Methodological issues in cross-cultural marketing research

A state-of-the-art review

Naresh K. Malhotra
Georgia Institute of Technology, Atlanta, Georgia, USA

James A. Garwal
University of Regina, Saskatchewan, Canada

and Mark Peterson
University of Texas at Arlington, Texas, USA

To be cross-cultural, the marketing research project must be conducted across nations or culture groups, rather than across provinces or ethnic groups (Berry, 1980). For the purpose of this review, international marketing research (research dealing with international issues), foreign research (research conducted in a country other than the country of the research-commissioning organization), or multinational research (research conducted in all or all important countries where the company is represented), and other similar terms, will be subsumed under the rubric of cross-cultural research. We do not deny the fine distinctions that can be made among these concepts (e.g., Terpstra and Sarathy, 1990). However, it is not necessary to make these distinctions as the methodological issues we consider apply in varying degrees to all of these types of research.

In recent years cross-cultural marketing research has assumed great importance in the academic and business worlds. Academically, cross-cultural research has gained wide acceptance both in international business journals and in specialized journals. A recent review identified 720 articles on the subject that were published in various academic journals between 1980 and 1990 (Aulakh and Kotabe, 1993). Businesswise, the USA accounts for only 39 per cent of the marketing research expenditures worldwide. About 40 per cent of all marketing research is conducted in Western Europe and 9 per cent in Japan. Most of the research in Europe is done in Germany, the United Kingdom, France, Italy, and Spain (Demby, 1990). With the globalization of markets, marketing research has assumed a truly international character and this trend is likely to continue (e.g., Malhotra et al., 1994).

Despite its burgeoning growth in recent years, further expansion and development of cross-cultural marketing research is being hampered by...
methodological problems. “While theory development in international marketing showed considerable progress in the last ten years, advancement in the research methodologies has generally tended to lag behind” (Aulakh and Kotabe, 1993, p. 24). Similar observations have been made by earlier reviewers. For example, Bodewyn (1981) lamented that research designs developed in the USA were adopted in other cultures without appropriate modifications. Bilkey and Ness (1982) in their review of country-of-origin (COO) studies, identified several methodological problems which plagued this stream of research[1]. Despite the clarion call of Bilkey and Ness (1982), methodological problems in this area have continued, limiting our understanding and constraining theory development. Even after 30 years of research in which about 200 articles have appeared on COO, a recent meta analysis concluded that “the inconsistent patterns of effect size differences found for the perception and intention variables across the study characteristics investigated suggest that the COO phenomenon is still not well understood” (Peterson and Jolibert, 1995, p. 894).

While articles dealing with cross-cultural research methodologies have appeared from time to time (e.g. Albaum and Peterson, 1984; Green and White, 1976; Mullen, 1995; Parameswaran and Yaprak, 1987; Sekaran, 1983; Sekaran and Martin, 1982; Singh, 1995) a unified treatment of such issues has been lacking in the marketing literature. As Aulakh and Kotabe (1993, p. 24) plead, “the research methodology issue is one area in need of immediate attention to make international marketing research more rigorous”.

This paper is an attempt to address this need. Our objective is to sensitize cross-cultural marketing researchers to the various methodological issues which should be considered if unequivocal findings are to emerge. We also provide guidelines for addressing these issues. These methodological issues are organized around the six-step framework describing the marketing research process: problem definition, developing an approach, research design formulation, field work, data analysis, and report preparation and presentation (Malhotra, 1992a, 1996). The issues involved in each step of the marketing research process are described in Figure 1.

**Problem definition**

The precise definition of the marketing research problem is more difficult, and more important, in cross-cultural marketing research than in domestic marketing research. Unfamiliarity with the cultures and environmental factors of the countries where the research is being conducted can greatly increase the difficulty of attaining comparability. To compare two phenomena, they must share some features in common, and they should also differ on some features. Comparability may be attained by adopting universals from other disciplines or by demonstrating equivalence of psychological concepts and data across cultural groups. To establish such dimensional identity, universals can be adopted from biology, linguistics, anthropology, or sociology. For example, one could adopt as universals a list of primary needs from biology, a list of common cultural components (language, myth, etc.) from anthropology, a set of
functional prerequisites for social life (e.g. role differentiation) from sociology
(Lonner and Triandis, 1980).

Further evidence of dimensional identity can be obtained by examining the
functional, conceptual, and metric equivalence (Berry and Dasen, 1972). Functional
equivalence implies that the phenomenon or behaviour in two or
more cultures is related to the same functional problem. Thus, functional
equivalence of behaviour exists when the behaviour in question has developed
in response to a problem shared by two or more cultures, even though the
behaviour in one culture may be superficially quite different from the behaviour
in another. Conceptual equivalence implies that the meaning of research concepts, stimuli, and materials should be equivalent across cultures. Thus, the marketing research problem should be defined in such a way that it is conceptually equivalent for individuals in the various cultures being investigated. It should be noted that many concepts are culture bound and are inappropriate for use on a cross-cultural basis. Metric equivalence exists when the psychometric properties of data sets from different cultural groups exhibit the same coherence or structure. The issue of equivalence is considered in more detail in the section on measurement.

Comparability is a prerequisite for valid cross-cultural comparisons (Berry, 1980). Many international marketing efforts fail not because research was not conducted, but because the issue of comparability was not adequately addressed in defining the marketing research problem. As a practical way of attaining comparability, it has been suggested that the researcher should isolate and examine the impact of the self-reference criterion (SRC), or the unconscious reference to one's own cultural values. The following steps help researchers to account for environmental and cultural differences and to define the problem in comparable ways (Douglas and Craig, 1983; Lee, 1966):

1. Define the marketing research problem in terms of domestic environmental and cultural factors. This involves an identification of relevant domestic country traits, economics, values, needs or habits.

2. Define the marketing research problem in terms of foreign environmental and cultural factors. Make no judgements. This involves an identification of the related traits, economics, values, needs or habits in the proposed market culture. This task requires input from researchers who are familiar with the foreign environment.

3. Isolate the self-reference criterion (SRC) influence on the problem and examine it carefully to see how it complicates the problem. Examine the differences between steps 1 and 2. If differences are found, they can be attributed to the SRC.

4. Redefine the problem without the SRC influence and address it for the foreign market situation. If the differences in step 3 are significant, the impact of the SRC should be carefully considered.

Consider the broad problem of the Coca-Cola Company trying to increase its penetration of the soft drink market in India. In step 1, the problem of increasing the market penetration in the USA would be considered. In the USA, virtually all households consume soft drinks, and the problem would be to increase the soft drink consumption of existing consumers. Furthermore, soft drinks are regularly consumed with meals and as thirst quenchers. So the problem of increasing marketing penetration would involve getting the consumers to consume more soft drinks with meals and at other times. In India, on the other hand (step 2), a much smaller percentage of households consume soft drinks. Water, rather than soft drinks, is consumed with meals. Soft drink consumption
is restricted to serving guests and to special occasions. Therefore, the management problem of increasing the market share of a soft drink brand would translate into a different marketing research problem in India from that in the USA. Thus, in step 3, the SRC can be identified as the American notion that soft drinks are an all-purpose, all-meal, beverage. In step 4, the problem in the Indian context can be defined as how to get a greater percentage of the Indian consumers to consume soft drinks (Coca-Cola products) and how to get them to consume soft drinks (Coca-Cola products) more often for personal consumption (Malhotra, 1996). When the problem has been appropriately defined, the task of developing an approach becomes easier.

Developing an approach

There are different approaches to conducting cross-cultural research including anthropological, sociological, and psychological perspectives (McCort, 1992). The anthropological approach attempts to make a direct assessment of the cultural processes and behaviours (e.g. Arnould, 1989). There are two schools of thought. The cultural realists equate culture with behaviour and conceptualize culture as a concrete reality of its own, independent of people. The people are viewed as catalysts to a cultural reality which is governed by its own laws. This thought leads to the idea that culture predetermines behaviour. The views of cultural realists have been challenged by intracultural and individual differences observed in motivations, cognition, and behaviours (Goodenough, 1981). The second school, the cultural nominalists, view culture as a system of meaning and that culture can be inferred or measured only indirectly from behaviour and psychological functioning (Poortinga and Van de Vijver, 1987).

The sociological approach focuses on the effects of social forces on behaviour (Lee, 1990). There is an emphasis on cultural universals which serve as prerequisites for social life such as role differentiation, normative regulation of behaviour, and socialization. A psychological perspective is concerned with processes through which people personalize social influences in their own cognitive organization (Triandis, 1984). Here again, there are at least two schools of thought: the postmodernists and the cognitive researchers (McCort, 1992). The postmodern approach relies on theory-driven interpretation of actual behaviour or elements of physical culture. Symbolic dimensions of consumption are explored by studying behaviour rather than cognition (Belk, 1987; Holbrook, 1989). On the other hand, cognitive researchers attempt to study cultural influences on cognitive processes and contents. External stimuli and internal psychological constructs are conceptualized in terms of cognitive structures and the process of interpretation is evaluated via explicit theories of cognitive relationships (Golden et al., 1989; Kleine and Kernan, 1991). It should be noted that the traditional anthropological and sociological approaches are at the group level, whereas psychological approaches are usually concerned with the individual.

From a marketing research viewpoint, it is appropriate to conceptualize culture as a knowledge system that is represented in cognitive processes and
expressed in behaviours (Triandis, 1984). Hence, the cognitive psychological perspective seems to be more appropriate for most marketing research applications. This perspective is also very useful for understanding basic processes of perception, cognition, emotion, and motivation (Lonner and Triandis, 1980), which are of major concern in most cross-cultural marketing applications. This, however, does not imply that marketing researchers should abandon the other perspectives. The other perspectives and approaches could be fruitfully used in a complementary manner. For example, the ethnographic field techniques in anthropology have generated rich cross-cultural data which could be very useful in conducting exploratory marketing research. In fact, marketing researchers should also tap into cross-cultural perspectives offered by other disciplines such as political science, criminology, economics, and public administration (Sekaran, 1983).

It is also important that the approach adopted should encompass both emic and etic viewpoints. The emic approach examines the phenomenon from within the system, investigates only one culture, and the criteria adopted are relative to internal characteristics of the culture. In contrast, the etic approach examines the phenomenon from a position outside the system, investigates many cultures, and the criteria adopted are considered absolute or universal. When these universals are assumed they have been termed imposed etic or pseudo etic (Berry, 1969). In contrast, a true etic or derived etic is one which is empirically derived from the common features of the phenomenon. While the emic and etic viewpoints are not without critics, as stated by Berry (1980), the very name cross-cultural embodies both these approaches. To be “cultural” requires the emic viewpoint, and “cross” requires the etic perspective. One could begin with the etic perspective characterized by the presence of universals. Emic descriptions can then be made by progressively altering the imposed etic until it matches an emic point of view. For example, with rapid globalization and technological networking, there seems to be a convergence of consumers’ preferences of products and services across cultures. In this case, consumer preference can be modelled from an etic approach. However, the motivations behind the preference may differ significantly across cultures. This would then require culture-based adaptation based on an emic approach. From a North American perspective, in general, one can hypothesize that cross-cultural research conducted in European countries would require relatively fewer emic adaptations as compared with Asian and Pacific-Rim countries. Research instruments should include items based both on etic and emic considerations. Approaches for the analysis of items applicable to all cultures and specific to each culture have been developed (e.g. Przeworski and Teune, 1970).

The approach developed should address and control for alternative explanations for the results. Plausible rival hypotheses for the results can be both substantive and methodological. A substantive alternative explanation may be that another cultural variable, not controlled by the researcher, accounts for the differential results obtained. Methodologically, the differences obtained may be attributed to sampling problems, translational inadequacies,
interviewing and data collection difficulties, and a host of other problems. Many stable reference points should be built into the approach so that valid interpretations of the results can be made.

**Research design**

In formulating a research design, considerable effort is required to ensure the equivalence and comparability of secondary and primary data obtained from different cultures. In the context of collecting primary data, qualitative research, survey methods, scaling techniques, questionnaire design, and sampling considerations are particularly important.

**Secondary data**

A wide variety of secondary data is available (e.g. Malhotra, 1992b; Takada and Jain, 1991). Evaluation of secondary data is even more critical for cross-cultural than for domestic projects. Different sources report different values for a given statistic, such as GDP, because of differences in the way the unit is defined. Measurement units may not be equivalent across cultures. In France, for example, workers are paid a thirteenth monthly salary each year as an automatic bonus, resulting in a measurement construct that is different from other countries (Douglas and Craig, 1983).

The accuracy of secondary data may also vary from country to country. Data from highly industrialized countries like the USA are likely to be more accurate than those from developing countries. For example, many countries attempt to attract foreign investment by overstating certain factors that make the economic picture look better than it is (Czinkota and Ronkainen, 1994). On the other hand, there may be some countries that understate certain factors, making their economic situation appear worse so that they can indicate a need for more foreign aid. Statistics could also be manipulated for political reasons. For example, a study conducted by the International Labour Organization found that the actual unemployment rate was over 10 per cent in Russia compared with official reported figure of 2 per cent. Still, in other countries where there is a lack of sophisticated data collection systems, estimates rather than precise readings are reported. Some international statistics may contain relatively high margins of error (such as 25 per cent) which may not be within the tolerance range of reliability for the international marketer. Business and income statistics are affected by the taxation structure and the extent of tax evasion. Population censuses may vary in frequency and year in which the data are collected. In the USA the census is conducted every ten years, whereas in the People's Republic of China, there was a 29-year gap between the censuses of 1953 and 1982.

An important source of cross-cultural secondary data of which marketing researchers may not be aware consists of ethnographic records describing various cultures. Ethnographic studies aim to outline a society's culture by describing what people must have learned in order to participate acceptably in the activities of that society. Observers isolate, inventory, and categorize people,
things, and events that characterize the society. Then the ways in which the
categories of people, things, and events are distributed on a variety of
dimensions are examined. Analyses of these distributions lead to formulations
of cultural expectations as rules of conduct and as principles governing group
membership, ritual performances, and all kinds of decisions. Several
ethnographic reports are available which can be used to code data from
different cultures. These data can then be analysed to develop generalizations
within a region or throughout the whole world. The various cross-cultural
ethnographic samples of secondary data include the Human Relations Area
Files (HRAF) Archive, the World Ethnographic Sample, Ethnographic Atlas (full
version, and the summary version), the Standard Cross-Cultural Sample, A
Standard Ethnographic Sample: second edition, the HRAF Probability Sample,
and Atlas of World Cultures. A brief description of each of these samples may
be found in Ember and Otterbein (1991). For a variety of marketing research
problems, analyses of such secondary data could provide useful insights for
formulating an appropriate research design.

Qualitative research
Because the researcher is often not familiar with the foreign market to be
examined, qualitative research is crucial in cross-cultural marketing research.
In the initial stages of cross-national research, qualitative research can provide
insights into the problem and help in developing an approach by generating
relevant research questions and hypotheses, models, and characteristics which
influence the research design. Thus, qualitative research may reveal the
differences between foreign and domestic markets. It may also help to reduce
the psychological distance between the researcher and the respondent/context.
In some cases the researcher must rely on qualitative research, since secondary
data may not be available. Goodyear (1982) points out some problems
associated with qualitative techniques in developing countries such as
accessibility (different concept of time), sampling (extended demographic
factors such as religion and tribal membership), shorter span of attention and
less familiarity with abstract thinking.

Focus groups can be used in many settings, particularly in industrialized
countries. The moderator should not only be trained in focus group
methodology but should also be familiar with the language, culture, and
patterns of social interaction prevailing in that country. The focus group
findings should be derived not only from the verbal contents but also from non-
verbal cues such as voice intonations, inflections, expressions, and gestures. In
some cultures, such as in the Middle or Far East, people are hesitant to discuss
their feelings in a group setting. In these cases, in-depth interviews should be
used.

The use of projective techniques in cross-cultural marketing research should
be carefully considered. Association techniques (word association), completion
techniques (sentence completion, story completion), and expressive techniques
(role playing, third-person technique) involve the use of verbal cues.
Construction techniques (picture-response and cartoon tests) employ non-verbal stimuli (pictures). Whether verbal or non-verbal stimuli are used, the equivalence of meaning across cultures should be established. This can be a difficult task if the sociocultural environments in which the research is conducted vary greatly. Establishing the equivalence of pictures can be particularly problematic. Problems in delineation, measurement, and control of relevant cultural variables are not easily addressed. Line drawings are subject to fewer problems of interpretation than photographs. The usual limitations of qualitative techniques also apply in the international context, perhaps to a greater extent. It is often difficult to find trained moderators and interviewers overseas. The development of appropriate coding, analysis, and interpretation procedures poses additional difficulties (Holtzman, 1980). Yet, projective techniques may be useful when employed in conjunction with focus groups and in-depth interviews. The use of qualitative techniques in cross-cultural research is easier to justify when they are used for exploratory purposes to understand cultural and group-level phenomena, rather than to make individual-level inferences.

Survey methods
We evaluate the major interviewing methods in the light of the challenges of conducting research in foreign countries, especially Europe and developing countries. The criteria used in evaluating the different methods focus on locating, contacting, and obtaining information from the respondents (Malhotra, 1994).

Telephone interviewing. In the USA and Canada, the telephone has achieved almost total penetration of households. As a result, telephone interviewing is the dominant mode of questionnaire administration. The same situation exists in some of the European countries (e.g. Sweden, The Netherlands and Switzerland). In many of the other European countries, telephone penetration is still not complete (e.g. Finland, Portugal). In most developing nations only a few households have telephones (e.g. many African countries, India, Brazil). Even in countries like Saudi Arabia, where telephone ownership is extensive, telephone directories tend to be incomplete and out of date.

Telephone interviews are most useful when employed with relatively upscale consumers who are accustomed to business transactions by phone or consumers who can be reached by phone and can express themselves easily. With the decline of costs for international telephone calls, multicountry studies can be conducted from a single location (De Houd, 1982). This greatly reduces the time and costs associated with the organization and control of the research project in each country. Furthermore, international calls obtain a high response rate, and the results have been found to be stable (i.e. the same results are obtained from the first 100 interviews as from the next 200 or 500). It is necessary to find interviewers fluent in the relevant languages, but in most countries this is not a problem.
Personal interviewing. Owing to high cost, the use of in-home personal interviews has declined in the USA, but this is the dominant mode of collecting survey data in many parts of Europe, newly industrialized countries (NICs), and the developing world (Hononichl, 1984; Monk, 1987). In-home personal interviewing is the dominant interviewing method in Switzerland. In Portugal face-to-face interviews are 77 per cent of the total interviews conducted (Queiros and Santos Lima, 1988). The majority of the surveys are done door-to-door, while some quick sociopolitical polls are carried out in the street using accidental routes.

In North America many marketing research organizations have permanent facilities in malls equipped with interviewing rooms, kitchens, observation areas, and other devices. While mall intercepts are being conducted in some European countries, such as Sweden, they are not popular in Europe or developing countries (Kaiser, 1988). In contrast, central location/street interviews constitute the dominant method of collecting survey data in France and The Netherlands.

Mail interviewing. Because of low cost, mail interviews continue to be used in most developed countries where literacy is high and the postal system is well developed. In countries where the educational level of the population is extremely high (e.g. Canada, Denmark, Norway, Sweden, The Netherlands, the United Kingdom and the USA), mail interviews are common (Vahvelainen, 1985; 1987). In Africa, Asia, and South America, however, the use of mail surveys and mail panels is low because of illiteracy and the large proportion of population living in rural areas. Mail surveys are, typically, more effective in industrial international marketing research, although it is difficult to identify the appropriate respondent within each firm and to personalize the address. Mail panels are extensively used in the UK, France, Germany, and the Netherlands (Bigant and Rickebusch, 1985). Mail and diary panels are also available in Finland, Sweden, Italy, Spain, and other European countries. Use of panels may increase with the advent of new technology.

Although mail surveys have been noted for low response rates, monetary incentives accompanying questionnaires consistently appear to be an effective technique to increase response rates. While this method has worked in some countries such as in the USA and Great Britain (Kanuk and Berenson, 1975), it may not necessarily work in other countries. In a comparative study of Japan and Hong Kong, Keown (1985) found that the response rate of the Hong Kong sample actually went down to 0 per cent with a one US dollar incentive and was 13 per cent without incentive. In comparison, the Japanese sample showed results in the expected direction: 50 per cent with incentive and 22 per cent without incentive. While monetary incentives may increase response rates in any culture, the type (local or foreign currency) and the amount may be important factors.

Selection of survey methods. No questionnaire administration method is superior in all situations. Table I presents a comparative evaluation of the major modes of collecting quantitative data in the context of cross-cultural marketing.
Cross-cultural marketing research (Malhotra, 1991; 1994). An important consideration in applying the criteria of Table I to select the methods of administering questionnaires is to ensure equivalence and comparability across cultures. Different methods may be appropriate in different cultures. In one country, a certain method, say mail survey, may be known to have a given level of suitability. In another country, in-home interviews rather than mail surveys may be equivalent in terms of the availability of sampling frame, sample control, accessibility of respondents, etc. Thus, mail surveys should be used in the first country, while in-home interviews should be conducted in the second country. In collecting data from different cultures, it is desirable to use survey methods which are equivalent in terms of flexibility of data collection, sample control, response rate, and potential for interviewer and respondent bias. Issues of equivalence are also salient in measurement and scaling.

Observational methods. Observational methods are well suited for cross-cultural research which involves sensitive topics. It is in these areas that subjects often do not provide truthful verbal responses owing to social desirability, the tendency to present a favourable image, and other legal and cultural constraints (Bochner 1986). Issues and problems facing domestic observational methods also apply to cross-cultural observational methods. Some of these are reactance, contextual effects owing to differences in natural settings, distorting the natural stream of behaviour, the difficulty of inferring motives and underlying cognitive processes for the behaviour observed, the relationship between variables over time, and individual differences. Some of these problems can be reduced by employing time-and-event sampling instead of studying the continuous stream of behaviour. However, in cross-cultural observation, another factor magnifies individual differences – cross-cultural differences.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Telephone</th>
<th>Personal</th>
<th>Mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>High sample control</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Difficulty in locating respondents at home</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Inaccessibility of homes</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Unavailability of a large pool of trained interviewers</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Large population in rural areas</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Unavailability of maps</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Unavailability of current telephone directory</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Unavailability of mailing lists</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Low penetration of telephones</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Lack of an efficient postal system</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Low level of literacy</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Face-to-face communication culture</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

**Note:** + denotes an advantage; – denotes a disadvantage
There are other observational methods which track and record behavior such as content analysis, trace analysis and audits. Some work has been done using content analysis in cross-cultural research although most of it dominates in cross-cultural advertising (Belk and Pollay, 1985; Gilly, 1988; Mueller, 1991; Tse et al., 1989). Brislin (1980) suggests the use of objectivity via predefined, explicit rules that need to be modified in different cultural settings. Bias-free predefined rules should be used to develop categories. Inter-judge or coder agreement involving one culture (such as domestic) generally tends to inflate the reliability. Hence, analysts and judges from different cultures should agree on coding the categories (Brislin, 1980; Gilly, 1988). It involves tedious coding and analysis, although microcomputers and mainframes can be used to facilitate the task. While content analysis has not been widely used in international marketing research, it offers promising potential.

Causal research design: experimentation. There are some inherent difficulties in conducting cross-cultural experimental designs. Malpass and Poortinga (1986) discuss these limitations as follows. First, unlike true experiments characterized by random assignment of subjects to treatments, in cross-cultural experiments researchers cannot gain control of which persons receive which cultural treatment and cannot manipulate treatments independently of other antecedent cultural variables. Second, because most cross-cultural research is generally characterized by quasi-experimental designs (since cultural groups are non-equivalent to begin with), the exchangeability of subjects across cultural treatments is low. In marketing, such sets of antecedent variables that identify a subject in a culture may include demographics, psychographics, and other culture-specific factors. A suggested strategy is to validate such differences by using the Human Relations Area Files. Third, in cross-cultural designs, because the researcher has no control over the effects of cultural factors on behaviour, only post-hoc inferences can be made. In testing differences on some hypothesized variable, because significant differences can also be a result of artifacts or bias variables, separate studies should be conducted to study such cultural variables.

There are three controls in cross-cultural experimentation (Brislin, 1980). These controls are related to location of stimulus, location of subjects, and location of reaction recording. Location of stimulus deals with issues such as the extent of stimuli comparability across cultures, attentiveness to the stimulus, and how aesthetic preferences affect the stimulus and its influence on the outcome. Questions related to location of subjects are: can some experimental situations be entirely artificial in some cultures and yet real in others? Are there special demand characteristics in some cultures not found in others? Do sequential effects make a difference in subjects’ responses? Finally, the location of reaction recording must be understood. An overt reaction or a word in one culture may have a quite different meaning in another culture.

We conclude this section on collecting primary data by stating that cross-cultural researchers would do well to employ multiple methods of data collection. Triandis and Brislin (1984) point out the need for multiple methods
of obtaining data. Mono-method bias can confound with the substantive issue. For example, using only one test to measure personality leads to results distorted by mono-method bias because the concept of self, and projection of the self in public, are subject to cultural inferences (Markus and Kitayama, 1991). In some cultures, people flatter and think highly of themselves while in others people are self-deprecating. A multimethod approach may involve several data collection methods (qualitative research, surveys, experiments, observation).

Measurement and scaling
Scientific observation and measurement are theory-laden; rationalism and empiricism interact to produce scientific explanation. Thus, researchers should critically examine the fit between the theoretical concepts they wish to study and the specific measures they employ. Whenever possible, the most direct measures should be used, using indirect measures as a last resort. In addition, several issues pertinent to equivalence, scale construction, and reliability and validity should be considered.

Equivalence. It is critical to establish the equivalence of scales and measures used to obtain data from different cultures. Construct equivalence deals with the question of whether the marketing constructs (e.g. variety seeking, brand loyalty) have the same meaning and significance in different cultures. In many countries, the number of brands available in a given product category is limited. In some countries the dominant brands have become generic labels symbolizing the entire product category. Consequently, a different perspective on brand loyalty may have to be adopted in these countries. As illustrated in Figure 2, construct equivalence requires an examination of functional, conceptual, instrument, and measurement equivalence (Drasgow and Kanfer, 1985).

Functional equivalence examines whether a given concept or behavior serves the same role or function in different cultures. For example, in many developing countries bicycles are predominantly a means of transport rather than recreation. Marketing research related to the use of bicycles in these countries must examine different motives, attitudes, behaviors, and even different competing products than such research would in the USA. Conceptual
equivalence deals with whether the concept or construct is expressed in similar attitudes or behaviours across cultures. For example, promotional sales are an integral component of marketing effort in the USA. On the other hand, in countries with shortage economies, where the market is dominated by sellers, consumers view sales with suspicion, because they suspect that the product being promoted is of poor quality.

Instrument equivalence deals with whether the scale items, response categories, and questionnaire stimuli such as brands, products, consumer behaviour, and marketing effort are interpreted identically across cultures. In the USA, the category of the principal shopper may be defined as either the male or female head of household. This category may be inappropriate in countries where routine daily shopping is done by a domestic servant. Furthermore, the category “household” itself varies across cultures.

Measurement equivalence examines whether each scale item measures the underlying construct equivalently in cross-cultural data. It consists of calibration equivalence, translational/linguistic equivalence, and scalar equivalence. Calibration equivalence examines whether the units of measurement are the same in different cultures. For example, if weight is measured in pounds in one culture and kilograms in another, then questionnaire items relating to weight should be carefully converted during translation. Translational/linguistic equivalence refers both to the spoken and the written language forms used in scales, questionnaires, and interviewing. The scales and other verbal stimuli should be translated so that they are readily understood by respondents in different cultures and have equivalent meaning.

Scalar equivalence, also called metric equivalence, examines whether the psychometric properties of data from the various cultures exhibit the same coherence or structure (Bhall and Lin, 1987). It examines whether the scores obtained from respondents in different cultures have the same meaning and interpretation. This involves demonstrating that two individuals from different cultures with the same value on some variable, such as brand loyalty, will score at the same level on the same test. The specific scale or scoring procedure used to establish the measure should be equivalent. The equivalence of response to a given measure in different cultures should be considered. For example, do the top-box or the top-two-boxes scores on a purchase-intent scale reflect similar likelihood of purchase in different cultures?

It has been suggested that the problem of equivalence in cross-cultural measurement can be addressed by assuming that, first, indicators of similar constructs are manifested in different ways in different cultures and, second, the influence of a third, or intervening, variable can be considered to account for moderating effects (Przeworski and Teune, 1970). Thus, cross-cultural measures should be composed of a set of cross-national etic indicators and a set of culture-specific emic indicators. A combination of the etic and emic indicators should result in a scale with improved reliability and validity in different cultures. The measurement is equivalent to the extent to which the
scale furnishes homogeneous indices for the various cultures; measures for specific cultures are equivalent to the extent to which the culture-specific emic measures are related to the identical etic measures.

To put the issue of equivalence in proper perspective, we refer to the paradox discussed by Sechrest et al. (1972). These authors note that important cultural differences may be obscured, even obliterated, by an attempt to achieve a rather misleading notion of equivalence. Thus, we should not be so obsessed by various types of equivalence that we preclude the cultural uniqueness of responses from surfacing.

Scale construction. In designing the scale or response format, respondents' educational or literacy levels should be taken into account. One approach is to develop scales which are pan-cultural, or free of cultural biases. Of the scaling techniques commonly used, the semantic differential scale may be said to be pan-cultural. It has been tested in a number of cultures and has generally produced similar results. However, Yu et al. (1993) found that attitude measures such as Likert and semantic differential scales are also culture-specific (with emic properties) even among countries with some commonality such as Japan, South Korea, and China. Significant main effects and interactive effects were observed owing to country and types of scale used on response ratings. Greater care should be used to minimize scale type effects across countries with substantial cultural differences (Western and Eastern countries). A similar problem could result from extreme response style (ERS) which is the tendency of a group to endorse extreme categories of responses in multiple response items. This can produce differences in group means and affect the level of item correlation within a measure (Samiee and Jeong, 1994). Courtesy bias is another source of variation common in Asia which raises concerns about biases that would affect comparability of results (Douglas and Craig, 1983).

When using a Likert-type scale or a semantic differential scale, international marketing researchers need to test the significance and appropriateness of anchors. For example, Johnson et al. (1993) in a study of US-Japanese channel relationships found after a preliminary questionnaire phase that Japanese managers did not adequately understand the scale anchors “agree/disagree”. As a result the anchors were changed to “definitely true”, “somewhat true”, and “not at all true”.

An alternative approach is to develop scales that use a self-defined cultural norm as a base referent. For example, respondents may be required to indicate their own anchor point and position relative to a culture-specific stimulus set. This approach is useful for measuring attitudes that are defined relative to cultural norms (e.g. attitude towards marital roles). In developing response formats, verbal rating scales appear to be the most suitable. Even less-educated respondents can readily understand and respond to verbal scales. Special attention should be devoted to determining equivalent verbal descriptors in different languages and cultures. The end points of the scale are particularly prone to different interpretations. In some cultures “1” may be interpreted as
best, while in others it may be interpreted as worst, regardless of how it is scaled. For measuring consumer perceptions and preferences in developing countries, the use of binary or dichotomous scales has been proposed (Malhotra, 1988a).

Reliability and validity. A critical review of the cross-cultural marketing research literature shows a lack of concern for reliability and validity issues (Aulakh and Kotabe, 1993; Parameswaran and Yaprak, 1987). This is not too surprising since only recently have these issues been accorded attention in domestic marketing research (e.g. Peter, 1979; 1981). However, these issues are crucial to making meaningful comparisons across cultures. It has been demonstrated that the same scales may have different reliabilities in different cultures. Davis et al. (1981) showed that two sources of measure unreliability (assessment method and nature of the construct) can confound the comparability of cross-cultural findings. Thus, substantive relationships among constructs must be adjusted for unequal reliabilities before valid inferences can be drawn. The adjustment factors to be applied are based on appropriate attenuation formulas (see Bollen, 1989; Nunnally, 1978). A practical approach to comparing the reliabilities of measures across cultures has been suggested by Parameswaran and Yaprak (1987). Essentially, this approach involves calculating the domain-composite reliability scores (e.g. Cronbach's alpha), splitting the sample in each culture into subsamples, and then using the various constructs as repeated measures in a mixed (between-subjects and within-subjects) analysis of variance design.

It has been argued that observations or measures of the phenomena in Culture B by a researcher in culture A are inherently ambiguous. The observations or measures may be a function of the real phenomena in culture B or a function of the observer bias derived from culture A (Berry, 1980; Brewer and Campbell, 1976; Campbell, 1970). Ideally, a research project should be conducted four times: twice in culture B (once with an observer from culture A and once with one from culture B), and twice in culture A (with the same set of observers). Only then can we determine the variation in the data which are due to the “real phenomena” and those due to “observer bias”. At the level of individual measurement, this strategy is well known in psychology and marketing research as the multitrait multimeasure approach (Malhotra, 1987a). However, the use of this approach in cross-cultural research has been limited.

Irwin et al. (1977) have proposed three types of validity based on the imposed etic, emic, and derived etic distinction discussed earlier in this paper. Support for imposed etic validity can be demonstrated by correctly predicting an outcome in culture B based on a theory or construct imposed from culture A. To show emic validity, the researcher must correctly predict an outcome in culture B based on theory or construct derived from culture B, i.e. that culture's own conceptual system. Derived etic validity can be demonstrated only after imposed etic and emic validity have been established. The derived etic validity must be based on the known validity in two or more cultures. According to
Davidson et al. (1976), to demonstrate derived etic validity, the researcher must generate emic content of the etic construct. Such content can then be employed to develop appropriate measures. These measures would be expected to have derived etic validity since they would have been developed on the basis of both etic and emic analyses.

**Questionnaire design**

The questionnaire or research instrument should be adapted to the specific cultural environment and should not be biased in terms of any one culture. This requires careful attention to each step of the questionnaire design process. It is important to take into account any differences in underlying consumer behaviour, decision-making processes, psychographics, lifestyles, and demographic variables. In the context of demographic characteristics, information on marital status, education, household size, occupation, income, and dwelling unit may have to be specified differently for different countries, as these variables may not be directly comparable across cultures. For example, household definition and size varies greatly, given the extended family structure in some countries and the practice of two or even three families living under the same roof.

While personal interviewing is the dominant survey method in cross-cultural marketing research, different interviewing methods may be used in different countries. Hence, the questionnaire may have to be suitable for administration by more than one method. For ease of comprehension and translation, it is desirable to have two or more simple questions rather than a single complex question. In overcoming the inability to answer, the variability in the extent to which respondents in different cultures are informed about the subject matter of the survey should be taken into account. Respondents in the Far East and the former Soviet Union may not be as well informed as those in the USA.

The use of unstructured or open-ended questions may be desirable if the researcher lacks knowledge about the determinants of response in other cultures. Unstructured questions also reduce cultural bias, because they do not impose any response alternatives. However, unstructured questions are more affected by differences in educational levels than structured questions. They should be used with caution in countries with high illiteracy rates.

The questions may have to be translated for administration in different cultures. A set of guidelines has been proposed by Brislin et al. (1973) for writing questionnaires in English so that they can be easily translated. These include:

- use short and simple sentences;
- active voice;
- repeat nouns rather than using pronouns;
- avoid metaphors;
- avoid adverbs and prepositions related to place and time;
• avoid possessive forms;
• use specific rather than general terms;
• avoid vague words;
• and avoid sentences with two different verbs if the verbs suggest different actions.

In addition, researchers should provide redundancy and add context for any difficult phrases. Redundancy helps since two phrases referring to the same concept allow the translator to be sure of the meaning. Likewise, it helps the translator-checker to identify errors when one of the phrases is different from the other. The principle of adding context states that translation improves when the word is part of a sentence, and is even better when the sentence is part of a paragraph.

Direct translation, in which a bilingual translator translates the questionnaire directly from a base language to the respondent's language, is frequently used. However, if the translator is not fluent in both languages and familiar with both cultures, direct translation of certain words and phrases may be erroneous. Procedures like back-translation and parallel translation have been suggested to avoid these errors. In back-translation, the questionnaire is translated from the base language by a bilingual speaker whose native language is the language into which the questionnaire is being translated. This version is then retranslated back into the original language by a bilingual whose native language is the initial or base language. Translation errors can then be identified. Several repeat translations and back-translations may be necessary to develop equivalent questionnaires, and this process can be expensive and time-consuming. The challenge of back-translation increases with an increase in the number of cultures being examined. An alternative procedure is parallel translation. A committee of translators, each of whom is fluent in at least two of the languages in which the questionnaire will be administered, discusses alternative versions of the questionnaire and makes modifications until consensus is reached. In countries where several languages are spoken, the questionnaire should be translated into the language of each respondent subgroup. It is important that any non-verbal stimuli (pictures and advertisements) are also translated using similar procedures.

It is important to demonstrate the equivalence of translated questionnaires to the original language version. Otherwise, the cross-cultural differences due to non-equivalent questionnaires might confound the results. For example, Grunert and Scherhorn (1990) found differences between the American and German samples when administering the List of Values (LOV). However, imprecise translation was more responsible for these results than actual differences in value orientations. Gentry et al. (1995) found that, although the scales developed in North America worked well across the Chinese-Thai and the Muslim-Thai subcultures, the use of a particular language was a complicating factor in their study of acculturation for ethnic groups in
The availability of three languages to acculturating people (such as immigrants) in Thailand made it different from the acculturating people in the USA, where only two languages are expected (English and the language of origin). They, therefore, caution that although language usage is a widely used measure of bi-directional acculturation in the USA, it should be used with care in countries that are multilingual.

There is also a need for instrument validation and equivalence. Jaffe and Nebenzahl (1984) demonstrated that alternative questionnaire formats used in country image studies lacked reliability and validity. Using five countries and 13 descriptors, the authors found that the two questionnaire formats were not equivalent. Country image studies that have different questionnaire formats may not be comparable, even if the scale items are the same.

Pretesting of the questionnaire is complicated in cross-cultural research, because the linguistic equivalence must be pretested. Two sets of pretests are recommended. The translated questionnaire should be pretested on monolingual subjects in their native language. The original and translated versions should also be administered to bilingual subjects. The pretest data from administration of the questionnaire in different countries or cultures should be analysed and the pattern of responses compared to detect any cultural biases.

Sampling

In cross-cultural psychology, sampling issues deal with the selection of cultures, individuals, stimuli, and responses. First, the cultural unit of analysis should be carefully defined. For example, the common practice of using a nation-state as a surrogate for culture may be inappropriate for countries with heterogeneous cultures (e.g., India, Malaysia, etc.). One suggestion is to use the cultunit as the unit of analysis. The cultunit is defined as people who are domestic speakers of a common distinct language and who belong either to the same state or the same contact group. Thus, the cultunit incorporates three criteria:

1. language;
2. territorial contiguity and, whenever there is sufficient authoritative political structure;
3. political organization (Naroll, 1970).

Alternatively, behaviour settings may be used as units for cross-cultural comparisons. A behaviour setting represents all the forces acting on individual members of a setting to enter and participate in its operation in particular ways. The criteria for identifying or defining behaviour settings may be found in Barker (1971). The selection of cultures (cultunits or behaviour settings) to be investigated should be based on the theoretical or applied objectives of the study. The latter point is noteworthy as the selection of cultures in cross-cultural research has been based primarily on convenience rather than on theoretical considerations. It is also important to select multiple cultures, as
studies which investigate only one or two cultures have limited usefulness (Samiee and Jeong, 1994; Sekaran, 1983).

The selection of individuals (or other sampling units) within a culture is also a salient issue. Sampling frames, often constructed to select individuals, should be equivalent across cultures. In many ways, the issues involved in sampling individuals in a cross-cultural setting are similar to those in domestic marketing research. Sampling of stimuli is also less of an issue since the stimuli are generally selected on the basis of the marketing research problem. However, if standard scales or tests are being administered, then the test items from which the total trait is to be measured must be representative and valid samples of the ability in question, as it is relevant within that particular culture (Goodenough, 1936). Similar arguments apply to the sampling of responses.

Implementing the sampling design process in cross-cultural marketing research is seldom an easy task. Several factors should be considered in defining the target population. The relevant sampling element (respondent) may differ from country to country. In the USA, children play an important role in the purchase of children's cereals. However, in cultures with authoritarian child-rearing practices, the mother may be the relevant element. Women play a key role in the purchase of automobiles and other durables in the USA; in male-dominated societies, such as in the Middle East, such decisions are made by men. Accessibility to respondents also varies across countries. In Mexico, many houses cannot be entered by strangers (i.e. interviewers) because of boundary walls and servants. Additional problems may be encountered. Dwelling units may be unnumbered and streets unidentified, making it difficult to locate designated households.

Developing an appropriate sampling frame is a difficult task. In many cultures, particularly in developing countries, reliable information about the target population may not be available from secondary sources. Government data may be unavailable or highly biased. Population lists may not be available commercially. The time and money required to compile these lists may be prohibitive. For example, in Saudi Arabia, there is no officially recognized census of population, no elections, and hence no voter registration records, and no accurate maps of population centres. If a sampling frame is not available, variations of the random walk plan can be adopted (Scheuch, 1968). For example, if villagers are to be interviewed, one could examine the routes on which the village residents walk. Then, houses from along the walk routes could be sampled. Another variation, which also works well in urban areas, is to select different starting points and identify sampling units according to some pattern (three dwelling units east, then five dwelling units north).

Given the lack of suitable sampling frames, the inaccessibility of certain respondents, such as women in some cultures, and the dominance of personal interviewing, probability sampling techniques, although more appropriate, are uncommon in cross-cultural marketing research. Quota sampling has been used widely in the developed and developing countries both in consumer and in industrial surveys. However, great caution should be exercised in using quota
Cross-cultural marketing research

Sampling to control for respondent selection bias. For example, only the most accessible individuals of a given type may be interviewed to fill the quota, although such individuals are not representative of the type. Snowball sampling is also appealing when the characteristic of interest is rare in the target population or when respondents are difficult to reach. For example, it has been suggested that, in Saudi Arabia, graduate students should be employed to hand-deliver questionnaires to relatives and friends. These initial respondents can be asked for referrals to other potential respondents, and so on. This approach would result in a large sample size and a high response rate (Tuncalp, 1988).

Statistical estimation of sample size may be difficult, as estimates of the population variance may be unavailable or may differ from country to country. Hence, the sample size is often determined by qualitative considerations such as:

- the importance of the decision;
- the nature of the research;
- the number of variables;
- the nature of the analysis;
- sample sizes used in similar studies;
- incidence rates;
- completion rates; and
- resource constraints (Malhotra, 1996).

It is important to recognize that the sample size may vary across cultures. Homogeneous cultures may require smaller samples while heterogeneous cultures require larger sample sizes.

Sampling techniques and procedures vary in accuracy, reliability, and cost from country to country. To achieve comparability in sample composition and representativeness, it may be desirable to use different sampling techniques in different cultures. In a study comparing perceived risk and brand loyalty in the four countries, Verhage et al. (1990) used a telephone directory-based random sample for the USA, judgemental sample followed by a random selection of households in Mexico, a judgemental sample followed by a store-intercept technique in urban centres in Thailand, and convenience and snowball sampling in Saudi Arabia.

Sometimes samples are drawn from an identifiable subgroup of the cultures’ populations, e.g. middle-class housewives. Even in such a case, sample comparability cannot be assumed but must be achieved by conscious effort on the part of the researcher. It has been suggested that an attempt should be made to hold age, sex, social class, and urban-rural residence constant across samples (Green and White, 1976). Sampling equivalence can be further refined by accounting for cultural and social systems within countries. When dealing with multiple countries with different levels of homogeneity or heterogeneity, the
researcher must account for such differences. Takada and Jain (1991) in their study of the diffusion process (using the Bass model) of consumer durable goods in the Pacific Rim countries found that the rate of adoption in countries characterized by high context culture and homophilous communication (such as in Japan, South Korea and Taiwan) was faster than in countries characterized by low context culture and heterophilous communication (such as in the USA). Understanding the interaction of the marketing efforts with culture can determine the degree of success in a foreign country.

In marketing studies involving organizational research and interfirm relationships, a suggested strategy has been to use a second informant who is both knowledgeable and willing to respond (Phillips, 1981). Johnson et al. (1993) used two informants (at the same level) to collect the same data in the study of US-Japanese channel relationships. This provided a way to assess the degree of response equivalence. Deshpande et al. (1993) extended the idea of dyads and used two pairs of dyads or a “quadrat” (two informants each for both the buyer and seller) as the unit of analysis. This concept of a sampling unit becomes more meaningful in international marketing where organizational managers are from diverse backgrounds (ethnic and otherwise) working in an overseas subsidiary.

An attempt should be made to minimize the total error, rather than just the sampling errors. In cross-cultural research, non-sampling errors which are due to lack of comparability of the questionnaires or measurements and variations in interviewer quality, can be particularly large. It is possible to reduce the magnitude of the total error by trading off a measured amount of sampling variance for better control of non-sampling error.

**Field work**

The quality of field work in cross-cultural research is affected by several factors including interviewer background, the interview and its setting, respondent background, and the cultural background (Brislin et al., 1973; Pareek and Rao, 1980). Interviewer background includes factors such as interviewer affiliation, interviewer image, respondent-interviewer distance, respondent relevance, and interviewer bias (e.g. owing to inadequate training). In a multicultural study, the responses may not be comparable if the image of the sponsoring agency, the interviewer, or the respondent-interviewer distance is perceived non-comparably in different cultures. If the respondents are sensitive to certain aspects of the interviewer background, care should be taken to bridge that distance. For example, in some cultures, it may be necessary to use female interviewers to interview women on matters of sex and family planning.

Important aspects of the interview and its setting include the thematic relevance and sensitivity, cultural relevance, social desirability, capacity to reach depth, length, and structure. In some cultures, but not in others, the presence of a third person may be a source of bias. The relevance of the topic, and hence the need to emphasize the relevance of the interview through initial
rapport building, may vary across cultures. In certain cultures some questions may be sensitive or taboo, and thus require special interviewer training. To achieve response equivalence, variance due to needless differences in data collection procedures should be minimized. Data from the different cultures should be collected within acceptable time frames.

The respondent-related factors of concern are private-public opinion gap, omniscience syndrome, previous experience, saturation, and response sets. The previous experience of the respondent may influence attitude towards the interview including the willingness to participate. In some cultures, respondents may be so accustomed to interviewing that they may answer in a superficial, routine way. Response-set biases, such as extreme checking style, and acquiescence response set, may vary across cultures. The field workers should be adequately trained to deal with such biases.

Important cultural factors relate to courtesy norms, reticence, and game playing. Courtesy norms, such as greeting respondents or accepting hospitality from them, are likely to vary across cultures. In some cultures, such as Japanese or the Chinese residing in Malaysia, people do not talk much or are slow in giving their responses. Game-playing norms such as leading strangers (interviewers) astray or lying to them, also vary. The field workers should be sensitized to such differences and trained accordingly.

The selection, training, supervision, and evaluation of field workers is critical in cross-cultural marketing research. Local field work agencies are unavailable in many countries. Therefore, it may be necessary to recruit and train local field workers or import trained foreign workers. The use of local field workers is desirable, as they are familiar with the local language and culture. They can thus create an appropriate climate for the interview and be sensitive to the concerns of the respondents. As far as possible, the interviewers selected and the respondents should have similar backgrounds. When interviewers and respondents of known dissimilar backgrounds are used, it is desirable to study the effect of this variable on response. For example, Webster (1996) found significant interaction effects of respondent ethnicity with interviewer ethnicity with respect to response rates and item response effort rates. An interview with an ethically similar interviewer and respondent produces higher response quality than an interview with an ethically dissimilar interviewer and respondent.

Interviewer biases are often due to communication problems between the interviewer and respondents. Several biases have been identified in cross-cultural research including rudeness bias, "I-can-answer-any-question" bias, courtesy bias, sucker bias, hidden premises bias, reticence-loquaciousness bias, social desirability, status difference bias, racial difference bias, and individual-group opinion bias. The reader is referred to Brislin et al. (1973) for guidelines for dealing with these biases.

Extensive training and close supervision of the interviewers and other field staff may be required to minimize these biases. Finally, interviewer cheating may be more of a problem in many foreign countries than in the USA.
Validation of field work is critical. In some cultures interviewers view validation procedures as insulting and expressive of distrust. However, this problem can be addressed by making explicit from the beginning the necessary and routine character of procedures adopted to verify individual work. Moreover, validation should be carried concurrently with interviewing so that it is not perceived by the interviewers as an empty threat which will not be carried out.

**Data preparation and analysis**

The main objective in cross-cultural research is identifying and measuring differences and/or similarities from various samples (Samiee and Jeong, 1994). The internal validity of international marketing research improves as fewer rival explanations for the results of a study remain. Several data analysis issues are pertinent including data preparation, standardization, sample comparability, construct equivalence, the level of analysis, and avoiding common methodological fallacies.

**Data preparation**

In preparing the data for analysis, unusually influential responses in the sample should be identified as they can be a problem. Outliers – observations with values that are distant from the bulk of the data – can distort results. Outliers are not necessarily a problem, but could indicate a problem in sampling (Mullen et al., 1995). There are four sources for the presence of outliers (Tabachnick and Fidell, 1989). First, data entry error could have occurred. Second, a missing value indicator could have not been specified, so the computer program reads the value as a real one. Third, and importantly for the cross-cultural marketing researcher, the outlier is not a member of the intended population. For example, a French expatriate may be included in a sample from a West African culture solely because he or she lived in the cultural community of interest in West Africa. Finally, the outlier could be from the intended population, but an unusual member of the population (i.e. an affluent, British-educated citizen of a developing country).

The decision to retain or delete outliers from the analysis is often difficult. Deletion of cases is a step of last resort, and must be described in the study report. Both the impact of outliers and the purpose of the research project must be considered. Transforming variables is one option to deleting cases. For groups of cases, discounted weighting could be another option.

While graphical plotting of distributions can be useful for identifying cases with extreme values on one variable, multivariate outliers (cases with unusual combination of two or more scores) can be extremely difficult to identify in this way. One statistical approach to identifying multivariate outliers includes the use of the Mahalanobis distance ($D^2$) which gauges the distance of cases from the centroid of the remaining cases while taking into account the covariance of the variables in question. If the set of variables has a high degree of covariation, the $D^2$ value will be adjusted downward. Conversely, if the set of variables is very distinct and has a small covariation, the $D^2$ value will be adjusted upward.
With multivariate normal data, the $D^2$ values follow a chi-square distribution with $p$ degrees of freedom ($p =$ number of variables). $D^2$ values which are significant at the 0.05 level indicate outliers, while values significant at the 0.01 level indicate extreme outliers.

Mullen et al. (1995) used the Mahalanobis distance measure along with two other statistical techniques – Bollen's $a_{ii}$ and their own Observed Covariance Ratio to identify multivariate outliers convergently. These authors integrated a multiple methods statistical approach to their macro-level study comparing the values for 96 countries on 12 economic and human welfare variables. Five countries including the Central African Republic, Libya, Nepal, Saudi Arabia, and Iraq were identified as outliers. In the second phase of this analysis, structural equation modelling with latent variables was used to assess the change in parameter estimates and overall model fit when outliers were included or dropped for the analysis. This tandem approach of using statistical tests and statistical models gives the researcher deeper understanding for the data being analysed.

Should the data be standardized?
Several issues should be considered in deciding whether the data should be standardized. Some researchers prefer to standardize the variables within each culture based on arguments of interpretability, common metric, or emic comparison (Singh, 1995). They argue that statistics based on standardized coefficients are easier to interpret. Furthermore, standardization converts the variables measured on scales to a common metric. This permits comparative analysis of the effects of different independent variables. For example, in regression analysis, under certain conditions it is meaningful to use the absolute values of the standardized regression coefficients, or the squared values, as measures of relative importance of the predictor variables (Malhotra, 1996). As the data for each culture are standardized separately, analyses based on such data reflect an emic comparison standard. Since the regression coefficients are adjusted on the basis of within-sample variability, the standardized coefficients have the same metric within a culture, but not across cultures. Standardized estimates eliminate any across-culture differences on account of differences in variances.

In contrast, statistics based on unstandardized data are preferred on the basis of arguments of comparability across cultures, structural invariance, and an etic comparison standard. Valid comparisons of statistics across cultures can be conducted only if these statistics are based on unstandardized data. Structural invariance means that statistics (e.g. unstandardized regression coefficients) represent structural parameters that are invariant across different samples obtained from the same culture. Structural invariance is much more likely to hold for statistics based on unstandardized data than on statistics based on standardized data (Alwin, 1988; Bollen, 1989; Singh, 1995). Statistics based on unstandardized data reflect an etic comparison standard because they are unadjusted for within-sample variability. However, it should be noted that
etic comparisons across cultures based on unstandardized data assume construct equivalence.

Thus, we recommend that general etic comparisons across cultures should be made on the basis of unstandardized data, assuming that construct equivalence has been achieved. However, emic comparisons within a culture should be made on the basis of standardized data. As opposed to standardizing the variables across the sample within each culture, sometimes to achieve scalar equivalence, it may be desirable to standardize the data by respondent. Such a procedure results in relative statistics (e.g., relative mean) for variables for each culture. These standardized statistics can be used for relative, rather than absolute, comparisons of the variables. See Kotabe et al. (1991) for an application of this approach. In the special case when the interest is to identify the etic dimensions of individual variation, a sequential standardization procedure discussed later is recommended.

Sample comparability
It is important that the samples in the cultures being studied should be equivalent in terms of basic socioeconomic, organizational, and other salient characteristics that may affect the findings. Such checks should always be made. In a study of strategy principles in Japan, Kotabe et al. (1991) performed three layers of analysis to examine the representativeness of the sample. First, the manufacturing sector firms were examined to verify the match with the industrial composition from which the samples were taken. Second, several demographic variables were measured to assess if any response pattern bias may have occurred across samples. Finally, the data were examined for any systematic differences in response patterns between early and late respondents.

In case the samples are not comparable, the analyses should be conducted on various subsamples to assess the impact of differences in terms of the identified variables. Sekaran and Martin (1982) illustrate this approach. Alternatively, it may be possible to control statistically for the effect of these variables. For example, partial correlations may be computed on the model items holding the background variables constant before further statistical testing of the models. Lee and Green (1990) provide an application of this approach. These authors used a series of ANOVAS with variables of the Fishbein model as the dependent variables and demographic variables in each culture as the grouping variables. Sex, status in society, and family income were found to be related to some of the variables in the Fishbein model. The effects of these demographic variables were partialled out of all the Fishbein model variables. The resulting partial correlation matrix was then converted to a covariance matrix and submitted to LISREL analysis. As another example, statistical control may be exercised by using an analysis of covariance framework.

Equivalence
The usefulness of employing common factor analysis and structural equations with latent variables in assessing construct equivalence is highlighted in the case of consumer ethnocentrism with the CETSCALE (Shimp and Sharma,
Netemeyer et al. (1991) collected data from the USA, France, Japan, and Germany using the CETSCALE. These researchers then employed multiple sample analysis (Jöreskog and Sörbom, 1989) to test for construct equivalence across these four culture groups.

Multiple group LISREL is a general technique for exploring the measurement model equivalence in multiple groups simultaneously (Bollen, 1989). This procedure assumes that independent, random samples are available from each population (culture). In structural equation terms, the measurement models of a latent variable for multiple groups are defined by the following parameter matrices:

\[ \Lambda^g \] (factor loading matrix)
\[ \theta^g \] (measurement error)
\[ \Phi^g \] (covariance matrix)

where, \( g = 1, 2, 3, \ldots, G \) is the number of groups (cultures).

For the measurement properties to be the same in all groups,

\[ \Lambda^1_x = \Lambda^2_x = \ldots = \Lambda^G_x \]
\[ \theta^1_0 = \theta^2_0 = \ldots = \theta^G_0 \]

The overall hypothesis for measurement equivalence is that the measurement models are invariant across groups. Bollen (1989) recommends a hierarchy of tests starting with the least demanding test of testing invariance of model form. This tests the equality of the number of factors but not the equality of constraints across parameters. If the model fit is poor, no further tests are done. If the fit is good, then the invariances of coefficients (factor loadings) are tested across groups. Finally, in a nested sequence, the measurement errors and covariances are tested across groups.

Singh (1995) notes that many researchers incorrectly regard factor equivalence as a sufficient condition for construct equivalence. Mullen (1995) elaborated on the distinctiveness of these aspects of construct equivalence by explaining that factor equivalence helps to diagnose translation equivalence in cross-national research. The more demanding test of error variance across groups is a joint test of systematic variance and random variance. Mullen asserts that the test of equivalent random errors addresses inconsistent scoring, while the test of systematic error variance addresses the validity issue of scalar equivalence.

Although Mullen (1995) reports a US-Japanese study where the third aspect of construct equivalence for a job satisfaction construct emerged, such results are rare at this stage of development for cross-cultural marketing research. This discussion of equivalence thus returns to the broader issue of internal validity presented earlier. Bollen proposed a definition of validity based on a structural equations approach. This approach incorporates the idea of an underlying
construct where the validity of a measure of $x_j$ for $\xi_j$ is the magnitude of the direct structural relation between $x_j$ and $\xi_j$ (Bollen, 1989, p. 197).

**Level of analysis**

The data analysis could be conducted at three levels:

1. **individual**;
2. **within-country or cultural unit**; and
3. **across-countries or cultural units**.

Individual level analysis requires that the data from each respondent must be analysed separately. For example, one might compute a correlation coefficient or run a regression analysis for each respondent. This means that enough data must be obtained from each individual to allow analysis at the individual level, which is often not feasible (Malhotra, 1986; 1987b). Yet it has been argued that in international marketing or cross-cultural research the researcher should possess a sound knowledge of the consumer in each culture. This can best be accomplished by individual-level analysis (Tan et al., 1987). A useful approach for obtaining and analysing data at the individual level is conjoint analysis (Baalbaki and Malhotra, 1995; Diamantopoulos et al., 1995). Other means for identifying latent segments would be traditional clustering (Malhotra et al., 1992; Yavas et al., 1992), and latent class analysis when behaviourally-based segments are obtained from discrete choice experiments. In these discrete choice designs, consumers make repeated choices from alternative configurations of a product (Swait, 1994). Usually the multinomial logit model is used, but choice experiments can also be run using conjoint analysis (Malhotra, 1984; 1988b).

In within-country or cultural unit analysis, the data are analysed separately for each country or cultural unit. This is also referred to as intracultural analysis. This level of analysis is quite similar to that conducted in domestic marketing research. The objective is to gain an understanding of the relationships and patterns existing in each country or cultural unit. In across-countries analysis, the data of all the countries are analysed simultaneously. Two approaches to this method are possible. The data for all respondents from all the countries can be pooled and analysed. This is referred to as pan-cultural analysis. Alternatively, the data can be aggregated for each country and these aggregate statistics analysed. For example, one could compute means of variables for each country, and then compute correlations on these means. This is referred to as cross-cultural analysis. The objective of this level of analysis is to assess the comparability of findings from one country to another (Netemeyer et al., 1991). The similarities as well as the differences between countries should be investigated. When examining differences, not only differences in means but also differences in variance and distribution should be assessed. All the commonly used statistical techniques can be applied to within-country or
cross-country analysis and, subject to the amount of data available, to individual-level analysis as well.

Methodological fallacies

Researchers whose focus is on understanding culture per se are keenly interested in the dimensions of culture. Hofstede's classic work (1980) comparing 40 countries identified four dimensions which explained half of the country-to-country differences. These dimensions were:

1. power distance;
2. uncertainty avoidance;
3. individualism and collectivism; and
4. masculinity versus femininity.

Eventually, a fifth dimension was added – long-term versus short-term orientation.

Hofstede (1994) warns about the different findings which researchers might find at the culture level and at the individual level. Citing individualism and collectivism as an example, Hofstede has made it clear that this dimension was proposed to describe societal contexts – not individuals. What characterizes a culture does not necessarily characterize individuals in the culture. As support for this assertion, Hofstede notes that correlations between variables at the culture level (based on mean scores of the individuals within the culture) can be completely different from the correlations of the same variables at the individual level. One set of variables may produce a bipolar dimension at the culture level and two or more unipolar dimensions at the level of individuals.

An ecological fallacy occurs when a researcher uses a culture-level correlation (GNP, epidemiological rates) without conducting individual-level analysis to interpret individual behaviour, while a reverse ecological fallacy occurs when researchers construct cultural indices based on individual-level measurements (attitudes, values, behaviours) without conducting culture-level analysis (Schwartz, 1994).

When focusing on cultural variation (e.g. the dimensions of culture), intra-cultural analysis (studying subjects in different populations separately), cross-cultural analysis (comparing the results of separate intra-cultural analyses), and pan-cultural analysis (mixing subjects from separate cultures to study) are appropriate. Counter-intuitively, pan-cultural analysis is not a suitable way for identifying universal dimensions of individual variation (Leung and Bond, 1989). To find dimensions of individual variation across which persons can be compared regardless of cultural background, Leung and Bond's individual-level analysis must be conducted. To accomplish this, two sequential standardizations are done before the data are aggregated into a pan-cultural data set and factor analysed. First, a within-subject standardization is done (case standardization) to remove response-style effects. Second, a within-culture
standardization is accomplished to set the mean on each item to zero and its standard deviation to one (Leung and Bond, 1989).

Because the mean value for each item within each cultural group is now zero, the positioning effect (differences in item means) of culture is removed. However, the patterning effect (correlations between items) of culture remains in the data set. In subsequent factor analysis, if a healthy factor structure results, this provides evidence for an etic or universal dimension of culture. If a poor factor structure results, then the dimension in question would likely be an emic or culture-specific construct (Leung and Bond, 1989; McCort, 1992).

It should be noted that two types of etic constructs exist. One, is termed a weak etic, which does not overlap with culture (i.e. price sensitivity). The other is termed a strong etic (i.e. individualism) because cultures, as well as individuals appear to manifest this construct. Etic phenomena usually have the stronger interest for cross-cultural marketing researchers, compared with emic phenomena, because of the multiple-culture focus of cross-cultural marketing researchers.

Report preparation and presentation
Interpreting and reporting data pose special problems in cross-cultural research. If only researcher(s) from one culture interpret the data gathered in two (or more) cultures, the researcher’s own biases may affect the implications derived. Values found in one culture may not be universally understood. Researchers themselves cause another problem based on their ethnocentrism. They report data based on their own frames of reference in journals from their country. To overcome such problems, it has been suggested that researchers from each country should interpret independently the data so that estimates of inter-interpreter reliability can be made (Brislin et al., 1973).

Interpretation should take into account the limitations of the study, particularly those of data collection. Even if diverse cultures have been examined, projecting the results to other cultures not included in the study may not be justified. Certain countries like India, Nigeria, and Thailand, whose governments have regularly admitted foreign researchers, have been overused. Other methodological problems may also plague interpretation. One serious limitation has been differences in non-completion rates. In one five-nation study, the non-completion rates varied from 17 to 41 per cent (Almond and Verba, 1963). Proper interpretation requires information on the type of people available and not available in different cultures. Perhaps a sample of those not interviewed after the required number of call-backs have been made should be the subject of a special study. The timing of data collection from culture to culture is also important and should be comparable (Scheuch, 1968). Interpretation is easier if consistency checks have been built into the research design to allow response validation.

The guidelines for preparing domestic reports also apply to cross-cultural marketing research, although report preparation may be complicated by the need to present reports for management in different countries and in different
languages. In such a case, the researcher should prepare different versions of the report, each geared to specific readers. The different reports should be comparable, although the formats may differ. The guidelines for oral presentation are also similar to those domestic presentations with the added proviso that the presenter should be sensitive to cultural norms. For example, making jokes, which is frequently done in the USA, is not appropriate in some cultures.

**Summary and conclusion**

It is clear from the issues we have discussed that cross-cultural marketing research cannot be considered merely as an extension of domestic research. In a cross-cultural setting the researcher is faced with several complex methodological issues in each phase of the marketing research process. Many of these issues are not salient in a purely domestic setting. Failure to address these issues adequately in cross-cultural research can lead to confounding, alternative explanations, and severely limit the usefulness of the marketing research project. However, new analytical techniques such as structural equation modelling with latent variables offer promise. Improvements in statistically identifying outliers, and more rigorous means of assessing the equivalence of factors between population groups are some of the capabilities of structural equation modelling. Other techniques such as conjoint analysis provide an effective means of conducting individual and group-level analyses in order to identify emic and etic constructs. Such methodological advances must continue if we are to overcome the many challenges which still remain.

However, methodological advances must be rooted in well-grounded theory in order to advance the discipline of cross-cultural marketing research. Because culture can be conceived as the context in which the social aspects of living occur for humans, better understanding of the dimensions of culture and their influences on consumer behaviour must be gained to formulate useful theories of cross-cultural marketing. Environmental adaptation holds promise as a valuable focus for future theoretical development. Cross-cultural marketing researchers should study the way in which adaptation to different biological and ecological environments results in the development of particular habits of perceptual inference, preference formation, cognitive processes, and psychological skills, which are thought to be adaptive for these environments. Such an effort would require the developments of new frameworks, similar in spirit to the work of Dawson (1969), which integrate findings from cultural anthropology with cross-cultural psychology. We end with a call for continued methodological advances and theory development to advance cross-cultural marketing research.

**Note**

1. While most of the COO studies involve domestic rather than cross-cultural samples, the methodological problems they present apply even more forcefully to cross-cultural research. As noted earlier, all such studies are being subsumed under the rubric of cross-cultural marketing research.
References


