation with back translation* is a technique that requires at least two individuals. In this method, a bilingual individual translates the instrument from the original or source language to the target language. A second individual, without knowledge of the contents of the original document, translates the document back to the source language. The original and back translations are compared and differences between the documents are addressed. The translation/back translation approach allows the researcher to better assess that the translation is true to the original. The cost of this method rises considerably due to multiple steps and time needed for the process to be finalized.
- *Translation by committee* requires a group of individuals to arrive at a consensus about the best translation of an instrument after a series of steps.

Harkness (2001) provides an original framework for understanding instrument translation procedures. Within this conceptual framework, existing instruments are either adopted or adapted. Adopting involves directly translating the existing instrument into the target language with little attention to culture. Adapting requires reviewing an existing questionnaire for cross-cultural appropriateness in the target language, adjusting the source language as necessary, then translating the questionnaire into the target language.

Translators can employ sequential, parallel or simultaneous development of a new survey instrument. Sequential development refers to instrument development procedures in which the instrument is designed and pretested in the source language only. Once the instrument is finalized or 'locked', it is translated. Thus, the sequential development process does not take into account cultural or linguistic issues until after the finalized source language instrument is translated. Sequential development is comparable to adopting an existing instrument. The approach is efficient, but minimizes attention to cross-cultural issues. The parallel development process incorporates input from all target cultures during instrument development. An instrument is designed and pretested in the source language with a multicultural team. As with sequential development, the instrument is locked prior to translation. The most culturally adaptive method of developing an instrument is the simultaneous approach. Survey designers using simultaneous instrument development seek to create more than one version of an instrument using decentring, a process in which cultural appropriateness is given equal importance in all languages. The

instrument is designed and pretested simultaneously in multiple languages (Harkness, 2001). This process offers tremendous benefits for developing valid questions that are appropriate within multiple cultures. However, it tends to be expensive and time consuming, and it may produce ambiguous questions. The need arises for a lower cost methodology to simultaneously develop new questionnaires and to adapt, rather than adopt, existing instruments.

3 Question Appraisal Systems

As the use of cognitive pretesting and evaluation methods became a matter of course for questionnaire development in the 1980s and 1990s, the need arose for low-cost methods to assess the cognitive characteristics of questions that might lead to response error. Methods such as cognitive interviewing, behaviour coding, interview observation, and embedded question wording experiments were invaluable, but were at times found to be expensive and time-consuming. Question appraisal methods were designed to meet the demand for a more cost-effective way to systematically assess cognitive problems with instruments.

A questionnaire coding system developed by Lessler and Forsyth (1996) was designed based on Tourangeau's model of the question response process. The four cognitive processes in Tourangeau's model are comprehension, memory retrieval, judgment, and response selection (1984). Lessler and Forsyth found that, compared to other pretesting methods, their coding system was less expensive to implement and identified similar problems. Forsyth and Hubbard (1992) had previously had similar findings when they validated a questionnaire appraisal system by using cognitive think-aloud interviews. In 1999, Willis and Lessler developed the QAS-99, a tool based on previous questionnaire appraisal systems. The QAS-99 was an effort to 'assist questionnaire designers in evaluating survey questions, and in finding and fixing problems, before questions "go into the field",' (1999). The QAS-99 differed from Lessler and Forsyth's system in that it had significantly fewer codes as a means to decrease the difficulty of the coding activity and improve inter-coder reliability. The QAS-99 was designed not to replace other interactive cognitive methods, but to provide questionnaire reviewers and developers a tool for systematically reviewing and improving survey instruments.

The QAS-99 is comprised of seven steps by which a user evaluates each item in a questionnaire. Each step maps to a specific stage in Tourangeau's question-response model (1984). The seven steps and their associated question-response stages are displayed in Table 1.

Table 1 Steps in the Question Appraisal System (QAS-99)

QAS Coding Step	Stage in Tourangeau's Question-Response Process
<i>Reading:</i> Determine if it is difficult for the interviewers to read the question uniformly to all respondents.	Occurs prior to question-response process
<i>Instructions:</i> Look for problems with any introductions, instructions, or explanations from the respondent's point of view.	Comprehension
<i>Clarity:</i> Identify problems related to communicating the intent or meaning of the question to the respondent.	Comprehension
<i>Assumptions:</i> Determine if there are problems with assumptions made or the underlying logic.	Comprehension
<i>Knowledge/Memory:</i> Check whether respondents are likely to not know or have trouble remembering information.	Memory/Retrieval
<i>Sensitivity/Bias:</i> Assess questions for sensitive nature or wording, and for bias.	Judgment
<i>Response Categories:</i> Assess the adequacy of the range of responses to be recorded.	Response Selection

Source: Tourangeau, 1984

An additional eighth step provides an 'other' category for problems that cannot be assigned to one of the previous seven steps (Willis & Lessler, 1999).

4 Expanding the QAS-99

As the volume of multilingual surveys has increased over the past several years, survey designers have recognised the advantages of working more closely with translators throughout the questionnaire development process. A best practice used by the authors of this research involves collaborating with translation experts as early as possible in the instrument development process. Recognizing that establishing cross-cultural and multilingual validity is vital to reducing survey response error, a natural next step was to expand the QAS-99. The system was revised to include codes to allow for a systematic evaluation of a questionnaire for problems that might emerge in translation and cross-cultural application.

Prior to expanding the QAS-99, common best practices synthesised recommendations from previous research (Brislin, 1993; Maxwell, 1996) to develop a series of practical guidelines to establish item and concept equivalence across languages. These guidelines are provided in Table 2 and Table 3. These practical guidelines form the basis of the new steps added to the QAS-99.

Table 2 Practical Guidelines for Developing Cross-Cultural Surveys

Use reference periods that are culturally relevant
Avoid making assumptions about knowledge that may only be applicable in the source culture (e.g. religion, sports, holidays, other customs)
Be aware that name formats vary
Avoid using pictorial information that may not be fitting to the population studied
Use seasonal and holiday references that are part of the targeted culture

Table 3 Practical Guidelines for Developing Questionnaires for Translation

Avoid using double negatives
Using short sentences of less than sixteen words
Employ active rather than passive voice
Repeat nouns instead of replacing them with pronouns
Avoid metaphor and colloquialism
Avoid adverbs and prepositions telling where or when
Avoid possessive forms if possible
Avoid sentences using the same verbs if the verbs are used to mean different actions
Establish measurement and weight equivalences for cultures using metric systems
Use nouns common to the culture
Be specific when using the word “you”
Remember gender specific references might create a need for longer sentence structures

Two new QAS steps incorporate criteria for evaluating questions for multilingual and cross-cultural administration:

- *Cross-cultural*: Assess questions for problems in the response process that may emerge when the instrument is applied to varied cultures.
- *Translation*: Identify areas that can be clarified for a more accurate translated instrument.

In addition, a step for cross-question problems was included (displayed in the Appendix). Step 8, Cross-Cultural Considerations, consists of seven codes. These are displayed in Table 4. Most of these seven codes address concept equivalence. Reference Periods (8a), Knowledge (8b), Measuring Units (8c), Assumptions (8d) and Response categories (8e) highlight concepts that tend to vary across cultures. For example, Assumptions (8d), can be used to identify emic population differences that may affect questions about religious

practices (such as in predominantly Christian or Muslim countries) or sports (such as European football versus U.S. football). Additionally, Name Format (8f) and Politeness (8g) address normative equivalence, such as the appropriate way to identify individuals and the perceptions of polite conversation within a culture.

Table 4 Cross-Cultural Codes

ID	Proposed Code	Definition
8a	Reference Periods	The use of seasons as a reference period might be ambiguous or uncommon. Consider converting into months. Date formats vary. Consider using words for the month to avoid misunderstandings (e.g., 15 April 2004 vs April 15, 2004).
8b	Knowledge	Knowledge may not exist. Respondent is unlikely to know the answer to a factual question because he/she not familiar with the source culture. A culture-specific example is health insurance in the US for respondents who originate from countries with nationalized health insurance.
8c	Measuring Units	Consider reporting measuring units in both the English system and the metric system.
8d	Assumptions	Consider revising culturally inappropriate assumptions, including statements related to: sports, drugs, foods, drinks, activities, meal time, music, family ties, holidays, religion, books, magazines, school system, health system, and history.
8e	Response Categories	There is no equivalent concept or rating scale in foreign language. Avoid rating scales with more than 5 categories.
8f	Name Format	Response categories lack a space for other types of names. For example, Spanish speakers use paternal last name as well as maternal last name. Consider other naming conventions.
8g	Politeness	Courtesy and politeness can differ between cultures. Consider adding a 'Please' before commands like, 'Do not include ...', 'Mark every ...', 'List all ...' Consider using 'could' instead of 'should' if possible. Some commands or instructions might be perceived as rude, and respondents could change their attitude towards participating.

In contrast, Step 9, Potential Translation Problems, is focused more on semantic equivalence. Table 5 displays the codes featured in Step 9. Step 9 features codes that help ensure that the words used in translation have the same or similar meaning as in the source language. Therefore, Idioms (9b) are identified as problematic because they tend to lack equivalent meaning in translation. Likewise, vague quantifiers such as Time Adverbs (9e) should be specified with time periods as much as possible to clarify meaning.

Table 5 Translation Codes

ID	Proposed Code	Definition
9a	Double Negatives	This type of construction is hard to translate and can easily cause misunderstandings in other languages.
9b	Idioms	Many idioms do not have an equivalent across languages.
9c	Acronyms	The acronyms have no meaning in other languages. Consider providing an explanation with the acronym.
9f	Term ‘You’ is not Defined	Need to define the word ‘You’ (i.e., plural, singular, feminine, masculine, formal, informal).
9h	Time Adverbs	Need to avoid adverbs in the use of time: recently, lately, usually. Time references might be understood somewhat differently between languages. Consider specifying time frame.
9j	No Equivalent Term or Concept	Consider including an additional explanation.
9m	Adjectives Modifying Other Adjectives	Using adjectives to modify other adjectives, (e.g. ‘house warming party’, which must be literally translated from English to Spanish as ‘A party in celebration of the purchase of a home in which guests take presents for the new home owner’) is an uncommon grammatical usage in languages other than English. Consider paraphrasing and clearly define each term.
9n	References Applicable only to English	Toll free numbers, Web sites, contact information, books and other references are only available in the source language. Consider verifying which services or references are available in the target language. Also consider using numbers instead of letters on phone numbers.

These two new QAS categories provide survey designers with a tool for facilitating multi-cultural instrument development within the source language prior to translation. The QAS-04 uses conceptual, normative, and semantic understanding of the target language to broaden an instrument’s applicability. Moreover, the relatively low costs associated with applying the QAS meet the need for an affordable way to conduct parallel development of a survey for translation.

5 Directions for Future Research

The QAS-04 provides a valuable tool for incorporating parallel question development into the source language questionnaire development and pretesting process. At minimum, it provides monolingual questionnaire developers with a tool by which they can anticipate question problems that may only show up in cross-cultural administration or translation. It provides a mechanism for bringing translation staff into the questionnaire design process at an earlier stage. The QAS-04 should be evaluated in a test application with an existing

questionnaire. It may be useful to compare the results of a test application to actual questionnaire data and interviewer experiences. Such a comparison would provide information on whether the QAS identifies the same problems interviewers and respondents experience in the field. The QAS-04 should also be assessed for internal consistency and reliability of the coding scheme, in order to assess the feasibility of using the codes.

In addition, the QAS-04 should be validated with reliability testing between monolingual and bilingual coders. At this time it is unclear whether monolingual coders and bilingual coders are equally capable of implementing the QAS-04. Moreover, the QAS-04 should be comparatively applied to questionnaires in the target languages as well as in English. Comparing the results could be a mechanism for validating translations as well as assessing the effectiveness of the tool itself.

Ultimately, the tool will be validated by its usefulness in developing cross-cultural and multilingual questionnaires. Only repeated use can determine whether it makes the survey development process easier or more efficient. It is our hope that the QAS-04 will provide the survey research community with a technique for systematically incorporating parallel questionnaire development into surveys that must be fielded in multiple languages and/or multiple cultures.

References

- Arias, I., E. Rodríguez, J. Padilla, N. González, and M. A. Rodríguez. 1999. "Translation to Spanish, Reproducibility, and Cross Cultural Adaptation of the Miller-Rahe Recent Life-Change Questionnaire." *Arthritis Care and Research* 12:153-170.
- Berry, J. W. 1969. "On Cross-Cultural Comparability." *International Journal of Psychology* 4:119-128.
- Braun, M., and P. P. Mohler. 2003. "Background Variables." Pp. 101-116 in *Cross-Cultural Survey Methods*, edited by Harkness, J. A. et al. Hoboken: John Wiley and Sons, Inc.
- Brislin, R. 1993. *Understanding Culture's Influence on Behavior*. 2nd ed. Fort Worth: Harcourt Brace College Publishers.
- Cronin, M. 2003. *Translation and Globalization*. London, New York: Routledge.
- Douglas, S. P., and C. S. Craig. 1983. *International Marketing Research*. Englewood Cliffs: Prentice Hall.
- Flores, G., E. Fuentes-Afflick, O. Barbot, O. Carter-Pokras, L. Claudio, L. Marielena, J. A. McLaurin, L. Pachter, F. R. Gomez, F. Mendoza, R. B. Valdez, A. M. Villarruel, R. E. Zambrana, R. Greenberg, and M. Weitzman. 2002. "The Health of Latino Children: Urgent priorities, unanswered questions, and a Research agenda." *Journal of the American Medical Association* 288:82-90.
- Fowler, F. J. 1995. *Improving Survey Questions: Design and Evaluation*. Thousand Oaks: Sage.

- Forsyth, B., and M. Hubbard. 1992. "A Method for Identifying Cognitive Properties of Survey Items." Pp. 470-475 in *Proceedings of the Section on Survey Research Methods, American Statistical Association*. Alexandria: American Statistical Association.
- Harkness, J. A., F. J. R. van de Vijver, and T. P. Johnson. 2003. "Questionnaire Design in Comparative Research." Pp. 19-34 in *Cross-Cultural Survey Methods*, edited by Harkness et al. Hoboken: John Wiley and Sons, Inc.
- Kumata, H., and W. Schramm. 1956. "A Pilot Study of Cross-Cultural Meaning." *Public Opinion Quarterly* 20:229-238.
- Lange, J. W. 2002. "Methodological Concerns for Non-Hispanic Investigators Conducting Research with Hispanic Americans." *Research in Nursing and Health* 25:411-419.
- Lessler, J. T., and B. H. Forsyth. 1996. "A Coding System for Appraising Questionnaires." Pp. 259-292 in *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by Schwarz, N., and S. Sudman. San Francisco: Jossey-Bass Publishers.
- Li, R. M., P. McCardle, R. L. Clark, K. Kinsella, and D. Berch. 2001. *Diverse Voices: The Inclusion of Language Minority Populations in National Studies: Challenges and Opportunities*. Report published jointly by the National Institute of Aging, the National Institute of Child Health and Human Development, and the National Center on Minority Health Disparities. http://www.nichd.nih.gov/publications/pubs/diverse_voices.htm.
- Locke, D. C. 1998. *Increasing Multi-Cultural Understanding: A Comprehensive Model*. 2nd ed. Thousand Oaks: Sage Publications, Inc.
- Marin, G., and B. V. Marin. 1991. *Research with Hispanic Populations*. Newbury Park: Sage Publications, Inc.
- Maxwell, B. 1996. "Translation and Cultural Adaptation of the Survey Instruments." Pp. 8-10 in *Third International Mathematics and Science Study (TIMSS) Technical Report, Volume I: Design and Development*, edited by Martin, M. O., and D. L. Kelly. Chestnut Hill: Boston College.
- Pan, Y. 2003. Paper presented at the Annual Meeting of American Association for Applied linguistics. *Sociolinguistic Research and Survey Development*. Arlington, VA.
- Peterson, R. A. 2000. *Constructing Effective Questionnaires*. Thousand Oaks, CA: Sage.
- Scott, K. M., D. Sarfate, M. I. Tobias, and S. J. Haslett. 1999. *A Challenge to the Cross-Cultural Validity of the Sf-36 Health Survey: Maori, Pacific, and New Zealand European Ethnic Groups*. Seventh Conference on Health Survey Research Methods. Hyattsville, MD: U.S. Department of Health and Human Services Publication No. (PHS) 01-1013.
- Smith, T. W. 2003. "Developing Comparable Questions in Cross-National Surveys." Pp. 69-92 in *Cross-Cultural Survey Methods*, edited by Harkness, J. A. et al. Hoboken: John Wiley and Sons, Inc.

- Tourangeau, R. 1984. "Cognitive Science and Survey Methods." Pp. 73-100 in *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*, edited by Jabine, T. et al. Washington: National Academy Press.
- Weech-Maldonado, R., B. O. Weidmer, L. S. Morales, and R. D. Hays. 1999. *Cross Cultural Adaptation of Survey Instruments: The CAHPS Experience*. Seventh Conference on Health Survey Research Methods. Hyattsville, MD: U.S. Department of Health and Human Services Publication No. (PHS) 01-1013.
- Willis, G., and J. T. Lessler. 1999. *Question Appraisal System: QAS-99*. Rockville: Research Triangle Institute.

Appendix: Questionnaire Appraisal System (QAS-2004)

- ▶ **STEP 1 – READING: Determine if it is difficult for the interviewers to read the question uniformly to all respondents or if the reading level is appropriate.**
 - 1a. **WHAT TO READ:** Interviewer may have difficulty determining what parts of the question should be read.
 - 1b. **MISSING INFORMATION:** Information the interviewer needs to administer the question is not contained in the question.
 - 1c. **HOW TO READ:** Question is not fully scripted and therefore difficult to read.
- ▶ **STEP 2 – INSTRUCTIONS: Look for problems with any introductions, instructions, or explanations from the respondent's point of view.**
 - 2a. **CONFLICTING OR INACCURATE INSTRUCTIONS,** introductions, or explanations.
 - 2b. **COMPLICATED INSTRUCTIONS,** introductions, or explanations.
 - 2c. **MISSING OR INCONSISTENT INSTRUCTIONS** for DON'T KNOW and REFUSED answers.
- ▶ **STEP 3 – CLARITY: Identify problems related to communicating the intent or meaning of the question to the respondent.**
 - 3a. **WORDING:** Question is lengthy, awkward, ungrammatical, or contains complicated syntax.
 - 3b. **TECHNICAL TERM(S)** are undefined, unclear, or complex.
 - 3c. **VAGUE:** There are multiple ways to interpret the question or to decide what is to be included or excluded.
 - 3d. **REFERENCE PERIODS** are missing, not well specified, or in conflict.
 - 3e. **PASSIVE VOICE:** Question is written in passive voice. Active voice is clearer both in source language and in translation.
- ▶ **STEP 4 – ASSUMPTIONS: Determine if there are problems with assumptions made or the underlying logic.**
 - 4a. **INAPPROPRIATE ASSUMPTIONS** are made about the respondent or about his/her living situation.
 - 4b. **ASSUMES CONSTANT BEHAVIOR** or experience for situations that vary.
 - 4c. **DOUBLE-BARRELED:** Contains more than one implicit question.

Appendix (continued)

- ▶ **STEP 5 – KNOWLEDGE/MEMORY: Check whether respondents are likely to not know or have trouble remembering information.**
 - 5a. **KNOWLEDGE** may not exist: Respondent is unlikely to know the answer to a factual question.
 - 5b. **ATTITUDE** may not exist: Respondent is unlikely to have formed the attitude being asked about.
 - 5c. **RECALL** failure: Respondent may not remember the information asked for.
 - 5d. **COMPUTATION** problem: The question requires a difficult mental calculation.
- ▶ **STEP 6 – SENSITIVITY/BIAS: Assess questions for sensitive nature or wording, and for bias.**
 - 6a. **SENSITIVE CONTENT (general):** The question asks about a topic that is embarrassing, very private, or that involves illegal behavior. If question will be applied across cultures, it may be sensitive in some cultures but not others.
 - 6b. **SENSITIVE WORDING (specific):** Given that the general topic is sensitive, the wording should be improved to minimize sensitivity.
 - 6c. **SOCIALLY ACCEPTABLE** response is implied by the question. If question will be applied across cultures, social acceptability could vary.
- ▶ **STEP 7 – RESPONSE CATEGORIES: Assess the adequacy of the range of responses to be recorded.**
 - 7a. **OPEN-ENDED QUESTION** that is inappropriate or difficult.
 - 7b. **MISMATCH** between question and response categories.
 - 7c. **TECHNICAL TERM(S)** are undefined, unclear, or complex.
 - 7d. **VAGUE** response categories are subject to multiple interpretations.
 - 7e. **OVERLAPPING** response categories.
 - 7f. **MISSING** eligible responses in response categories.
 - 7g. **ILLOGICAL ORDER** of response categories.
- ▶ **STEP 8 – CROSS-CULTURAL CONSIDERATIONS: Assess questions for inappropriate or ineffective cross-cultural references.**
 - 8a. **REFERENCE PERIODS:** The reference period uses seasons, American MM/DD/YYYY format, or may be otherwise ambiguous or unusual in other cultures.
 - 8b. **KNOWLEDGE** may not exist: Respondent is unlikely to know the answer to a factual question because he/she not familiar with the American culture. Example: health insurance.
 - 8c. **MEASURING UNITS:** Measuring units are from English system. If surveying Latin Americans or Western European populations, the metric system should be used.
 - 8d. **ASSUMPTIONS:** The question includes culturally inappropriate assumptions or graphics. All statements related to sports, drugs, foods, drinks, activities, meal time, music, family ties, holidays, religion, books, magazines, school system, health system, and history should be evaluated.
 - 8e. **RESPONSE CATEGORIES:** There is no equivalent concept or rating scale in foreign language. Avoid rating scales with more than 5 categories.
 - 8f. **NAME FORMAT:** Response categories lack a space for other types of names. Spanish speakers use maternal last name as well as paternal last name, and other cultures list the family name as the first name.
 - 8g. **POLITENESS:** Courtesy and politeness can differ in other cultures. Consider adding a 'Please' before commands like, 'Do not include ...', 'Mark every ...', 'List all ...' Consider using 'could' instead of 'should' if possible. Some commands or instructions might be perceived as rude, and respondents could change their attitude towards participating.

Appendix (concluded)

- **STEP 9 – POTENTIAL TRANSLATION PROBLEMS: Identify problematic question characteristics.**
 - 9a. **DOUBLE NEGATIVES:** This type of construction is hard to translate and can easily cause misunderstandings in other languages.
 - 9b. **IDIOMS:** Many idioms do not have an equivalent in other languages.
 - 9c. **ACRONYMS:** The acronyms have no meaning in other languages. Consider providing an explanation with the acronym.
 - 9d. **UNCLEAR USE OF THE TERM ‘YOU’:** ‘You’ not defined as plural, singular, feminine, masculine, formal, informal – a necessary step for translation.
 - 9e. **TIME ADVERBS:** Question or response categories use adverbs to describe time, such as recently, lately, usually. Consider specifying time frame with number of days, weeks, etc.
 - 9f. **NO EQUIVALENT TERM OR CONCEPT** in foreign language. Text may require an additional explanation.
 - 9g. **REFERENCES APPLICABLE ONLY TO ENGLISH:** Toll free numbers, Web sites, contact information, books and other references are only available in the source language. Consider verifying which services or references are available in the target language. Also consider using numbers instead of letter on phone numbers.
 - 9h. **ADJECTIVES MODIFYING OTHER ADJECTIVES:** Using adjectives to modify other adjectives, (e.g. ‘house warming party’, which must be literally translated as ‘A party in celebration of the purchase of a home in which guests take presents for the new home owner’) is an uncommon grammatical in usage languages other than English. Consider paraphrasing and clearly define each term.
- **STEP 10 – CROSS-QUESTION: Look for cross-question problems in the entire questionnaire.**
 - 10a. **QUESTION PLACEMENT.** The questions are not positioned in the most adequate section or order.
 - 10b. **DATA COLLECTION MODE:** Sensitive question may be more effective if it was administered through another data collection mode.
 - 10c. **INCONSISTENCY WITH OTHER QUESTIONS:** Wording, or response categories lack consistency.
 - 10d. **CONTENT OF PREVIOUS QUESTION AFFECTS MEANING:** Does the content of the previous question or section affect the interpretation of the current question.
 - 10e. **SKIP PATTERN PROBLEM:** Skip pattern is illogical or inadequate.
 - 10f. **FORMATTING:** Layout or formatting is difficult to follow.
- **STEP 11 – OTHER PROBLEMS**
 - 11a. **QUESTION CONTAIN IRRELEVANT INFORMATION**
 - 11b. **INAPPROPRIATE READING LEVEL**
 - 11c. **OTHER PROBLEMS**