Comparability in Cross-Cultural Qualitative Marketing Research: Equivalence in Personal Interviews

Pia Polsa
Swedish School of Economics and Business Administration, HANKEN

Acknowledgements: I would like to thank Dr. Ann-Charlotte Lindeberg for not only editing the text but also providing ideas concerning illustrative examples of differences in small talk across cultures. Furthermore, I am grateful for the anonymous reviewers and the editor for providing helpful comments. Department of Marketing and Corporate Geography, P.O.Box 479, 00101 Helsinki, Finland, tel. +358-40- 3521341, mobil +358-50-538 0842, fax. +358-9-431 33287, email: pia.polsa@hanken.fi

EXECUTIVE SUMMARY

Increasingly qualitative inquiry has became of interest in marketing research, and a number of cross-cultural comparisons of qualitative data have been published. Therefore, it is important to address the issue of comparability in qualitative data. When assessing the comparability of two or more data sets across cultures, in quantitative research different types of equivalences have been widely examined, but this has not been common in qualitative inquiry. The present paper focuses on how the concept of equivalence can be applied in qualitative cross-cultural research. The comparability criteria found to be relevant in this context are: conceptual, functional, sample, contextual, researcher, and response equivalences. Thus, the paper contributes to qualitative research methodology in marketing by proposing the criterion of cross-cultural comparability in qualitative research methods in marketing.

Keywords: comparability, cross-cultural, equivalence, qualitative marketing research
Comparability in Cross-Cultural Qualitative Marketing Research: 
Equivalence in Personal Interviews

Traditionally, cross-cultural research has paid a great deal of attention to the comparability of data and findings across cultures, nations, regions, subcultures, groups and time. A widely used method to strengthen comparability has been the rise of different types of equivalences. In cross-cultural marketing research, research equivalence is an important concern, because it enables arguing that similarities and differences in results are caused by actual differences in two cultures, and not by invariance in the methods (cf. Baumgartner and Steenkamp, 2001; Kumar, 2000; Brislin, 1993; Craig and Douglas, 2005; Hui and Triandis, 1985; Mullen, 1995; Sekaran, 1983; Steenkamp and Baumgartner, 1998; Usunier, 1998, Usunier and Lee, 2005). As Harkness, Mohler and Van de Vijer point out: “Cross-cultural survey research is required to pursue strategies that try to come to terms with the fact that concepts may not be identical or comparable and that an instrument appropriate and adequate in one context (temporal or spatial) may not be adequate in another (Harkness, et al. 2003, p. 8).”

However, with some exceptions (Sparks, 2002; Malhotra Agarwal and Peterson, 1996) equivalences have mainly been studied in the context of quantitative marketing research (e.g. Davis, Douglas and Silk 1981; Mullen, 1995, Singh, 1995). At the same time, several cross-cultural researchers agree that quantitative methods are not applicable in all cultures or that they can be biased (e.g. Arnould, 2001; England and Harpaz 1983). At times, qualitative methods¹ are necessary because of the types of research questions asked, and because a qualitative approach uncovers emic issues of society more easily. Although there is a great deal of cross-cultural qualitative marketing research and a great need for such studies, with exception of Kjeldgaard, Csaba & Ger (2006) there appears to be no discussion or use of the concept of equivalences in qualitative studies of cross-cultural marketing.

As qualitative research methods have emerged in marketing (Douglas and Craig, 2006), studies using cross-cultural qualitative methods have begun to appear in international marketing journals (see for example Knight, Holdsworth and Mather, 2006, Owusu and Welch, 2007). Further, as qualitative inquiry is considered as crucial in comparative research (Malhotra et al. 1996), the examination of equivalencies in qualitative data is likely to increase the rigour of analysis in cross-cultural qualitative marketing research. This paper explores the possibility of applying the concept of equivalence in cross-cultural qualitative marketing research, in order to initiate a discussion about the applicability of equivalence of methods to qualitative research. In addition, the paper contributes to qualitative methodology by including the criterion of cross-cultural comparability in qualitative research methods in marketing. While quality criteria of qualitative research have been examined for example by Lincoln & Guba (2000 &1985) and in consumer behaviour research by Wallendorf & Belk (1989), these studies have not examined the quality of comparability in qualitative data, which is the specific focus of the current paper. The present paper is limited to covering literature mainly in the field of marketing. Therefore, the literature on methods in these fields as well as the illustrative examples of equivalences in qualitative cross-cultural research focus on personal interviews, a major tool in marketing research.

The next section offers a definition of the construct of equivalence and presents a typology of equivalences frequently used in marketing research. The subsequent section presents and examines the equivalences of qualitative cross-cultural research. Finally, as a conclusion the suggested contributions to qualitative marketing methodology are summarised.

¹ A generic definition of qualitative research is used here: Qualitative research “consists of a set of interpretive practices that make the world visible” by the participating role of an observer. These practices “turn the world into a series of representations”, such as field notes, interviews, memos of self, and other types of collections of data. The world is studied in its “natural settings”, and the aim is to “make sense of” it, or interpret the world “in terms of the meanings people bring to it”. (Quotations from Denzin and Lincoln, 2000, p. 3)
THE CONCEPT OF EQUIVALENCE

In a review of the use of equivalences in survey research, Johnson identified more than 40 different definitions of the construct of equivalence (Johnson, 2003). In survey research, the concept of equivalence generally refers to “the comparability of measures obtained in different cultural groups” (Harkness, et al., 2003, p. 14). However, in the current study the construct is used in a broader sense. Here the concept of equivalence in the context of cross-cultural methods refers to the comparability of two or more data sets from two or more distinctly different nations or cultures. Reaching equivalence in comparative research guarantees that, instead of possibly being caused by unequal methods, results show differences or similarities across compared samples. A failure to reach equivalent methods is called non-equivalence, while the term dissimilarity refers to a difference that yields an interesting finding per se.

In comparative quantitative research a large number of different types of equivalences exist. Table 1 summarises a typology of equivalences examined in the literature of marketing and management research. However, not all the equivalences are applicable to or relevant for qualitative comparative research. Equivalences that directly address the comparability of a scale, such as instrument, measurement and scalar/metric equivalences, are not transferable to research that does not use measurements. Thus, in addition to offering definitions of equivalences and marketing literature references to equivalences, the table also highlights equivalences that can and should be addressed in qualitative cross-cultural marketing research.

TABLE 1
Overview of equivalences and their applicability in qualitative research

<table>
<thead>
<tr>
<th>Equivalences</th>
<th>Definitions</th>
<th>References</th>
<th>Applicable in qualitative research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construct</td>
<td>Constructs have the same meaning and significance across cultures</td>
<td>Singh, 1995; Malhotra, Agarwal and Peterson, 1996; Cavusgil and Das, 1997; Craig and Douglas, 2005; Kumar, 2000; Herk van, Poortinga, and Verhallen, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>Conceptual</td>
<td>Concepts in two different cultures can be meaningfully used in the same contexts of these cultures</td>
<td>Hui and Triandis, 1985; Peng, Peterson, and Shyi, 1991; Brislin, 1993; Singh, 1995; Malhotra, Agarwal and Peterson, 1996; Cavusgil and Das, 1997; Usunier, 1998; Salzberger, Sinkovics, and Schlegelmilch, 1999; Kumar, 2000; Steenkamp and Ter Hofstede, 2002; Schaffer and Riordan, 2003; Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>Functional</td>
<td>Similarity between the goals of two behaviours</td>
<td>Hui and Triandis, 1985; Peng, Peterson, and Shyi, 1991; Singh, 1995; Malhotra, Agarwal and Peterson, 1996; Cavusgil and Das, 1997; Usunier, 1998; Salzberger, Sinkovics, and Schlegelmilch, 1999; Kumar, 2000; Steenkamp and Ter Hofstede, 2002</td>
<td>yes</td>
</tr>
<tr>
<td>Category</td>
<td>Similarity in categories in which objects are placed</td>
<td>Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>Salzberger, Sinkovics, and Schlegelmilch, 1999; Kumar, 2000; Steenkamp and Ter Hofstede, 2002; Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td></td>
<td>-----</td>
</tr>
<tr>
<td>Instrument</td>
<td>Similarity in instruments</td>
<td>Singh, 1995; Malhotra, Agarwal and Peterson, 1996; Cavusgil and Das, 1997</td>
<td>no</td>
</tr>
<tr>
<td>Measurement</td>
<td>Same operationalization of measures (sometimes used as synonyms to calibration equivalence)</td>
<td>Mullen, 1995; Malhotra, Agarwal and Peterson, 1996; Steenkamp and Baumgartner, 1998; Usunier, 1998; Steenkamp and Ter Hofstede, 2002; Craig and Douglas, 2005; Ewing, Salzberger and Sinkovics, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>no</td>
</tr>
<tr>
<td>Calibration</td>
<td>Same measurement units</td>
<td>Mullen, 1995; Malhotra, Agarwal and Peterson, 1996; Steenkamp and Ter Hofstede, 2002; Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>Translational</td>
<td>Same meaning of measurements after translation</td>
<td>Peng, Peterson, and Shyi, 1991; Mullen, 1995; Malhotra, Agarwal and Peterson, 1996; Usunier, 1998; Salzberger, Sinkovics, and Schlegelmilch, 1999; Steenkamp and Ter Hofstede, 2002; Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>Scalar/metric</td>
<td>Respondents respond to scales in the same way</td>
<td>Peng, Peterson, and Shyi, 1991; Mullen, 1995; Malhotra, Agarwal and Peterson, 1996; Usunier, 1998; Salzberger, Sinkovics, and Schlegelmilch, 1999; Steenkamp and Ter Hofstede, 2002; Craig and Douglas, 2005; Herk van, Poortinga, and Verhallen, 2005</td>
<td>no</td>
</tr>
<tr>
<td>Contextual</td>
<td>Similarity in cognitive and sometimes even emotional processes that respondents experience while responding to questions.</td>
<td>Braun, 2003</td>
<td>modified</td>
</tr>
<tr>
<td>Sample</td>
<td>Equivalence in sample frames and sample selection in order to establish comparability.</td>
<td>Cavusgil and Das, 1997; Ember and Ember, 1998; Kumar, 2000; Schaffer and Riordan, 2003; Craig and Douglas, 2005</td>
<td>yes</td>
</tr>
<tr>
<td>Response style</td>
<td>Similarity in the ways and manners of respondents when answering questions</td>
<td>Sekaran, 1983; Usunier, 1998; Baumgartner and Steenkamp, 2001; Smith, 2003</td>
<td>yes</td>
</tr>
</tbody>
</table>

It is important to discuss the epistemological assumptions of the concept of “equivalence”, as the concept has mainly appeared in cross-cultural quantitative research. Traditionally, the objective of quantitative cross-cultural
research has been to replicate the same study in different cultures, in order to establish whether a phenomenon can or cannot be generalised across several cultures. Quantitative research assumes that, given equivalent methods, cross-cultural comparisons are possible.

The goals of cross-cultural qualitative research and its view of comparability are different. Further, both the goals and the view of comparability vary across different qualitative methodological approaches. Even if generalizations are not explicitly sought some qualitative approaches assume that, at least findings can be transferred across cultures. In a similar vein, as results from one case can be transferred to other similar cases (see Lincoln and Guba, 1985; Miles and Huberman, 1994; Wallendorf and Belk, 1989), transferability refers to the possibility of comparison. The goal may not be to transfer findings but to compare these with established knowledge, and if possible to bring new insights into existing knowledge (for an example see Joy, 2001 and Arnould, 1989) or to deliver a broad view on a new concept (see Belk, Ger, Askegaard, 2003). Quantitative research compares research findings across samples, whereas qualitative research compares new findings or locally valid substantial theories with established theories (cf. Glaser & Strauss, 1967). Even though the primary goals of quantitative and qualitative research are different, both compare findings with existing theory. Thus, both may view comparability as etic\(^2\), meaning that methods can be universally applied across cultures and that comparability is possible. However, there exist some qualitative perspectives contending that cultures are so unique that they cannot and should not be compared, i.e. they have a purely emic\(^3\) view of comparability. Alternatively, they argue that in the course of globalization the entire field of cross-cultural research should be challenged, and replaced by multi-sited ethnography (Kjeldgaard et al., 2006).

As face-to-face personal interviews are perhaps the most commonly used qualitative marketing data collection method, the present paper focuses on interviews. Thus, the conceptual analysis is exemplified and illustrated by examining the interview in qualitative marketing research and further illustrated by examples from the article by Knight et al. (2006), which compare interview data from several cultures.

**EQUIVALENCE IN CROSS-CULTURAL QUALITATIVE RESEARCH**

This section discusses the application of conceptual, functional, sample, contextual, researcher and response equivalences to cross-cultural qualitative marketing research. The other equivalences that are applicable to cross-cultural qualitative research (cf. Table 1), i.e. construct, calibration and translational equivalences, have been excluded in the discussion for the following reasons. Construct equivalence refers to similar meaning and significance of terms. The significance of a construct can be detected when doing quantitative research, whereas meaning is partly covered by an examination similar to the one concerning conceptual equivalence. Thus, to avoid repetition, construct equivalence is not included in the examination below. Translational and calibration equivalences face similar concerns in both qualitative and quantitative marketing research and are therefore excluded here.

The methods section by Knight et al. (2006) is used to exemplify ways of addressing equivalences. The study compared food imports in Germany, the Netherlands, Greece, Italy, and the UK. The material consisted of 17 personal interviews: four in Germany, one in both the Netherlands and Greece, five in Italy, and six in the UK.

**Conceptual equivalence**

The first issue in cross-cultural marketing research is to determine whether the concepts have similar connotations in different cultures. Conceptual equivalence means that concepts can be meaningfully examined in the same contexts in two different cultures (Hui and Triandis, 1985). However, conceptual equivalence does not mean that the concepts in the study have *exactly* the same meaning in both cultures, as the same concept may have different as-

\(^2\) The term etic originally comes from anthropology and refers to the development of pan-cultural methods (De Beuckelaer, 2005).

\(^3\) The term emic refers to culture-specific measures and can be seen as the opposite to etic (De Beuckelaer, 2005).
pects and dimensions in different cultures, even if it serves the same purpose in those cultures (Brislin, 1993). For example, perceived power as a concept has a variety of factor structures in different cultures (cf. Johnson, Sakano, Cote, and Onzo, 1993). In qualitative research, disparities in the definition of concepts and content of concepts across cultures may be an interesting finding per se. As the aim of qualitative research is to uncover substance-specific theories, different or similar concepts drawn from different cultures challenge and extend theoretical foundations and established knowledge.

Thus, in qualitative cross-cultural research equivalences and dissimilarities are examined for the purpose of theory building and not only in order to establish data comparability. While conceptual non-equivalence is seen as failure in equivalence (Usunier and Lee, 2005), in qualitative research it is a source for unexpected findings (see for example Eisenhart, 1989). In addition, qualitative methods are often used to uncover differences in a concept or to detect emic codes that may lead to the discovery of new dimensions and constructs in established concepts (see for example Price, Arnould & Tierney, 1995).

However, even when the concepts of multiple comparable studies are equivalent, their underlying latent constructs or dimensions may not be exactly similar. Qualitative research methods uncover codes which emerge through abstraction into constructs that form concepts (Spiggle, 1994). In qualitative research, like conceptual equivalence, dimensional dissimilarity is an interesting finding per se. The varying internal structures of a concept across cultures uncovers the emic meaning of the concept, and thus points to cultural differences.

Existing data such as metaphors, proverbs, and comparisons of antonyms and synonyms may uncover differences in concepts across cultures even before data collection, or may strengthen the findings of interviews. For example, a finding drawn from the analysis of jargon may lead to a novel dimension in the concept of mobile phones. A Chinese nickname for mobile phones, dageda, ‘big-elder brother-big’, reveals the high status of a cell phone. In addition to dimensions of functionality, usability and so on, the dimension of prestige could be added to this concept.

To sum up, in qualitative cross-cultural marketing research the findings from the assessment of conceptual equivalence are interesting, even when a dissimilarity is found. All other conditions being equal across cultures, conceptual dissimilarity may provide new aspects of abstract theoretical concepts and their meanings. Different qualitative approaches provide new views on how to seek conceptual equivalence, dissimilarity and meaning in concepts both due to external conditions (such as cultural, historical and societal) and research-coresearch-respondent internal relations, which will be examined later in the paper.

**Functional equivalence**

Functional equivalence has been given several definitions. For example, in cross-cultural psychology it is defined as a synonym for construct equivalence (Johnson, 2003). However, construct equivalence refers merely to equivalence in concepts, whereas functional equivalence also encompasses the idea of dissimilarity and similarity in goal antecedents and outcomes of the same behaviour in different cultures. This paper draws on Hui and Triandis’ (1985) definition of functional equivalence, in which the goals of similar behaviours are compared. *Functional equivalence* which means ‘similarity between the goals of the two behaviours’ has been distinguished from conceptual equivalence (Hui and Triandis, 1985). Conceptually equivalent concepts in different cultures may have different functions and also different outcomes. In other words, functional equivalence is interpreted by Hui and Triandis (1985) to include the assessment of equivalence in the nomological network of constructs and concepts (Cronbach and Meehl, 1955; Steenkamp and van Trijp, 1991). For example, perceived power influences satisfaction perceptions in some cultures but may be more acceptable in other cultures, thus also influencing other aspects of a relationship such as the intensity of a relationship. For example, the function of television is perceived differently in different cultures. In one culture, television is considered a means of information and entertainment, whereas in another it is a gathering-point for social activities (Gould & Wong, 2000). This conceptual dissimilarity in function influences the studies that use the concept of television and therefore the theory and interpretation of results.
The nomological network provides information on similarity in statistical associations between multiple concepts and constructs (Singh, 1995) but not on the goals that are embedded in respondents’ behaviour. Personal interviews may produce clues for conscious goals of behaviours, and participant observations may uncover unintentional and tacit objectives of actions. The findings of a dissimilarity/equality assessment is the discovery of new functions, goals, and consequences for known concepts and behaviours, which will lead to varying nomological networks across cultures. The specific type of network of concepts and constructs can be identified from interview narratives by the integration of categories (see Spiggle, 1994) but can be further strengthened by using supplementary methods. Such methods could be the analysis of respondent-created fairy tales, folktales, traditional fairy tales, proverbs, and high and popular art (news, short stories, visual art, music, films, TV programmes). For example, a Chinese pop-song on TV tells about a son’s gratitude and obligation towards his parents, which reveals an important conceptual network. Gratitude can be interpreted as the antecedent of obligation in China, in parent-child relationships in particular, but also in any relationships in general. Another example is proverbs, which often express traditional cause-consequence events and value systems in a culture.

Similar to conceptual equivalence, functional equivalence in qualitative cross-cultural research is not sought for methodological purposes only, but because the similarity or dissimilarity found is a result in itself. Functional equivalence is not only a question of comparability across populations. Qualitative data are functionally comparable even when there are differences in the functions of a phenomenon. It is important, however, to discuss the functional equivalence or dissimilarity of findings, and to strive to offer explanations for similarities or differences in different cultures.

Sample equivalence
In order to establish comparability it is important to have equivalence in sample frames and sample selection. However, rather than being identical, sampling frames in different cultures should be relevant and suitable for the aims. It has been argued that random sampling is preferable and that it provides fewer systematic biases when generality is sought (Ember and Ember, 1998). However, probabilistic sampling is not always possible in less developed countries (Craig and Douglas 2005), and therefore other types of sampling methods need to be considered. In addition, identical sampling procedures or methods in each culture are not as important as ‘equivalent levels of accuracy and reliability’ (Ibid).

A determinant of sampling in cross-cultural qualitative design is the research approach or design. If the approach is inductive and emic, as it often is in qualitative research, then purposeful sampling could be serendipitous and induce a snowball or chain strategy (Patton, 2002). Snowball or chain sampling is drawn via respondents who provide contact information of other potential respondents (Miles and Huberman, 1994). During fieldwork, serendipitous sampling allows the researcher to follow new leads and unexpected findings, which suits emic and inductive approaches well.

Serendipitous sampling is also purposeful in a creative way (as suggested by Spiggle, 1994 and Wallendorf and Belk, 1989). Serendipitous sampling refers to an emergent sampling design, which follows the new leads based on preliminary emergent results. However, chain or serendipitous sampling may also constitute results. For example, when distribution channels are studied in different countries, managers of retailing outlets could be a starting-point of sampling and data gathering, as they can be easily found at street-level shops. They provide information on the retailing level and give access to respondents at wholesale and production levels. Hence, snowball sampling is not only sampling, but at the same time a means to uncovering distribution channel structure, which could be part of the results of a study (cf. Arnould, 2001). Similarly, a possible and interesting result of extreme or typical case sampling could be that a case that is typical in one culture is an extreme case in another culture (cf. Patton, 2002).

When detecting results yielding the sampling design, the process of serendipitous sampling and detailed reasoning should be documented throughout the fieldwork. In order to assess sample equivalence, the track of chain sampling should be recorded and compared across cultures. However, the main point here is that, in qualitative cross-cultural research sampling, equivalence is not always the goal, because tracing and comparing multiple and vary-
Ing sampling processes may provide important findings. Depending on the purpose of a study, these findings may contribute emergent theory (theoretical sampling) and/or the contribution may consist of empirical data (chain and serendipitous sampling). Carefully recorded and reported sampling procedures across cultures permit determination of whether detected dissimilarity in the sampling process qualifies to become a result or is caused by bias in the sampling procedure.

Knight et al. (2006) provide a detailed description of a sample in the form a table that displays interviewee positions, sectors of food industry, company locations in a given country, and business activities. Thus, the article provides information for assessing equivalence in the sample frame and processes. As the access to respondents in the Netherlands differed from the access to informants in other countries, non-equivalence could be detected only in the case of the sampling process. The sampling was purposeful but not serendipitous or theoretical, which would have enriched the findings that aimed to revise existing models of buying behaviour.

Contextual equivalence
Contextual equivalence refers to a cognitive and sometimes even an emotional process that a respondent goes through while responding to questions. The respondent subconsciously recalls different contexts in which the question is embedded, i.e. the respondent tries to ‘clarify the meaning of a question from its context’ (Braun, 2003, p. 60). Respondents may employ at least three different contexts in interaction with each other: textual, experiential, and cultural. The textual context of any part of a questionnaire is the rest of the questionnaire. The experiential context refers to a respondent’s past experiences, which may be activated in memory. The cultural context consists of the external conditions that may influence the responses to the questions (Braun, 2003). While the two latter contexts exert direct influence on attitudes and values, the first context may create biases in terms of comparability. Therefore, in order to ensure the comparability of cross-cultural data, the ensuring of textual equivalence will be demonstrated below.

In qualitative research the textual context is labelled the verbal context. The verbal context not only refers to the textual context in the form of order of questions, but to a wider verbal and oral environment. In addition to questioning during an interview, the interviewer can chat, talk, and discuss the topic of the interview, and therefore the verbal context may consist of a variety of interaction during the course of the interview, from its initial introduction to the final closing of the interview. This verbal context around the questions should be as equivalent as possible across the investigated cultures. It should be noted that the verbal context in different cultures indeed differs a great deal. For example, in an Anglo-American context, more small talk in order to make respondents feel at ease is expected than is the custom in, for example, the Finnish culture.

In addition to the verbal context, there is another type of context that may influence the results of an interview, and consequently reduce comparability. The physical location of an interview may have an influence on the responses received. In some cultures it is acceptable to conduct interviews at the homes of respondents, whereas in other cultures the home is a private sphere and thus inappropriate for researchers to enter. For example, while conducting marketing channel research in China, I was invited by a respondent to conduct the interview in a hospital, where the respondent’s wife was hospitalised. This type of physical context would be unacceptable in most Western cultures. In terms of comparability the influence of the physical context varies on a continuum from appropriate to inappropriate seen from the point of view of the respondent, who should feel comfortable in the physical environment where the interview is conducted. To secure comparability, it is not important to seek identical physical locations for the interviews but to find locations that in the two cultures to be compared are equally placed on the scale of appropriateness.

In practice, equivalence of verbal context can be assessed from detailed interview transcripts and analysis of videotaped interviews, if available. When thorough transcripts consist of all the interactions (overlaps, laughs, and pauses) during an interview, they allow analysis of what happened before, during, and after a question was posed and an answer was given (Silverman, 2006). The verbal environment may influence both the way a question was posed and, more importantly, what and how the answer was given, since the phrases, body language, and small statements before a comment may influence answers. On the basis of detailed transcripts or videos, the equivalent
lence of verbal context can be assessed by analysing the verbal environment around a response. In contrast, the assessment of the equivalence of the physical interview context needs to be determined before the data collection. On the scale from appropriate to inappropriate, researchers can evaluate possible interview locations and choose those in the compared cultures that score similar values on the scale of appropriateness.

In the study by Knight et al. (2006) all but four of the interviews were conducted on the respondent companies’ premises, thus indicating a fairly good equivalence of physical context. As the personal interviews were ‘semi-structured’ and ‘steered to a large extent by the interviewee’ (Knight et al. 2006: 115), and no report or assessment of interview processes across cultures were given, the equivalence of verbal context cannot be assessed. However, the interview protocol was published in an appendix, which enables an evaluation of the verbal context at a generic level.

**Researcher equivalence**

In qualitative cross-cultural research, it is critical to seek equivalence of the researchers who collect and analyse the data. In most qualitative research traditions, the process of collecting and analysing data is considered to be subjective, which further strengthens the importance of researcher equivalence.

In qualitative research, researcher equivalence refers to equivalence in researchers’ traditions of collecting and analysing qualitative data. There are two major ways of collecting and analysing cross-cultural data. Firstly, a single researcher can conduct a study in several cultures. Secondly, a multi-cultural research team can be organized, consisting of researchers from different cultures who conduct research in their own native environment (cf. Easterby-Smith and Malina, 1999). In both cases researcher equivalence should be established. This equivalence and the discussion that follows below should not be seen to exclude the well-established benefits of multiple analysts of interdisciplinary, international, or different gender teams or the use of external auditors (e.g. Belk, et al. 2003; Belk, Sherry, & Wallendorf, 1988). Whereas multiple coders work with the same data, and gain insights from discussing the discrepancies in their analyses of those data, researcher equivalence is required when researchers work with separate data sets from several different cultures.

Depending on the interviewer who conducts the research, response effects may vary considerably and it is likely that such variation is inconsistent across cultures (Smith, 2003; Usunier, 1998). By *researcher equivalence* is here meant that those who conduct interviews or observations also have a similar or equivalent status in the societies where they perform the study. This does not mean identical equivalence in gender, age, position, or academic status, but rather an emic equivalence with etic dissimilarity. In other words, researchers do not need to be identical, but of a kind that provides maximally equivalent stimulus in an interview or observation situation. For example, in some cultures age is appreciated over professionalism, and in other cultures professionalism over age. If respondents’ respect of interviewers is required, then interviewers should be older professionals in one culture, whereas younger professionals may gain respect in another culture and are therefore appropriate for conducting interviews. Researcher qualifications can thus be different in different cultures (etic dissimilarity), even if the goals are the same (emic equivalence). This means that it is crucial that researchers adopt identical roles on one but not on another level. As has been pointed out: ‘Traditional rules mandate that methods should be identical across surveys, but the challenge is to identify cases in which methods identical on one level are not identical on other levels affecting measurements.’ (Smith, 2003, p. 82)

In some qualitative approaches, researcher equivalence in a multicultural team is emphasized further by stressing the cocreation of reality through participation in the phenomenon under investigation. Participation in a phenomenon of interest and in the context around the phenomenon is not possible for a single researcher, if s/he is not a member of the contexts that are compared. Although increasingly, only on rare occasions is one interviewer representative of more than one distinctive culture, and therefore in a multi-cultural research team participatory cocreation is limited. In the current global world, there exist individuals who are members and even insiders of several
cultures. This recent resource of the multi-cultural researcher has probably rarely been utilised in cross-cultural research, but is already acknowledged in multi-sited ethnography (Ekström, 2006). On an abstract level, as in the assessment of conceptual equivalence, the true insider’s role is crucial.

In practice, researcher equivalence can be assessed during different phases of the research procedure. First, before data collection an evaluation of researcher roles in the studied countries should be determined. In other words, the research design should first include the setting of the goals for researcher equivalence, depending for example on whether the informants respect ethnographers, whether the informants should feel equality with researchers, or whether both parties should come from the same social or age group. Next, the research design should determine which researcher qualifications meet the goals set. For example, if equality between the respondent and the researcher is sought, then the assessment should determine whether age, status, education or gender trigger inequality in the respective cultures. When the triggers of inequality have been detected, suitable researchers to conduct the study can be sought. Rarely, if ever, perfect researcher equivalence can be found. Therefore, when a single researcher conducts studies in all compared cultures, it is probable that his role varies somewhat across cultures. Therefore, possible non-equivalences during the data collections should be identified and documented. One solution is the careful analysis and comparison of detailed interview transcripts by researchers. However, more rigour can be achieved by using video recording of the interviews (see Belk & Kozinets, 2005). Analysis of interview films is one way to become aware of non-equivalences in interviews across cultures which might influence the comparability of results.

Along similar lines, Ember and Ember (1998) have discussed coder equivalence, which can be generalised to mean data analysis equivalence. If the cross-cultural researchers are monolingual, there is a need for more than one coder or interpreter of data, and then problems of coder and data analysis equivalence may arise. Preferably, coders and those who interpret the data should have equivalent knowledge of theory, setting, context and substance, be equally experienced in coding and analysing data, have similar knowledge of the society where the data come from, and be provided with the same coding and analysing instructions and rules (Cavusgil and Das, 1997; Ember and Ember 1998). These instructions are applicable to structured coding and, even more importantly, to open coding when inductive results are sought. For inductive findings, however, coding and data analysis instructions and rules should not be strictly defined beforehand, but coder equivalence should be controlled more stringently. Coder and data analysis equivalence is important for all qualitative cross-cultural research approaches and data collection methods, from observations to autoethnographic texts, legends, and stories.

In the study by Knight et al. (2006) all the interviews were tape-recorded and transcribed as well as coded by a single researcher. The validity of the work was checked by a reflective analysis of the role the interviewer played in the data collection. However, the reflections on his role as a coder remain unanalysed. The check of validity is not a check of comparability or researcher equivalence, as there is no discussion on how a single researcher was received in the studied countries. As German, Greek, Dutch, Italian and British cultures vary (see for example Hofstede, 2001), it is probable that respondents from these cultures do not perceive the interviewer in an equivalent manner. Further analysis of the researcher’s role separately for each country would enable assessment of researcher equivalence.

Closely related to equivalence in interviewer behaviour are respondent reactions. Therefore, response equivalence will be examined next.

Response equivalence

Response equivalence refers to the manners of respondents when answering questions. Previous research has identified eight response effects that influence the comparability of data (Smith, 2003). Some of these effects vary across cultures (Baumgartner & Steenkamp, 2001), and some are assumed to do so, but have not been tested across different populations. Depending on their cultural background, respondents may provide answers that are

---

4 Here the term multi-cultural researcher refers to an individual with multiple cultural, linguistic and national identities. The term does not refer to the use of native researchers which is common in cross-cultural research.
socially desirable, tell the interviewer what they assume the interviewer wants to hear (‘yea-saying’), use extremes or middle points in scales differently, may give no responses at all (‘no opinion’ and ‘don’t know’), tend to prefer first (primary effects) or last responses (recency effects), or may be influenced by physical context, the order of the questions, and the mode of administration (Smith, 2003). In their review of response styles, Baumgartner and Steenkamp (2001) additionally include ‘nay-saying’, directional bias (tendency to show acquiescence), use of narrow or wide range of response categories, and tendency to respond carelessly.

The definition of response equivalence is directly applicable to cross-cultural qualitative research. However, the response styles presented above are relevant for qualitative studies only to some extent. The responses to scales (due to use of extremes or middle points, primary and recency effects, and use of narrow or wide range of categories) cannot directly be applied to comparative qualitative research. However, the response styles of social desirability, yea-saying and nay-saying, no-opinion and don’t knows, and directional and careless responses can be applied to comparative qualitative research.

The mode of administration (i.e. whether the informant is responding, in person or over the telephone) also influences the styles of responses (Smith, 2003). In traditional qualitative interviews, face-to-face contact during interviews is the mode of administration. While Baumgartner and Steenkamp (2001) have studied response styles in survey research, and there is advice on how to overcome response biases in monocultural interview, face-to-face interview response styles across cultures are an uncharted territory. In the digitalised world of today, qualitative interviews are also conducted over mobile phones, the internet, and by using email, i.e. without interaction in face-to-face (see Masten and Plowman, 2003). At present it is only possible to speculate in what ways social desirability, yea-saying and nay-saying, no-opinion and don’t knows, and directional and careless responses during interviews face-to-face, over the phone, the internet or email differ across cultures.

Face-to-face contacts usually increase and intensify interaction between an interviewer and an interviewee. This interaction is the point of departure for the comparability of response effects, i.e. response equivalence. Depending on the role of the researcher as being superior, inferior or equal, of the same or the opposite sex, and having similar or dissimilar education, response styles may vary. Even though not yet studied empirically across cultures, it can be assumed that face-to-face contacts intensify response styles, in the sense that for example person-to-person responses tend towards more socially desirable answers or seek more acquiescence. However, this assumption is made under the cultural condition that in face-to-face contact people prefer to please the interviewer, but are more neutral and free in their responses when only responding to a questionnaire. The ways in which individuals across cultures respond to questions over the phone or the computer, and interact with technical devices, such as phones and computers, is still largely an unstudied subject.

While the penetration rates of phones, internet and email have been investigated across cultures, further research on response styles in face-to-face, phone, internet, and email interviews needs to be conducted, in order to identify the differences in response styles across cultures and modes of interview. Only then will it be possible to evaluate the role of response style differences regarding the comparability of cross-cultural interview data, and to determine whether general interview guides apply in cross-cultural settings.

Knigth et al. (2007) provide a nicely detailed report of how they tried to prevent response biases such as social desirability. However, the report does not contain country-level analyses, and thus does not provide tools for assessing response equivalence across the German, Dutch, Greek, Italian and British respondent styles studied. As there is evidence from quantitative studies that response bias varies across cultures (Baumgartner and Steenkamp, 2001 and Smith, 2003), country-level analyses of response bias in qualitative inquiry should be considered of central interest.

Summary of equivalences in cross-cultural research
Table 2 summarizes the types of equivalences useful for qualitative cross-cultural research proposed in this paper. As some equivalences are modified to be applicable for qualitative cross-cultural inquiry, other equivalences have
been redefined (cf. Table 1). In the second column the function of each equivalence is summarized, and finally the ways to assess each equivalence are briefly presented.

**TABLE 2**

Summary of examined equivalences for qualitative inquiry

<table>
<thead>
<tr>
<th>Equivalence</th>
<th>Definition in qualitative cross-cultural research</th>
<th>Function</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual</td>
<td>Two concepts can be meaningfully examined across cultures</td>
<td>Examined for the purpose of theory building (new concepts, dimensions, constructs)</td>
<td>Through coding interview transcripts and analysis of metaphors, proverbs, and antonyms/synonyms</td>
</tr>
<tr>
<td>Functional</td>
<td>The goals of two behaviours are similar</td>
<td>Examined for the purpose of theory building (new nomological networks, functions, goals, and consequences)</td>
<td>Through integration of categories (axial and selective coding) and analysis of fairy tales, folktales, traditional fairy tales, proverbs, and high and popular art</td>
</tr>
<tr>
<td>Sample</td>
<td>Two sample frames and sample selections are comparable</td>
<td>Examined for the purpose of theory building</td>
<td>Through detailed documentation of the process of sampling and justification of each step in the process</td>
</tr>
<tr>
<td>Contextual</td>
<td>Two contexts are comparable</td>
<td>Equivalence is sought</td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>Two oral and verbal environments of a question are comparable</td>
<td>Equivalence is sought</td>
<td>Through detailed analysis of interview transcripts and/or video tapes</td>
</tr>
<tr>
<td>Physical</td>
<td>Two physical locations of interviews are comparable</td>
<td>Equivalence is sought in appropriateness</td>
<td>Through detailed analysis of culturally appropriate locations of interviews</td>
</tr>
<tr>
<td>Researcher</td>
<td>Researcher(s) who collect and analyze data are comparable across cultures</td>
<td>Emic equivalence (which does not mean etic equivalence) is sought</td>
<td>Before data collection through evaluation of researcher roles, and after data collection detailed analysis of interview transcripts or/and video tapes</td>
</tr>
<tr>
<td>Response</td>
<td>Ways and manners of respondents are comparable</td>
<td>Equivalence in social desirability, yea- and nay-saying, no-opinion and don’t knows, and directional and careless responses</td>
<td>Through detailed analysis of interview transcripts or/and video tapes</td>
</tr>
</tbody>
</table>

The equivalences presented above, i.e. the conceptual, functional, sample, contextual, researcher and response equivalences, are not examined or assessed simultaneously, but rather studied in stages, during the different phases of the research process which will be presented next.

**EQUIVALENCES AND THE RESEARCH PROCESS**

Table 3 illustrates the courses of action in the qualitative research process, and positions the equivalences in the process. The sequence of equivalence assessments is central to the entire examination of equivalences. While some of the equivalences can only be studied after the data collection and analysis, the examination of others is vital before the initiation of the investigation, and still others need consideration throughout the research process.
The qualitative research process is compared with the quantitative process (from Cavusgil and Das, 1997), in order to illuminate the dissimilarities of the two research paradigms.

### TABLE 3
Recommended research steps

<table>
<thead>
<tr>
<th>Cross-Cultural Research</th>
<th>Qualitative</th>
<th>Quantitative (Cavusgil and Das, 1997, pp. 89-92)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set up the areas or topics of research interest, assess researcher equivalence</td>
<td>‘set up the theoretical domains of the research construct’</td>
<td></td>
</tr>
<tr>
<td>Acquire the applicability and equivalence in applicability of the research topic</td>
<td>‘acquisition and application of substantive knowledge about the conceptual and functional equivalence’</td>
<td></td>
</tr>
<tr>
<td>Serendipitous sampling and chain sampling, not a priori determined design, assess sampling equivalence and dissimilarity</td>
<td>‘creation of an effective and cost-efficient sampling design’</td>
<td></td>
</tr>
<tr>
<td>Decide research areas, interview guides, and observations guidelines, sampling of metaphors, proverbs, synonyms/antonyms, fairy tales, folklore, art and popular art in cultures of interest, assess sampling equivalence of data that are not gathered through interviews or observations</td>
<td>‘developing a sound instrumentation design’</td>
<td></td>
</tr>
<tr>
<td>Data collection takes place as concurrently as possible, if serendipitous findings leads to different research tracks in samples that are compared, assess context (verbal and physical), response, and researcher equivalences</td>
<td>‘data collection takes place as concurrently as possible’</td>
<td></td>
</tr>
<tr>
<td>Coding, data analysis, interpretation of meaning, and explanations by coders with similar knowledge background, assess coder equivalence, and evaluate conceptual and functional equivalence</td>
<td>‘use of multivariate methods’</td>
<td>‘confirmatory factor analysis’</td>
</tr>
<tr>
<td>Discussion of comparability of methods</td>
<td>‘external validity’</td>
<td></td>
</tr>
</tbody>
</table>

The first stage of a research project is the decision on the research aim and the applicability of the aim to various cultures of interest. While quantitative research starts by a theoretical discussion of constructs and hypothesis-building, qualitative inquiry is more inductive by nature. The starting-point is usually the assessment of areas or topics of interest that meet the importance and applicability criteria for the local communities to be compared. In other words, research topics need to be emic and substantive topics (Glaser and Strauss, 1967). When determining emic research topics, the assessment of researcher equivalence is crucial. Are the researchers equally knowledgeable about the compared cultures capable to assess the suitability of the research topic in the cultures to be studied? Do researchers have similar backgrounds to determine what is interesting to study in these cultures? These questions need to be answered before the final decision on the focus of the study.

The second stage is closely connected with the previous stage, as without researcher equivalence the entire research question cannot be addressed adequately. Whereas at this phase quantitative comparative research examines conceptual and functional equivalences, in qualitative cross-cultural inquiry only the significance of the research area is investigated. Is the research topic equally applicable, significant, interesting and important to the cultures to be compared?

The third stage involves sampling. While sampling design in quantitative research is often random, qualitative inquiry provides more alternatives to create an effective and purposeful sampling design (Patton, 2002). The sampling strategies that further enable the inductive and emergent nature of qualitative research are chain and serendipitous samples, which may lead to unexpected theoretical findings, and thus evolve into theoretical or conceptual samples (Patton, 2002). At this phase, a sampling strategy can be chosen which should be equally applicable in all the cultures under investigation. However, the entire process of sampling continues to the very end of data collection, i.e. it is initiated at this stage but overlaps the following two stages, to be presented next.
The fourth stage involves the design of data collection instruments. In quantitative research this means instrumentation and assessment of various equivalences connected to the measurement designs. In qualitative research it is necessary to plan a variety of guides and guidelines for collecting data, in the present case mainly interviews, and interview protocols. These protocols should meet the requirement of equivalence. Further, other sources of data that may shed light on the meanings and functions of concepts of interest can be collected at this phase. Sampling procedure and frame of metaphors, proverbs, synonyms/antonyms, fairy tales, folktales, art and popular art should meet the criteria of equivalence.

At stage five data collection starts. In both research paradigms data should be collected in all the different fields as concurrently as possible. However, data collection in the qualitative domain calls for the evaluation of a number of equivalences. These are the context, response and researcher equivalences. At this phase it can be assessed which types of physical environments are suitable and appropriate in order to yield good quality interviews. From detailed video- or audiotape transcripts, response, researcher, and verbal equivalences should be assessed. Going through raw data in the form of transcripts is important for the evaluation of equivalence in interviews across cultures. Reading interview transcripts for the purpose of equivalence assessment is a different process than final data analysis, and is solely done to find out how similar the interview situations were in different cultures. Similarity adds to the comparability of the results.

At the sixth stage all data have been gathered and the process of analysis can start. A number of statistical analyses assess equivalences in quantitative data (e.g. Mullen, 1995) but the discovery of equivalence in qualitative data is grounded in the researchers who code, categorize and interpret the data. Thus, again at this point the researcher equivalence is evaluated but not in their role as experts of a given culture but in their role as analyzers. Do coders have the same background? Do coders have the same theoretical background? Is their knowledge of the substance area and of the studied culture the same? Similarities in knowledge and substance background of coders facilitate the equivalence in the process of analysis.

The process of coding leads to categorizations, abstractions, and dimensionalizations (Strauss and Corbin, 1990; Spiggle, 1994) that yield concepts and the functions of these concepts. In contrast to the quantitative approach, conceptual and functional equivalences are addressed at this final stage of the qualitative research process. Depending on the aim of the analysis of concepts and their functions, findings may become results per se or can be treated as investigations of conceptual and functional equivalences.

The final stage in any research ends with an evaluation of the validity of the data, methods, and findings. The key dimension of evaluation is the comparability of two or more separate data sets. While traditionally the quality criterion of qualitative inquiry has been trustworthiness, credibility, transferability, dependability, conformability and integrity (Lincoln and Guba, 1985; Belk and Wallendorf, 1989), to the best of my knowledge comparability has not been previously discussed as an additional quality criterion to be considered. Thus comparability as a novel quality criterion in qualitative cross-cultural marketing research and as presented above is the key result and contribution of the present paper.

**CONCLUSIONS AND IMPLICATIONS**

This paper has presented some preliminary observations on cross-cultural qualitative research method more specifically on equivalences that have traditionally been used to assess comparability in quantitative cross-cultural research. Thus, the paper has not adopted a totally emic view based on the assumption that cultures are incomparable. Neither has it adopted the perspective of multi-sited ethnography, which challenges the basic premises of
The overall purpose of the paper was to examine whether the concept of equivalence could be applied to evaluate comparability in qualitative cross-cultural studies. The conclusion is that the concept of equivalence can indeed be applied to qualitative research. Further, using different types of equivalences would increase comparability, and consequently improve the quality of qualitative cross-cultural research. The practical guidelines provided in this paper suggest that the methods usually employed in order to increase trustworthiness in qualitative research can be utilised to increase comparability in cross-cultural qualitative marketing research (cf. Lincoln and Guba, 1985; Belk and Wallendorf, 1989; Miles and Huberman, 1994).

At the level of each specific equivalence the conclusions are as follows. Firstly, conceptual, functional and sample equivalences were here examined for the purpose of theory building. Therefore, the examination resulted in interesting findings rather than merely an assessment of comparability. Secondly, contextual equivalence was redefined in order to contain both verbal and physical equivalences. The concept of verbal equivalence does not only concern the order of questions, as does textual equivalence, but refers to the entire verbal context in which a question is embedded. The physical context refers to places where interviews are conducted, and it is assessed on a scale of appropriateness. People in different cultures vary in their perceptions of what is a suitable place for an interview and therefore the physical context should be controlled. Thirdly, as the role of the researcher is crucial in qualitative inquiry researcher equivalence is the equivalence that should be given most consideration. Researchers who collect data should have similar roles across cultures in their relations to their informants. Researchers who code and analyze data should have similar experiences and a comparable education, as well as the same guidelines for coding, in order to produce comparable results. The same applies both to a single researcher in two cultures and to multi-cultural research teams. Finally, response equivalence refers to the response options informants adopt while interacting with researchers. Whereas response equivalence has previously been studied cross-culturally, it has not been studied in the frame of qualitative inquiry or when using novel modes of communication, such as mobile phones, video conferencing and the internet.

Suggestions for Further Research
Clearly the preliminary observations presented above on equivalence and dissimilarity in cross-cultural qualitative marketing research call for further studies and further development of methods. Because of its exploratory nature, the paper invites further investigations of the topic. The paper has covered only some areas of the literature in cross-cultural marketing research. It has mainly referred only to those articles or books that deal with methods in cross-cultural research, but has here excluded other types of studies, such as empirical cross-cultural research using qualitative data or innovative data-gathering methods such as ‘netnography’ and mobile phones. A review of such material would provide useful documentation of the current practices in the field of cross-cultural marketing research.

The present paper considered only one type of qualitative data, that is face-to-face individual interview data. Also other qualitative empirical data should be examined regarding their comparability across cultures. Several methodological traditions exist in qualitative inquiry (for an overview, see Patton, 2002 or Lincoln and Guba, 2000), and qualitative methods also apply a variety of means for collecting data. For example, the following data-collecting methods are used in qualitative research: Face-to-face interviews (individual or focus group), web panels or ‘netnography’5 in general (Kozinets, 2002), observations (in a variety of forms, e.g. factory tours), images (photos, collages or videos), texts (narratives and interview transcripts), and secondary material, (such as advertising, online chat, editorials, letters to editors, novels, news paper/magazine articles, diaries, TV programs, art, music videos, and novels) among others. For example, the use of focus groups is influenced by both the social hierarchies of cultures and by the interdependence of the representatives of these cultures (Eckhardt, 2004), and therefore do not produce equivalent social interactions in a group. Further, qualitative researchers who use existing material (verbal or visual) should account for equivalences in sampling across cultures. For example, when

---

5 This term refers to ethnography over a digital device such as mobile phones and internet.
analyzing photos in family albums across cultures in order to compare consumer behaviour in travelling, it is necessary to establish the socio-demographic equivalence of the samples of families and even the equivalence of the samples of albums inside the chosen families.

The paper is purely theoretical and at present lacks empirical support. Therefore, it is suggested that, in addition to the theoretical development of the exploratory ideas of this paper, empirical data should be studied in the further development of the concepts of qualitative equivalence or dissimilarity. Further studies could also include a meta-study of recent cross-cultural qualitative studies in marketing, and other relevant disciplines, similar to the reviews of quantitative methods in cross-national marketing research (e.g. Peng, Peterson and Shyi, 1991).

REFERENCES


