1 Mixed Methods Cross-Cultural Research and Discourse

It is a truism that "data do not speak for themselves," implying that it is necessary for the researcher to interpret them. Nevertheless, some data do speak. In social science, vast amounts of data originate in talk. We interview people, and they respond, and all that talking eventually becomes the data – quantitative data in spreadsheets and qualitative data in coded transcripts. In cross-cultural research specifically, the talking either originates in bicultural, bilingual encounters or is very soon translated by bicultural bilinguals. In successive steps, natural languages are rendered into other natural languages, and folk languages are rendered into scientific languages, and talk becomes text. In all of these transformations, however, the actual experience of talking and listening is often lost, and (I would argue) a great deal of the cultural meaning is also lost. This book is about returning to the kind of survey and interview data that we already collect to find that meaning. To accomplish that, I describe a way to conduct linguistic analyses of mixed methods data, namely, a series of techniques focused on discursive interaction for interpreting both qualitative and quantitative data and for making more sophisticated cross-cultural comparisons. In this chapter, I discuss mixed methods cross-cultural research, and I sketch the notions of discourse that ground a discourse-centered integration of qualitative and quantitative data. Finally, because I use data from some of my own previous studies as examples throughout the volume, I provide some background to the cross-cultural Alzheimer's Beliefs Study (Schrauf and Iris 2011a, 2011b, 2014a, 2014b).

Quantitative Cross-Cultural Comparison

The standard quantitative, cross-cultural research design is a form of between-groups comparison. From a design perspective, such between-groups comparisons are more or less like an experiment, except that history or nature or circumstance has assigned the people to either group. In social science research, humans cannot be assigned randomly to cultural groups, as might be done in a true experiment (e.g., comparing mice or rats in different

experimental conditions). For instance, if the study compares parental authority and child-rearing in China versus the United States, the parent and child triads – the participants – are all living human beings in China and the United States, with personalities, life histories, and goals and aspirations. Cross-cultural comparisons are therefore quasi-experiments, and quasi-experiments comprise the majority of between-group comparisons throughout social science. The reasoning behind such designs involves testing for relationships between independent (predictor) variables and dependent (outcome) variables, and these relationships are specified during the research design process in the study's hypotheses. Continuing the example, we might hypothesize that vis-à-vis their own children, American parents are more likely than Chinese parents to feel ambivalent about the role of extended family members as disciplinarians.

The comparison itself requires careful consideration of whether and how the relevant beliefs or practices are similarly present in both cultural contexts. This is the task of establishing construct equivalence. In this example, we would have to examine what constitutes parental authority, family, and discipline in each culture. Perhaps the core features of parental authority are assumptions about age-appropriate behaviors of children, expectable forms of parent–child talk, norms of emotional expression, and beliefs about human development. If parents in both cultures exhibit or express such beliefs and practices, we would have a testable cross-cultural construct, and we might expect variations in these features in each culture.

Data-gathering instruments in the quantitative approach usually consist of closed-ended surveys, scales, checklists, or inventories that can be quantified. Such instruments need to be carefully translated, pilot tested, and their administration formalized so that the answers of every participant are independent and comparable to the answers of every other participant. If, for instance, the researchers administering the instruments in one culture were women of childbearing age, it would be important that researchers in the other culture be women of childbearing age as well. Were older men to gather the data in one culture, but young women in the other, we might expect some response biases in the data, relative to cultural assumptions about age and gender.

Finally, at the stage of data analysis, quantitative researchers use both descriptive and inferential statistics to assess the relationships between independent and dependent variables, as well as other variables, such as mediating, moderating, or confounding variables. For example, a researcher might use an analysis of variance to test for differences between Chinese and American parents (the independent variable of group) in their attitudes toward the extended family acting as disciplinarians (the dependent variable), while simultaneously considering the gender of the parent (a moderating variable). Of course, there are many other kinds of statistical analyses that can be

performed with such a data set (e.g., correlational analyses, multiple regressions, structural equation modeling), but the goal is generally the same: to model the relations between variables and to quantify the effects of predictor variables on outcome variables, in accordance with the researcher's theory-driven hypotheses articulated in the research design.

These, then, are the hallmarks of quantitative methods as applied to cross-cultural comparison: hypothesis-driven research, with attention to independent and dependent variables; designs that include quasi-experiments and correlational studies; attention to closed-ended data collection instruments; and descriptive and inferential statistical analyses. There are many variations on these methods that are part of a large and growing literature on cross-cultural quantitative methods (Matsumoto and van de Vijver 2011). In practice, the scope of quantitative cross-cultural research ranges from studies that mine existing, large-scale international data sets to test hypotheses at the level of multiple societies (see Ember and Ember 2009; van de Vijver, van Hemert, and Poortinga 2008) to face-to-face, small-scale comparisons of two or three cultural contexts. Many of the issues remain the same, however.

To exemplify the actual practice of quantitative research, allow me to offer a personal anecdote. Every Saturday morning for five years, I spent a couple hours reading manuscripts for the *Journal of Cross-Cultural Gerontology*, hunting for appropriate reviewers, reading their reviews, and composing decision letters to the authors. I was the associate editor, and my responsibility was to review the quantitative submissions, while the editor would review the qualitative submissions. I would carefully scrutinize their research design, preparation of instruments and stimuli, varieties of nonrandom sampling, multilingual data collection, and the statistical analyses that they used to model the relations between sociocultural variables and to suggest credible explanatory (though ultimately correlational) accounts.

Across a few hundred decision letters to manuscript authors, I found myself making the same comments over and over again – comments having to do specifically with language. For instance, what procedures did you employ to ensure that the translation of instrument X or survey Y from English to (whatever language) was appropriate for this community? What sorts of differences in cultural or linguistic norms did the bilingual translators encounter? (There *had* to be some bilinguals involved, right?). Some comments had to do with the traditional issue of construct equivalence. For example, what empirical evidence do you have that ensures that the concept of (for instance) social support or *nervios* or intergenerational remittances or filial piety or successful aging (or whatever) actually covers the same meanings and behaviors in the cultural contexts in which you collected data? Unfortunately, it is fairly common that a construct that has been operationalized in one culture (usually the United States) is only approximated in another culture. It is also fairly common

for researchers to acknowledge these differences, minimize them, and proceed apace. But, to be blunt, such discrepancies are exactly what are most interesting about cultural difference. So, I would say to those authors: tell us a bit about these differences, let us see what decisions you have had to make. Finally, some comments had to do with the representativeness of samples. Cross-cultural designs are quasi-experimental at best. Participants cannot be randomized into groups, and in practice the representativeness of participants in convenience samples (i.e., whoever volunteers) or even stratified convenience samples is often questionable because participants self-select (what are the cultural characteristics of the people who did *not* volunteer to participate?). Nevertheless, on the basis of such samples we often make broad comprehensive statements about whole populations. Ultimately, I would tell the authors of the manuscripts "give us more information about who these people are and what their cultural contexts are."

In my decision letters, I tried never to ask authors to collect additional data, especially if that involved doing additional interviews or surveys. Granted, *any* study can be improved by expanding the sample size or adding more items to the survey, but in cross-cultural research this is usually nearly impossible without the outlay of considerably more money and more time. Rather, reanalysis of the existing data could often address my – and the reviewers' – concerns. Further, although it might be frustrating, scaling back the claims made about group differences might be more scientifically defensible. More often than not, I was convinced that the answers were in the data somewhere and that a resubmission or revision would make a valuable contribution to the literature.

Upon reflection, the concerns I mention – translation, construct equivalence, and sampling – all trace back to questions of meaning and context. Do the words in these two data collection instruments mean the same things to speakers of two different languages? Does this construct have the same meaning to people in these two different cultural contexts? What makes these two cultural contexts comparable? These are the sorts of issues dealt with in the methods sections of quantitative articles, and, presumably, once they are resolved, one can get on with the business of doing the statistical analysis. However, the value of the quantitative modeling or analysis depends critically on precisely how these issues are solved. As I indicated earlier, there is a growing literature on methods for solving these issues, but, to anticipate my argument, I would like to suggest that the incorporation of qualitative methods would go far in addressing issues of meaning and context.

Qualitative Cross-Cultural Comparison

Qualitative cross-cultural comparisons are designed around research questions that seek to explore patterns of meaning, beliefs, behaviors, and practices of

people in different cultural contexts. As Creswell has pointed out, words such as "explore," "discover," and "describe" are of the essence of qualitative work (Creswell 2003). The focus of such research is on the experience of participants – their perceptions, opinions, evaluations, emotional framings, expectations, and agendas relative to the topic of the study. Not surprisingly, because meanings, beliefs, and behaviors are discovered, understood, and interpreted by researchers throughout their projects, it is not uncommon that the research questions themselves evolve over time.

Generally speaking, qualitative cross-cultural comparison shares the between-groups approach of quantitative research, if only in the sense that participants are recruited based on their membership in a particular group or participation in a particular context. However, whereas quantitative approaches assume a measureable central tendency (a mean, median, or mode) with quantifiable variation around it, and then analyze the differences in that tendency between groups, qualitative approaches seek to represent, with as much descriptive detail as necessary, the range of variation across a group. Comparison between groups then involves displaying the similarities and differences exhibited by individuals within each group.

Research designs in qualitative approaches vary according to the methodological orientation of the researcher. Some standard approaches (see Lichtman 2013) are *narrative analysis*, in which researchers collect participants' stories about a particular phenomenon and examine the cultural logic that undergirds those stories; *phenomenological analysis*, which seeks to capture the lived experience of individuals in different cultural contexts; *grounded theory*, which seeks to derive from participant interviews their own theory about a phenomenon or process (instead of imposing one from without) and then compare these theories across cultural groups; and *ethnography*, in which participant observation, usually over an extended period, gives the researcher sufficient exposure to the phenomenon to describe and interpret it.

Qualitative data almost always consist of text (rather than numbers). Interviews, discussions, social and political events, and rituals, even when video-recorded, are usually transcribed. Events and conversations experienced during participant observation are described in field notes. Obviously, archival and documentary artifacts (e.g., letters, emails, newspaper articles, pamphlets, and websites) are also considered texts. Analysis of textual data invariably requires a coding process, and codes reflect a mix of sources, including the methodological orientation of the researcher (i.e., coding for narrative development differs from coding for grounded theory), the disciplinary commitments of researchers (e.g., anthropologists code for different themes than sociologists), the participants themselves (this is called *in vivo* coding, where a participant's articulation itself becomes a code), and the research topic itself (e.g., where codes derive from the literature in that topic).

Finally, the qualitative between-group comparison involves matching coded material, some showing between-group similarities and some showing between-group differences, and then interpreting these differences. What explains cross-cultural differences is an interesting question. Sometimes the participants themselves offer their own explanations for their beliefs and practices, but often researchers reach into the literature for explanations. For instance, in a study of Chinese parenting practices, a Chinese participant may invoke a Confucian theme or the influence of Western media to explain his or her parenting behavior. Or, for example, in a study of water conservation, a Native American may invoke a deep traditional sense of unity with the environment to explain his or her behavior around conserving water, but more often than not, it is the researcher who, observing these behaviors, invokes Confucianism or Native American beliefs to explain them. The latter practice is perhaps more questionable; however, there is little question that appeals to the literature are a common means of interpreting observed differences between cultures. Finally, the presentation of results usually takes the form of excerpts of data (e.g., quotations from participants embodying a particular theme) or tables summarizing such themes, or perhaps both.

Not surprisingly, the majority of qualitative data come from interviews. Narrative analysts usually gather stories during interviews; phenomenologically oriented researchers do multiple interviews with a small numbers of participants; researchers employing grounded theory probe for local theories via interviews; and ethnographers mix interviews with participant observation: "If you want to know, ask." However, as any cross-cultural researcher knows, it is more complicated than that. Although we often frame interviews as seemingly natural conversations, there remain certain formalities, role expectations, clear power differentials, and often (gasp) monetary compensation. There is also the task of translating written documents or interpreting oral communication from one language to another. Oddly, in much cross-cultural research these issues are often dispatched as merely technical problems. In analyzing what participants say, researchers tend to ignore the role of the interviewer and focus exclusively on the respondent's answers. Interviews are conducted by bilinguals and often transcribed and translated, usually into English. As I noted above, these translations are then coded for content, usually via a mix of the researchers' codes and participants' categories – innocently enough – into social scientific English. In the end, qualitative approaches to cross-cultural comparison, while certainly getting us closer to the "real" cultural experience of the participants, are nevertheless problematic.

Because of the emphasis on face-to-face data collection and the time and energy it takes to conduct, transcribe, and translate the interviews, the scope of most qualitative cross-cultural comparisons is relatively small. Generally, researchers compare only two or three cultural groups, although there are

examples of team-based projects involving multiple societies (e.g., Hirsch et al. 2010). Further, in comparison to quantitative studies, sample sizes are almost inevitably small (i.e., fewer than thirty), even in team-based projects, and the question of the generalizability of results lingers. Thus, qualitative research usually involves a practical trade-off between time-intensive, face-to-face personal involvement, followed by careful transcription and coding, all of which are necessary to capture the lived experience of participants, and the generalizability of findings to the larger cultural group. Marrying qualitative cross-cultural research with quantitative research would thus be an attractive option.

Mixed Methods Cross-Cultural Comparison

Creswell (2015) defines mixed methods as research in which quantitative data (usually generated via close-ended questions or items) and qualitative data (usually understood to result from open-ended queries) are integrated and interpreted to address research questions. In these reflections on quantitative and qualitative approaches in cross-cultural research, I have drawn attention to the virtues and shortcomings of both methods and suggested that mixed methods approaches combine the best of both worlds. This balancing of strengths and weaknesses is a standard way of introducing the benefits of mixed methods approaches. The signal strength of quantitative studies is that they offer the possibility of modeling social and behavioral phenomena on a larger and perhaps more precise scale than do qualitative studies. The principal virtue of qualitative studies is that they provide a window into the personal meaning and lived experience that is lost in the abstractions of quantitative research. The combination of approaches capitalizes on the strengths and overcomes the weaknesses of either approach separately, as quantitative modeling is balanced by qualitative attention to lived experience, and the small samples and resultant problems with generalizability of qualitative work are offset by the larger, representative samples of quantitative work. In cross-cultural comparisons, where social and cultural contexts are key factors, an approach that captures the range of macro- and micro-influences on human behavior is unquestionably ideal.

As Creswell (2015) notes, mixed methods require attention to rigorous methods on both the qualitative and quantitative sides, and in particular, a well-articulated rationale for combining the data types and analysis. Helping researchers to deal with these issues of data integration and theoretical frameworks is one of my purposes in this book, which is to develop an overarching framework for integrating data. In that vein, this book is part of a growing literature on the design and practice of mixed methods research. Several handbooks (Creswell and Plano Clark 2010; Morse and Niehaus 2009; Tashakkori

and Teddlie 2010; Teddlie and Tashakkori 2009), a collection of seminal readings (Plano Clark and Creswell 2008), and journals dedicated specifically to the topic, the *Journal of Mixed Methods Research, International Journal of Multiple Research Approaches*, and *Quality and Quantity*, are also available.

In early work on the use of mixed methods in evaluation research, Greene, Caracelli, and Graham (1989, p. 251) articulated five common purposes for integrating data types. In what follows, I list the key words from Greene et al., but provide my own examples using the term "surveys" to represent quantitative research and "interviews" to represent qualitative approaches.

1. *Triangulation* – as when a survey and interviews with samples from the same population show similar results. That is, results from both methods corroborate one another.
2. *Complementarity* – as when interviews provide illustrations or elaborations of survey results. Oftentimes, qualitative data provides explanations of patterns in survey data.
3. *Development* – as when interviews or focus groups are used to develop items for a survey with a larger sample. This is a common use of qualitative data in the field of psychometrics for the purpose of constructing tests or assessments.
4. *Initiation* – as when either the survey or the interviews contradict one another or suggest paradoxical results. In these cases, novel findings trigger new questions and insights.
5. *Expansion* – as when a survey is followed by interviews with a particular strata or subsample from the survey population.

In practice, there can be multiple purposes for combining data types, as for instance, it is easy to see that triangulation and complementarity are both at work when an investigator provides an excerpt from an interview that illustrates and explains a survey result.

The purpose(s) of mixing methods are embodied and initially articulated in the research design of a study, and typologies of research designs and notations for expressing them have been a consistent feature of the mixed methods literature. A number of taxonomies exist for modeling the relations between quantitative and qualitative phases: Morse (1991) articulated one of the earliest examples, and Creswell et al. (2003) provide a more recent example. The most common notation system is as follows: first, the methods are abbreviated **qual** for qualitative, **quan** for quantitative. Second, the concurrence or sequentiality of the methods is indicated either by a plus (+) sign, indicating that both types of data are collected simultaneously (**qual + quan**), or by an arrow, indicating that one type of data collection precedes the other (**qual → quan** or **quan → qual**). Third, priority of method is signaled by capital letters. For instance, if the primary purpose of the study is to produce and validate an instrument, and qualitative interviews are conducted to generate items for the instrument, the

qualitative strand supports the quantitative strand, and the notation would be **qual → QUAN**. Conversely, if the study seeks to explore the creative habits of major figures in a field and then assess whether these habits are commonly found in practitioners of the field, the study design might prioritize the qualitative interviews, giving a **QUAL → quan** design.

Although the prototypical designs generally include two phases or strands, more may be added. In a recent introduction to the field, Creswell (2015) suggested that there are three basic research designs. A brief consideration of them gives a flavor for the terminology and range of options.

- The *convergent design* involves collecting both kinds of data, followed by analyzing each and then comparing and integrating the results. It is easy to see how this design serves the purposes of triangulation (each data set corroborates the results from the other). Priority might be given to either method (**QUAL** or **QUAN**), and either sequencing or concurrence (**→**, **+**) might be appropriate.
- The *explanatory sequential design* involves using qualitative methods to explain the results from a quantitative study, giving a **QUAN → qual** approach if emphasis is on the quantitative results or a **quan → QUAL** approach if the emphasis is on the qualitative findings.
- The *exploratory sequential design* begins with qualitative methods to investigate a little-known phenomenon or perhaps to solicit local articulations, and then uses these descriptions or articulations to develop a quantitative instrument (scale, survey, or checklist), which is validated in a larger population. The resulting design is **qual → QUAN**, with the quantitative stage stretching throughout the creation, piloting, and validation of the instrument.

From these three basic designs, Creswell (2015) develops three advanced designs that complexify them in various ways. He describes an *intervention design* in which the investigator embeds a qualitative component (convergent, explanatory, or exploratory) before, during, or after an experimental trial. In a *social justice or transformative design*, the study as a whole, and each of its stages, is framed around the goal of attending to the rights and opportunities of people in society. Finally, *multistage evaluation designs* extend and expand the qualitative and quantitative stages over time to assess and improve a program or service. Interestingly, what makes each of these an advanced design differs in each case: the first adds an experiment, the second elaborates a framework, and the third is longitudinal. Given this variety of criteria, I suggest that cross-cultural designs are an additional category of advanced designs.

The distinguishing feature of a cross-cultural design is its purpose to make between-group cultural comparisons, which in practice amounts to conducting the basic or advanced design (convergent, exploratory, explanatory, intervention, social justice/transformative, or multistage evaluation) in parallel across

two or more cultural or ethnic groups or contexts. Adopting the notation system described above, and assuming for the sake of simplicity a two-stage design in which priority is given to the qualitative component, we might picture a typical cross-cultural design as shown in Figure 1.1.

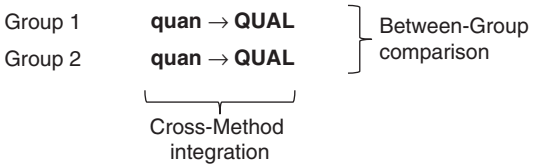

Figure 1.1 Example of a Cross-Cultural Design

Within-stage/within-method comparisons derive methodologically from the kinds of comparisons already used in qualitative and quantitative research. For instance, in the qualitative strand, the between-group comparisons are driven by comparing and contrasting thematically selected textual excerpts, and in the quantitative strand, the between-group comparisons are driven by standard statistical procedures, such as t-tests; analysis of variance; or even visual comparison of data maps as in multidimensional scaling, correspondence analysis, or principal components analysis. Across methods, the data types are integrated according to requirements of the purpose of the design (e.g., triangulation, complementarity, and expansion) and possibly according to the overarching framework that drives the study (see next section). Finally, cross-method integration and between-groups comparison are brought together in a final cross-cultural synthesis.

The question of integrating data types goes beyond the preceding articulation of purposes and typologies for mixing methods and includes more extensive, overarching frameworks that drive researchers' explicit choice of mixed methods over either qualitative or quantitative methods alone. I will explore such frameworks in detail in Chapter 2, but here I note that the mixed methods literature is currently grappling with this issue, and a consistent language for dealing with the issue has yet to emerge. Creswell (2015) uses the term "frameworks" to encompass both theories and philosophical perspectives. On the one hand, theories seem to be disciplinary and project-specific, as "for example, a researcher may use a leadership theory to advance an explanatory sequential design and to present both the quantitative and qualitative results" (p. 8). On the other hand, "philosophical frameworks are general beliefs and assumptions about research, such as how researchers discover knowledge" (p. 8). For instance, qualitative research has been associated with the philosophical perspective of constructivism (the notion that reality is a human social construct versus a pre-given, uniform world to which all adapt), and mixed methods