

Collaborative and Iterative Translation: An Alternative Approach to Back Translation

ABSTRACT

Back translation is the technique most commonly used to check the accuracy of translation in survey research. Although it results in a direct or literal translation, it does not address issues of conceptual equivalence. Furthermore, if it is not combined with pretesting, it does not address issues of comprehension and meaning to the respondent. This article discusses some of the problems of relying exclusively on back translation in isolation. It suggests the use of a collaborative, iterative approach for translating questionnaires for international marketing research.

**Susan P. Douglas and
C. Samuel Craig**

As companies expand further into international markets, penetrating emerging markets, such as China, India, and Brazil, as well as markets within the former Soviet Bloc, the Middle East, and Latin America, the conduct of international marketing research is becoming increasingly important and more complex. Not only are researchers increasingly using cross-national surveys to investigate customers' attitudes and interests, but advertising and product concepts are also being tested on a global basis. The diversity of languages spoken in these countries means that the effective translation of questionnaires and instruments used in experiments is a central concern. To obtain meaningful results, it is essential to establish equivalence of meaning in each language and to ensure that each respondent and interviewer clearly understands the questionnaire and instructions.

A procedure that is commonly used to test the accuracy of translation in multicountry research is back translation (Brislin 1970, 1980). In this procedure, a bilingual native of the target country translates a questionnaire into the target language.¹ A bilingual native speaker of the source language then translates it back into the source language. The original and back-translated versions are then compared for differences and comparability. The accuracy of the back-translated version is considered an indicator of the accuracy of the target translation. Back translation was initially developed for situations in which a researcher was not familiar with the target language but wanted some assurance that respondents were indeed being asked the same question in that language (Harkness 2003). The procedure was believed to provide insights into potential errors when no other means were available to assess the accuracy of the translation.

*Submitted April 2006
Accepted September 2006*

Journal of International Marketing
© 2007, American Marketing Association
Vol. 15, No. 1, 2007, pp. 30–43
ISSN 1069-031X (print)
1547-7215 (electronic)

Back translation is the most commonly used approach in marketing to help identify problems and egregious errors in translation. However, it does not necessarily ensure equivalence in meaning and concepts in each country, and thus the translation that is adapted to the culture of the target language might not be accurate. This is particularly likely to be an issue when idioms or colloquial language is used, as is often the case in marketing surveys. The purpose of this article is to describe the use of back translation and the problems associated with reliance on back translation in isolation. We propose an alternative, collaborative, iterative approach to translation that should result in a translated instrument that more accurately reflects the meaning of the source language in the target culture. First, we review commonly used translation practices and identify their limitations. Second, we discuss the steps needed to ensure appropriate translation, depending on the specific situation. Finally, we recommend a five-step iterative procedure that can lead to improved translations.

Historically, in the social sciences, back translation has been widely used to test the accuracy of the translation and to detect errors in translation (Brislin 1970, 1980). However, in recent years, the use of parallel or double translation (i.e., two translated versions) has been advocated as a preferred method of achieving equivalence in meaning (Hambleton 1993, 1994; Van de Vijver and Hambleton 1996). This can be used in conjunction with a committee approach to review the meaning and equivalence of translations and to select the “best” translation (Harkness 2003; Harkness and Schoua-Glusberg 1998). Extensive checking, pretesting of the translation, and debriefing are essential to ensure a reliable and accurate translation.

In marketing, back translation is still the primary method used to check translation accuracy. A review of *Journal of International Marketing* since its inception (1993–2005; i.e., 52 issues) identified a total of 45 articles that report surveys that use multiple languages. Of these surveys, 17 were in European languages, 3 were in Russian and Kazakh, 1 was in Arabic, and 28 were in Asian languages (Japanese, Thai, Chinese, Vietnamese, and Korean). In 75% (34) of these, back translation was used, and in 6 cases, a committee checked items in the back translations for equivalent meaning across languages. In 5 cases, parallel or double translation was undertaken, and a committee reviewed and adjudicated these to determine which translation should be used. In 1 case, a computer system (IADS) capable of handling multiple languages was used, and in 3 cases, no back translation was undertaken. Only 14 cases used pretesting of the questionnaire to ensure that the respondents clearly understood the questions. This typically entailed asking respondents the

RELIANCE ON BACK TRANSLATION

meaning of the question and further probing their understanding of the meaning or significance of the questions.

The review suggested that the most common form of checking the accuracy of translation in marketing studies was back translation. However, although there were limited details in many cases, it appears that beyond some modification after a single translation or back translation iteration by bilingual translators, relatively few additional checks or tests were undertaken to ensure the adequacy of the translation. A mechanism that is particularly useful for detecting translation problems, pretesting, was used in fewer than one-third of the studies.

As we noted previously, although back translation can potentially identify translation errors, in isolation, it is subject to several limitations (Brislin 1970; Van de Vijver and Hambleton 1996). Because bilingual translators are fluent in both the relevant languages, they do not necessarily use a language in the same way as monolingual people and may have common ways and conventions for moving from one language to another. Bilingual translators may also be able to make sense of a poorly written target translation (Brislin 1970). Because back translation provides a direct or literal translation from one language to another, it is possible to move from one language to another and back again without capturing the intended sense of the statement. For example, this may occur with idioms. Harkness (2003) notes an example from a German general social survey. An item in the survey is “*Das Leben in vollen Zügen genießen.*” Literally translated into English, this means “Enjoy life in full trains.” (One reading of *Zügen* is the plural of “train”; the other, idiomatic meaning is “in full draughts.”) A back translation may produce the same wording as the original German, which would suggest that the translation was accurate. However, the more appropriate translation into English would be “Live life to the full” or, in American English, “Live life to the fullest.”

The limitations of translating an existing questionnaire by relying solely on back translation can be observed in the work of Douglas and Nijssen (2003). In research conducted in the Netherlands, they used the CETSCALE (Shimp and Sharma 1987), which was initially developed in the United States. They tested two versions of the CETSCALE, a literal back-translated version and a modified version that attempted to address differences in consumer ethnocentrism in the Netherlands, namely, the absence of domestic brands in certain product categories. Consequently, an item such as “It is not right to buy foreign products” had limited relevance. Because of this problem, the authors conducted in-depth interviews to examine the meaning and significance of the construct to Dutch consumers. They also obtained reac-

tions to the Dutch version of each back-translated question. On the basis of qualitative research, they developed a modified version of the questionnaire that they believed to be more faithful to the construct as it existed in the Netherlands. Although the back translation did not change the meaning of any of the original scale items, the modified questions were clearer and better adapted to the Dutch concept of consumer ethnocentrism. The analytical portion of the study found that the modified questionnaire was also superior.

In addition, back translation assumes an etic approach to linguistic translation or, in other words, that there is always an equivalent word or construct in the target language. As a result, in some cases, a complete phrase may be necessary to clarify a construct. For example, Triandis (1972) states that the concept of *philotomo* is unique to the Greek culture and may need to be translated as a phrase such as “the love of honor” to capture the full meaning. Consequently, back translation is likely to be most useful when a literal or direct translation is required, but it is less helpful when idioms need to be translated or when the equivalence of a term or construct in another language needs to be established.

The problems associated with back translation suggest that simply relying on the back translation of a questionnaire without a careful review and an equivalence check is unlikely to provide an appropriate solution to determining whether a translation is adequate. Subtle nuances in the use of language, idioms, and the lack of equivalence of words or phrases also need to be considered. This implies that direct or literal translation is rarely adequate in and of itself. Rather, some adjustment in the use of language structure or phraseology is typically necessary to ensure equivalence in meaning (Harkness 2003).

The complexity of translation, the subjective character of translation assessment, and the multiple skills required (i.e., both linguistic ability and skills in questionnaire development) imply that a team-based approach to translation is needed. In the social sciences, this has been found to provide the richest output in terms of alternative translations and a balanced critique of alternative versions (Harkness and Schoua-Glusberg 1998; McKay et al. 1996). A team can bring together the mix of required skills and the disciplinary expertise required. Team members need to have knowledge of the study and of questionnaire design, as well as the cultural and linguistic skills to translate the questionnaire into appropriate versions of the target language when necessary.

Two main forms of collaborative-based work can be used in questionnaire translation: the committee approach and the

COLLABORATIVE APPROACHES TO INSTRUMENT TRANSLATION

expert team approach. In the committee approach, much of the work is done together, with collaborators working in a group. In expert team approaches, team members work individually rather than as a group. This procedure may be appropriate, for example, when the team is geographically scattered and operates as a virtual team.

Regardless of the approach taken in developing a final translation of a questionnaire, five basic stages are involved: translation, review, adjudication, pretesting, and documentation (Harkness 2003). Documenting translation procedures is particularly desirable when the questionnaire is likely to be repeated (e.g., an annual or biannual survey). As we noted previously, these stages often result in an iterative process. Review or adjudication may result in retranslation, and pretesting may result in a further round of translation review and testing.

A committee-based application of this procedure begins with a form of translation, commonly, parallel translation (Schoua-Glusberg 1992). In this case, two or more translators make independent, parallel translations of the questionnaire into the target language (Brislin 1980). A review meeting is then held in which the translators and at least one independent reviewer discuss the various versions and decide on a version for final review. An adjudicator may be present at the meeting and be a reviewer, or the final review version may be passed on for adjudication. In some cases, the researchers conducting the study may act as adjudicators, especially if they are familiar with the appropriate languages.

In many cases, committee-based review and adjudication are merged, depending on team expertise, schedules, previous experience with similar surveys or languages, and so forth. Alternatively, two committee rounds can be used, one to review and revise the translated versions and one to decide whether to accept the revised version. This may be appropriate if the researchers need time to consult on which version to adopt—for example, if the questionnaire is complicated or highly technical. This may also be desirable if the researchers are not familiar with one or more of the languages involved.

NEW VERSUS EXISTING RESEARCH INSTRUMENTS

In approaching multicountry research translation, researchers may encounter two different situations. The first situation occurs when a new instrument is being developed in multiple languages, and the second situation occurs when an instrument has already been developed in one language and needs to be translated into another language or multiple languages.

In the first case, there is an opportunity to eliminate the dominance of the structure of a source language or culture and to establish equivalence of different questions in each language (Werner and Campbell 1970). When the content of the questionnaire has been established, a team of bilingual or multilingual researchers and translators, preferably a researcher from each country and/or linguistic context, should examine it. A researcher establishes a version in his or her own language and then translates it, if necessary with a translator, into an agreed-on common language. The team then reviews, discusses, and harmonizes each version until a comparable and, as far as possible, equivalent version is established in each language (Van de Vijver and Leung 1997). In some cases, when the same concept cannot be expressed or tapped in the same way in each language, this may result in emic questions in a version. This is likely to occur with attitudinal scales or with activities such as sports, which vary in popularity from one country to another.

In the second, more common case, in which an instrument has already been established in one language, questions need to be translated so that they capture the same meaning in each language. As we noted previously, this is complex and typically requires an iterative process and multiple revisions of the translation to ensure that an appropriate translation is achieved, particularly when multiple languages are involved. In the remainder of the article, we focus primarily on this case, which is the situation that most international marketing researchers face.

Most of the literature related to translation and, in particular, to the use of back translation has focused on translation of the instrument in general. However, there are important issues related to the type of data being collected. International surveys typically involve one or more of three different types of data: socioeconomic, behavioral, and attitudinal data. Socioeconomic data are usually the easiest to translate because similar or equivalent categories can typically be identified. Behavioral categories require the identification of similar types of behavior that are equivalent. This may be more problematic, especially when the relevant behavior is cognitive rather than physical. Most difficulties are likely to occur with attitudinal data because the construct may differ from one language to another. Consequently, it is not just linguistic equivalence that needs to be considered but also other equivalence issues, such as category equivalence, functional equivalence, and construct equivalence (see Craig and Douglas 2005).

The relevance of each of these issues depends to a substantial degree on the nature of the question. Sociodemographic questions are most likely to raise the issue of category equiv-

INDIVIDUAL ITEM ISSUES

Types of Questions

alence related to employment, position with the firm, marital status, and so forth. Here, it is important to ensure that categories are equivalent, have similar status, and perform similar functions in each context under examination. For example, the tasks performed by a lawyer in the United States would be performed by either a barrister or a solicitor in the United Kingdom. Racial or regional demarcations also may be somewhat ambiguous, and in some cases, there may be reluctance to answer such questions.

Questions related to lifestyle and behavior are the most likely to raise the issue of functional equivalence. For example, questions related to washing or doing the laundry would need to consider the mode and circumstances of washing—for example, whether hot water is available, whether washing machines are commonly used, or whether washing is commonly done by domestic help or by services outside the home. Similarly, engaging in sports may be viewed either as a purely social activity or as a means of gaining exercise and improving health. Therefore, questions about leisure-time activities would need to take this into consideration.

Attitudinal or abstract questions are the most likely to raise the issue of conceptual equivalence. The linguistic equivalence of words such as “happiness,” “grief,” or “mourning” may have different connotations and tap other related constructs (Smith 2003). Equally, an apparently equivalent word may reflect different levels of intensity. Use of multiple indicators that are linguistically distinct may help resolve this problem (Przeworski and Teune 1966) but only if different terms do not replicate the same differences in intensity.

Individual words can also present difficulties in obtaining an accurate translation. Soriano and Foxall (2002) apply Mehrabian and Russell’s emotionality scales in Venezuela. They employ what can best be described as an iterative back-translation approach with pretesting to arrive at the final instrument. Of the 36 adjectives used in the English version of the scale, 20 were translated exactly. Either the remaining 16 adjectives did not have a direct Spanish equivalent, or the preferred usage varied between Spain and Venezuela. For example, the English word “annoyed” could be translated as *enfadado* or *molesto*. The preferred choice in Spain would be *enfadado*, but in Venezuela, it should be translated as *molesto*. Some word pairs had no direct equivalent in Spanish, and a different pair needed to be used. Notably, most of the translation difficulties were detected through pretesting and not through back translation, which accurately reproduced all but three of the adjective pairs used in the scale.

Scale Anchoring

In addition to considering the translation of questions in a survey, establishing equivalency in translating response

scales and response format is also important. Specifically, it is necessary to ensure that endpoints and intervals on a scale are equivalent (Mullen 1995). In this context, three alternatives have been suggested to improve equivalence and to mitigate language-related issues. These include using nonverbal scales, limiting the number of response options, and better calibrating label strength.

First, the use of nonverbal scales, such as visual or numerical scales, can aid in avoiding some of the difficulties in establishing linguistic equivalence of response scales. However, they do not completely eliminate translation issues, because the instructions for using the scale need to be translated as well as the meaning of the scale and the endpoints of the scales. Nonverbal and numeric scales are also subject to their own problems of differences in meaning and equivalence across cultures. Numeric meaning of endpoints of scales may vary from country to country. For example, in Germany, “1” typically refers to the first or best, resulting in a reversal of numeric scales compared with the order that is typically implied in the United States. Consequently, they may not necessarily be equivalent.

A second possible solution is to use dichotomous scales with responses, such as yes/no, favor/disapprove, or agree/disagree, which have equivalents across languages. However, this assumes the feasibility of writing a question with a dichotomous response, which may not necessarily be easy with attitudinal questions. It also requires the establishment of an equivalent “no-opinion” or “indifferent” category and also a “do-not-know” category.

The third alternative is to calibrate the response scale so that the strength of verbal scales used is equivalent across languages. Here, two cases may be identified. In the first case, only the ends of the rating or ranking scale are labeled. This typically poses fewer problems than when each position on the scale needs to be translated. However, there may be difficulties in finding translations of points such as “strongly agree/strongly disagree” and “somewhat agree/somewhat disagree” that capture the same level of intensity across languages. For example, in the Euro barometer, which measures social and political attitudes in the European Union on an annual basis, the French and English scales differ in structure and, to a lesser extent, in semantics (Harkness 2003). Both use the semantic dimension of agree, but in the French scale, the use of *d'accord/pas d'accord* suggests a unipolar scale, whereas the English scale uses a bipolar construction in which the wording is linguistically symmetrical with the endpoints modified by “strongly.” Equally, the “do-not-know” category is “cannot choose” in English compared with *ne sais pas* or “do not know” in French.

Even if the words are accurately translated, there still may be issues of equivalence in meaning. Reliance on back translation will provide accurate translations but may fail to provide words that generate equivalent responses across countries. As we indicated previously, words in isolation pose greater difficulty because they lack a context to help the translator render an accurate translation. Voss and colleagues (1996) examine the semantic comparability of adjectives used to label points of scales in three languages (English, Japanese, and Chinese). They employ two techniques, numeric estimation and line production, to assess interword distances between adjectives. Both techniques produced comparable results. This approach enables the researcher to calibrate the meaning more precisely and to develop scales with equivalent scale points across countries. In addition to using a formal methodology to achieve metric equivalence, Voss and colleagues use meetings and discussions among the translators to arrive at comparable words.

Pretesting the Questionnaire

After the final version of the translation has been approved, the next stage is to pretest the translation for comprehension, clarity, and coverage. Here, it is important to ensure not only that the translation represents a faithful rendering of the original text but also that the relevant target population clearly understands its meaning. Attention should be paid to everything included in the study, such as interviewer instructions, show cards, and other visual aids. Various procedures can be used to evaluate a translated questionnaire, including both qualitative and quantitative approaches, though qualitative procedures are more common because they provide richer insights into the exact nature of translation issues.

Qualitative tests include, for example, feedback (debriefing) from monolingual or bilingual respondents and field staff. In this case, the questionnaire is administered to either monolingual or bilingual respondents, who are then asked to provide feedback about their understanding of the meaning of questions, the ease of comprehension, clarity, and so forth. Typically, it is desirable to conduct pretests with both monolingual and bilingual respondents because, as we noted previously, bilingual people's use of language, vocabulary, and interpretation of terms may differ. In the case of bilingual respondents, half are given the question in one language, half are given the question in another language, and the similarity of responses is compared. Field staff who have administered a pretest can provide feedback on the clarity of questions, difficulties in respondents' comprehension, and so forth. This may be particularly desirable with semiliterate or low-income respondents because they may not understand certain words, concepts, or terms. Local terms or usage may also be common among such populations. Similarly, the ability to understand and use support materials, such as show cards

and response scales, can be evaluated. Focus groups can also be used to evaluate questionnaires and to probe the meaning of different words and phrases. This may be especially helpful in the case of abstract attitudinal questions.

Quantitative methods can also be used to evaluate pretests when a sufficiently large sample is available. These include tests that are based on item response theory, double ballot administration, and bilingual fielding splits (Harkness and Schoua-Glusberg 1998). In the case of double ballot administration, bilingual respondents are provided with both language versions of the questionnaire, and differences in responses are compared. The results of pretesting may result in some modification or readjustment of the questionnaire, in which case it could be pretested again. As we noted previously, this becomes an iterative process until a satisfactory version of the questionnaire is developed.

In international marketing research, translation is a critical issue in the development of a valid and effective data collection instrument. Although the starting point for any research project is the underlying issue or conceptual framework under investigation, the vehicle that drives survey research is the questionnaire. International marketing research necessarily involves different countries in which different languages or language variants are spoken. It is critical to ensure that translated questions have equivalent meaning in each context, are devoid of cultural bias, and can be clearly understood by all respondents.

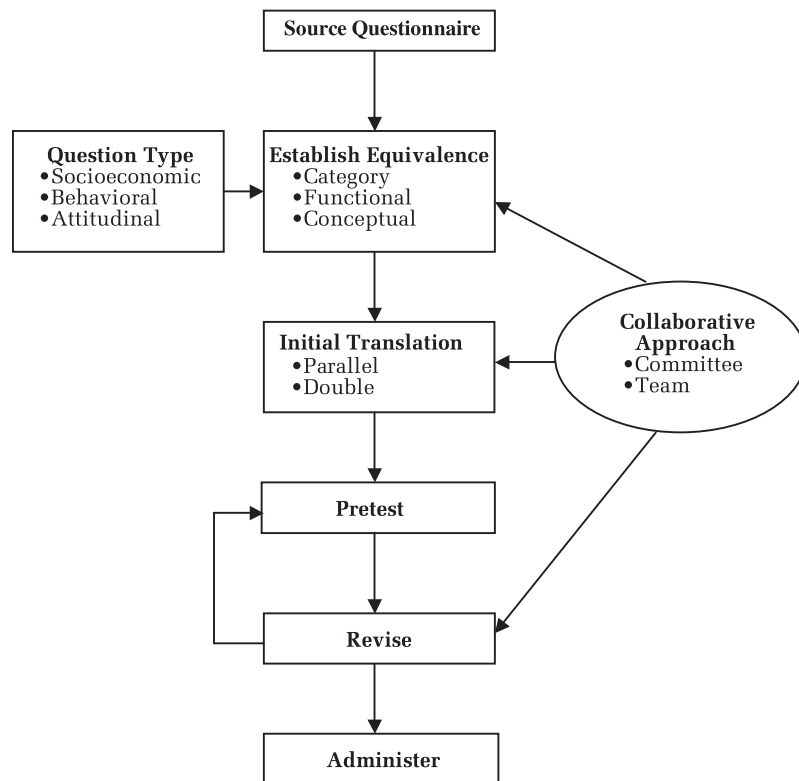
For the reasons we discussed previously, a single back translation in isolation does not provide a reliable instrument in the target language. Although this helps ensure a correct literal translation, it does not help assess whether the questions have equivalent meaning or are devoid of cultural bias. Even in catching errors, much depends on the back translator's understanding of a question and its purpose. Furthermore, a mechanistic application of back translation, without adjustments, is unlikely to result in an instrument that will produce reliable and valid results. Rather than continue to employ back translation, we recommend that researchers follow the steps we outline subsequently when translating questionnaires into other languages. Following these steps will result in a questionnaire that produces reliable and valid results. We summarize the steps in Figure 1 and elaborate on them in the following subsections.

The limitations of back translation mean that it is critical to use a team or committee approach to ensure a reliable and effective translation. This team or committee should consist of people who are familiar with each of the cultures in which research is being conducted and the different languages

RECOMMENDED APPROACH TO QUESTIONNAIRE TRANSLATION

Employ a Collaborative Approach

Figure 1.
Collaborative and Iterative
Questionnaire Translation



used. The collaborative approach ensures that different points of view are represented. Furthermore, it is iterative by nature and helps ensure that the “best” translation evolves.

When a new set of questionnaires is being developed for a project that will be conducted in multiple languages and different cultural contexts, it is important to ensure not only that the questions have equivalent meaning in each context but also that the questionnaire is decentered, which often requires multiple iterations of translation. This helps ensure that one culture or language structure does not dominate other questionnaire versions. When research has already been conducted in one country and a questionnaire needs to be translated into another language or multiple other languages, the committee needs to ensure not only that the target language version is faithful to the source version but also that it will elicit equivalent types of responses.

Establish Equivalence

Different types of data (i.e., questions related to socioeconomic, behavioral, and attitudinal data) all pose different issues and thus need to be examined to ensure that they are couched in terms that are similar in their context to the source questions and will elicit equivalent responses. Category equivalence and functional equivalence are relatively straightforward to achieve, but they still require knowledge of the local conventions and the cultural context. Construct equivalence is more difficult to achieve because there are

instances in which the construct under examination is unique to a particular culture. A collaborative approach with researchers from each culture can help identify and resolve issues of construct equivalence. Similarly, scale anchors, visual stimuli, and interviewer instructions need to be translated so that they are understood and elicit equivalent responses in each language. Again, some degree of cultural adaptation may be required.

After the equivalence of the questions has been examined to ensure that they are expressed in similar terms and are likely to elicit similar responses, the next step is to conduct an initial translation. When there is an existing questionnaire, there should be independent, parallel translation into the target language. Two separate translators should each translate the questionnaire into the target language. A review meeting should then be held with the translators and an independent reviewer to decide on the final version. A further round may be needed to resolve inconsistencies and ensure that the questionnaire accurately captures the same meaning in each country.

After the researchers and translators arrive at what they believe is an appropriate instrument, the next step is to pretest each language version of the questionnaire. This is an important step and one that is often ignored. Only 31% (14 of 45) of the studies reported in *Journal of International Marketing* incorporated this step. The initial version of the questionnaire can be pretested on monolingual or bilingual respondents. Pretesting, followed by debriefing, can find problems of comprehension or meaning that were not identified previously. Issues identified in pretesting are referred back to the committee to be incorporated into the final version of the questionnaire. Without this step, there is no way to know how actual respondents (versus translators) will react to the instrument.

In addition to employing a collaborative approach and establishing equivalence, it is important for the researcher not to view the initial translation as the final translation. There is a natural tendency to want to get the research into the field and to get the results as quickly as possible. However, without an accurate translation of the instrument, the results of the research may not be valid. Committee and team approaches are intrinsically iterative because there will inevitably be differences of opinion and interpretation that need to be resolved. Multiple iterations, though time consuming, help ensure that the translated instrument is appropriate for the new context.

Although this approach may appear somewhat complex and time consuming, it is nonetheless essential to consider the

Conduct an Initial Translation

Pretest the Questionnaire

Adopt an Iterative Approach

issues involved at each step to ensure that a reliable and conceptually equivalent translation is achieved. The time taken in this approach will depend to a large degree on the experience of the researcher or company that is undertaking the approach and the number of languages involved. When a market research company or academic researcher conducts a survey in a country or language for the first time, it is important to undertake this approach to establish equivalence and to ensure the integrity of the translation. A local research company, a local manager, or an academic researcher familiar with the language can be used at the committee stage to check the translation, help select the final version, and decide whether revisions are needed after the pretesting. Conversely, if the company or academic researchers already have extensive experience in conducting similar surveys in a given country or language, they are likely to be already familiar with relevant linguistic equivalents; thus, parallel translation may not be necessary, and the checking process is likely to be more expeditious. However, pretesting is essential to detect any potential errors, though fewer modifications and only one iteration are likely to be needed.

CONCLUSION

Instrument translation is a crucial element when conducting international marketing research. Unless careful attention is paid to equivalence in meaning and to the constructs tapped in each context, the results will not be comparable from one linguistic context to another. As a result, inappropriate conclusions may be drawn, thus negating the value of the research. Reliance on back translation as the sole means of achieving an accurate foreign language translation is unlikely to result in equivalence. Consequently, researchers need to employ the proposed iterative approach to ensure that valid and reliable multicountry instruments are developed.

NOTE

1. Target language is the language into which the translation is being made.

REFERENCES

- Brislin, Richard W. (1970), "Back Translation for Cross-Cultural Research," *Journal of Cross-Cultural Psychology*, 1 (3), 185–216.
- (1980), "Translation and Content Analysis of Oral and Written Material," in *Handbook of Cross-Cultural Psychology*, Vol. 1, Harry C. Triandis and J.W. Berry, eds. Boston: Allyn & Bacon, 389–444.
- Craig, C. Samuel and Susan P. Douglas (2005), *International Marketing Research*, 3d ed. Chichester, UK: John Wiley & Sons.
- Douglas, Susan P. and Edwin J. Nijssen (2003), "On the Use of 'Borrowed' Scales in Cross-National Research: A Cautionary Note," *International Marketing Review*, 20 (6), 621–42.
- Hambleton, R.K. (1993), "Translating Achievement Tests for Use in Cross-National Studies," *European Journal of Psychological Assessment*, 9 (1), 57–65.

- (1994), "Guidelines for Adapting Educational and Psychological Tests: A Progress Report," *European Journal of Psychological Assessment*, 10 (3), 229–44.
- Harkness, Janet A. (2003), "Questionnaire Translation," in *Cross-Cultural Survey Methods*, Janet A. Harkness, Fons J.R. van de Vijver, and Peter Ph. Mohler, eds. New York: John Wiley & Sons, 35–56.
- and A. Schoua-Glusberg (1998), "Questionnaires in Translation," in *ZUMA-Nachrichten Spezial, No. 3: Cross-Cultural Survey Equivalence*, Janet A. Harkness, ed. Mannheim, Germany: ZUMA.
- McKay, R.B., M.J. Breslow, R.L. Sangster, S.M. Gabbard, R.W. Reynolds, J.M. Nakamoto, and J. Tarnai (1996), "Translating Survey Questionnaires: Lessons Learned," *New Directions for Evaluation*, 70, 93–105.
- Mullen, Michael R. (1995), "Diagnosing Measurement Equivalence in Cross-National Research," *Journal of International Business Studies*, 26 (3), 573–96.
- Przeworski, A. and H. Teune (1966), "Equivalence in Cross-National Research," *Public Opinion Quarterly*, 30 (Winter), 551–68.
- Schoua-Glusberg, A. (1992), *Report on the Translation of the Questionnaire for the National Treatment Improvement Evaluation Study*. Chicago: National Opinion Research Center.
- Shimp, T.A. and S. Sharma (1987), "Consumer Ethnocentrism: Construction and Validation of the CETSCALE," *Journal of Marketing Research*, 24 (August), 280–89.
- Smith, Tom W. (2003), "Developing Comparable Questions in Cross-National Surveys," in *Cross-Cultural Survey Methods*, Janet A. Harkness, Fons J.R. van de Vijver, and Peter Ph. Mohler, eds. New York: John Wiley & Sons, 69–92.
- Soriano, Mirella Yani and Gordon R. Foxall (2002), "A Spanish Translation of Mehrabian and Russell's Emotionality Scales for Environmental Consumer Psychology," *Journal of Consumer Behavior*, 2 (1), 23–36.
- Triandis, Harry C. (1972), *The Analysis of Subjective Culture*. New York: John Wiley & Sons.
- Van de Vijver, Fons and R.K. Hambleton (1996), "Translating Tests: Some Practical Guidelines," *European Psychologist*, 1 (4), 89–99.
- and Kwok Leung (1997), *Methods and Data Analysis for Cross-Cultural Research*. Thousand Oaks, CA: Sage Publications.
- Van Herk, Hester (2003), *Equivalence in a Cross-National Context: Methodological and Empirical Issues in Marketing Research*. Tilburg, the Netherlands: Tilburg University Press.
- Voss, Kevin A., Donald E. Stem, Lester W. Johnson, and Constantino Arce (1996), "An Exploration of the Comparability of Semantic Adjectives in Three Languages: A Magnitude Estimation Approach," *International Marketing Review*, 13 (5), 44–58.
- Werner, Oscar and Donald Campbell (1970), "Translating, Working Through Interpreters and the Problems of Decentring," in *A Handbook of Method in Cultural Anthropology*, R. Naroll and R. Cohen, eds. New York: Columbia University Press, 398–420.

THE AUTHORS

Susan P. Douglas is Paganelli-Bull Professor of Marketing and International Business (e-mail: sdouglas@stern.nyu.edu), and **C. Samuel Craig** is Catherine and Peter Kellner Professor of Marketing (e-mail: scraig@stern.nyu.edu), Stern School of Business, New York University.

ACKNOWLEDGMENTS

The authors acknowledge helpful comments from Dawn Lesh, Durairaj Maheswaran, and Tom Meyvis on a previous draft of this article.