

A Basic Guide to Cross-Cultural Research

Carol R. Ember and Melvin Ember

This brief guide takes you through the basic steps of a cross-cultural study using the HRAF Collection of Ethnography — on paper, fiche, or online (*eHRAF World Cultures* --- <http://ehrafworldcultures.yale.edu>).

After reviewing the history of the HRAF Collection of Ethnography, we start with a discussion of the kinds of questions cross-cultural researchers investigate, and why they want to investigate them. Then we discuss how to choose a sample, how to develop measures, and how to analyze results.

- **Introduction**
- **Organization and Classification**
 - Selection of Cultures
 - Source Materials
 - Classification
- **What's the Question?**
- **Sampling**
 - Why use the HRAF Collection of ethnography?
 - Sampling within the HRAF Collection
- **Measures**
- **Analysis of Results**
- **References**

Introduction

The growing concern of students, scholars, and the general public to understand ethnic conflict, cultural diversity, and global problems has generated a demand for educational and research programs emphasizing the worldwide, comparative study of human behavior and society. The development of cross-cultural and area studies requires a large mass of readily available, organized cultural information; conventional sources of such information are widely scattered and often inaccessible, and often too expensive to assemble and utilize effectively. The HRAF Collections are designed to overcome this traditional barrier to research.

The HRAF Collection of Ethnography is a unique source of information on the cultures of the world, and as of April 2008 the complete collection contained over a million pages of indexed information on about 400 different cultural, ethnic, religious, and national groups around the world. The collection was developed by the Human Relations Area Files, Inc. (HRAF), a non-profit research organization based at Yale University. For almost fifty years, HRAF has served the educational community and contributed to an understanding of world cultures by assembling, indexing, and providing access to

primary research materials relevant to the social sciences, and by stimulating and facilitating training and research in these fields.

Development of the HRAF Collections began with the belief that enduring generalizations about human behavior and culture will emerge from a wealth of knowledge about the ways in which the different peoples of the world live. In 1937 at the Institute of Human Relations, Yale University, under the direction of the Institute's Director, Mark A. May, and Professor George Peter Murdock, a small group of researchers attempted to design a system by means of which the cultural, behavioral, and background information on a society might be organized. A fundamental part of that system was a universal topical classification scheme, the *Outline of Cultural Materials--OCM* (Murdock et al. 2008), which is still integral to the work HRAF does today.

In 1949, the Human Relations Area Files was incorporated in the State of Connecticut, with Harvard University, the University of Oklahoma, the University of Pennsylvania, the University of Washington, and Yale University as its founding member institutions. These five were joined within the year by the University of Chicago, the University of North Carolina, and the University of Southern California. Today, hundreds of colleges, universities, libraries, museums, and research institutions in the United States and other countries have full or partial access to the HRAF Collection of Ethnography. (See the member list at http://www.yale.edu/hraf/collections_body_members.htm for institutions that are active members of the online version, eHRAF World Cultures.)

The HRAF Collection of Ethnography contains mostly primary source materials—mainly published books and articles, but including some unpublished manuscripts and dissertations—on selected cultures or societies representing all major regions of the world. The materials are organized and indexed by a unique method designed for rapid and accurate retrieval of specific data on given cultures and topics. HRAF's system of organization and classification of source material presents information in a manner that significantly increases the usefulness of original source materials. Researchers can use the Collection of Ethnography in four different media: the original paper files, fiche, and on the World Wide Web. Until 1958, the HRAF Collection was produced and distributed as paper files: source materials were manually reproduced on 5" x 8" paper slips called File pages, and then filed by subject (OCM) category and by culture. Wider distribution of the collection was facilitated in 1958 with the development of the HRAF Microfiles Program. Materials from the paper files were processed into microfiche and issued in annual installments to participating institutions; Installment 42 was the last microfiche series issued to members.

In the 1980's, HRAF began developing an electronic publishing program with the intention of distributing the HRAF Collection of Ethnography exclusively through electronic means. The Cross-Cultural CDs were the first result of this effort, providing researchers with ten collections on such topics as old age, marriage, religion, and human sexuality, excerpted from HRAF's 60-Culture Probability Sample Files (PSF). In 1993, the first installment of the full-text HRAF Collection of Ethnography on CD-ROM (eHRAF) was issued to members with the plan of converting the entire 60-Culture PSF,

plus new files covering North American immigrant groups, by the year 1999. Additional installments are added annually. As of April 2008, there were 165 cultures online (<http://ehrafworldcultures.yale.edu>).

Using eHRAF is a relatively straightforward process. Mechanics of use and research techniques are similar in many respects to standard library practices; searching follows the same principles and techniques, such as Boolean logic, that are used for other electronic educational collections.

Organization and Classification

In the paper and fiche versions of the HRAF Collection of Ethnography all documents that contain information about a particular culture are grouped together in a collection for that culture. Each culture collection is identified by a unique alphanumeric code according to the Outline of World Cultures--OWC (Murdock 1983) . In the OWC all the cultures are classified according to geographical regions:

A - Asia

E - Europe

F - Africa

M - Middle East

N - North America

O - Oceania

R - Eurasia (cultures located in the former Soviet Union and Russia)

S - South America

There is one exception to this system—Muslim societies in Africa are classified as being in the Middle East. In its recent literature, HRAF has begun to organize those Muslim cultures under Africa, although they retain the same OWC code.

All the cultures in the paper and microfiche version of the HRAF Collection of Ethnography are grouped into these eight regions. Thus, all the documents pertaining to African cultures are grouped together and their OWC begins with "F." Each of the major regions is then subdivided, usually on a political basis, into sub-regions designated by the addition of a second letter: "FF" designates the country of Nigeria and its component cultural units, while "SC" indicates that the culture described is in the South American country of Colombia. Finally, within each sub-region, more specific units are defined and

assigned a number; these may be country entities, such as "RD01" for Ukraine, or "cultural" units such as "FL12" for the Maasai. Each culture is therefore listed in its regional, political, and cultural context within the Collection.

In eHRAF World Cultures, the OWC number is de-emphasized and cultures are ordered by major geographical regions arranged in alphabetical order: Africa, Asia, Europe, Middle America and the Caribbean, Middle East, North America, Oceania, and South America. The OWC number is listed in the Culture Profile (Browse/Cultures).

A. Selection of Cultures

Several thousand cultures are listed in the OWC, but not all the cultures on the list are included in the HRAF Collection of Ethnography. The cultures in the Collection are selected mainly on the basis of the following criteria:

(a) Maximum cultural diversity—the cultures should represent, as far as possible, the known range and variety of cultural types in terms of language, history, economy, and social organization.

(b) Maximum geographical dispersal—the cultures should be geographically representative of all major world areas and all major ecological settings.

(c) Adequacy of literature—within the scope of the two preceding criteria, the cultures should have a quantitatively and qualitatively adequate literature coverage.

B. Source Materials

Once the decision has been reached to build a collection on a particular culture, extensive bibliographic research is undertaken to identify as thoroughly as possible all of the significant literature on that culture. HRAF also solicits the advice and expertise of specialists. As always, researchers are encouraged to inform HRAF of any salient material which might have escaped notice.

The materials processed for the Collection of Ethnography are largely descriptive rather than theoretical, with the great majority being primary documents resulting from field observation. The ideal document is one which consists of a detailed description of a culture, or of a particular community or region within that culture, written on the basis of prolonged residence among the people documented by a professional social scientist. Many documents which do not meet all the criteria are included in the Collection of Ethnography because they are still important pieces of information—in fact, it is likely that they may be the only sources available for particular time periods, regions, or subjects. Thus the collection for each culture may contain documents written by travelers, missionaries, colonial officials, traders, etc. The Collection of Ethnography provides researchers with a comprehensive picture of life in one or more communities and in one or more time periods.

C. Classification

Every page in each document is indexed and assigned any number of appropriate subject category codes according to the classification scheme in the Outline of Cultural Materials (OCM) (Murdock et al. 2008; online versions); the subject codes are sometimes referred to as OCMs. The OCM consists of 710 subject categories plus a category numbered "000" for unclassified materials. The 710 categories are grouped into seventy-nine major subject divisions, each assigned a three-digit code ranging from 100 (Orientation) to 880 (Adolescence, Adulthood, and Old Age). Within each major subject division, up to nine more specific categories are defined. For example, the (Family (590) division is subdivided into seven more specific subject categories as follows: Residence (591), Household (592), Family Relationships (593), Nuclear Family (594), Polygamy (595), Extended Families (596), and Adoption (597).

Each category in the OCM includes a brief descriptive statement, indicating the range of information which may be classified under that category. Beneath this statement is usually a list of cross-references to other categories under which related information may be classified.

The OCM contains a detailed index which directs the researcher to OCM numbers relevant for their search. The OCM subjects are clearly defined in the OCM, but a few are essential to effective use of the HRAF Collection of Ethnography and bear mentioning here.

Every document page has at least one OCM assigned to it. If there are no pertinent subject categories, "000" indicating non-classified data is applied. In the paper and microfiche, the OCMs are written in roughly where the subject starts. Sometimes an OCM will apply to a particular sentence, although most OCMs apply to at least a section of a paragraph. For eHRAF all OCMs are located at the paragraph level. If five consecutive paragraphs discuss categories 585, 578, and 602, all three OCMs will appear at the beginning of each of the five paragraphs until the subject changes.

What's the Question?

Cross-cultural (worldwide comparative) researchers ask four kinds of questions. The first is descriptive and deals with the prevalence or frequency of a trait: *What percentage of the world's societies practice polygyny? Which is the most important subsistence activity among food collectors — gathering, hunting, or fishing? How common is female infanticide?* A second kind of question considers the causes of a trait or custom. Questions: *Why is polygyny permitted in most societies known to anthropology? Why do women (as opposed to men) do most of the agricultural work in some societies? Why is the extended family the customary form of household in many societies?* The third kind of question explores the consequences or effects of a particular trait or custom. *What are the effects on infant care of high involvement of women in subsistence activities? Does punitive childtraining affect the frequency of warfare?* The fourth question, which is not

significantly different from the second and third, is a relational question. Rather than postulating causes or consequences, a researcher may simply ask how a particular aspect of culture may be associated with some other aspect(s). For example: *Is there an association between most important subsistence activity and level of political complexity?*

Of these four questions, the causal question is the most challenging because it does not completely specify what the researcher needs to do. The descriptive question tells the researcher what to count. The "consequence" and "relational" questions both specify two sets of phenomena that may be related. But the causal question does not tell the researcher where to look for causes. It only specifies what scientists call the *dependent* variable (the thing to be explained).

Think of the causal question as analogous to the format of a detective story. After a murder is committed the detective may know a lot about the crime, but not "whodunit" or why. Finding the solution usually entails hypothesizing about suspects and their possible motives and opportunities, eliminating the implausible possibilities, and concluding who is probably the culprit.

Similarly, in science, the pursuit of causes involves the testing of alternative explanations or theories which purport to say why something is the way it is. The researcher who chooses a causal question needs to identify plausible explanations or theories to test and to decide on a strategy (for collecting and analyzing data) that could falsify or disconfirm explanations. If all theories fail, researchers must come up with new theories. Although these requirements may suggest that the researcher who searches for causes may need to act differently from other researchers, this is really not the case, as we shall see.

The basic strategy for examining relationships in cross-cultural research is the same, whether the relationship involves presumed causes, consequences, or just hypothesized association. To illustrate that strategy, let us turn to an example of a test of a causal explanation.

In the first study we did together (M. Ember and C.R. Ember 1971), our question was: *Why do some societies practice matrilineal residence and others patrilineal residence?* We started where most people start — with explanations found in the literature. One of the most common was the idea that the division of labor based on gender in primary subsistence activities would largely determine residence after marriage (Lippert 1931: 237; Linton 1936: 168-69; Murdock 1949: 203ff.) In other words, female dominance in subsistence should produce matrilocality; male dominance should produce patrilocality. What makes this a causal explanation are the words "determine" and "should produce," which are equivalent to using the word "cause." But, as philosophers of science tell us, causes cannot be directly verified. Even if we can be sure that presumed causes preceded the presumed effects, we cannot rule out the possibility that something else is the real cause.

So how do we test such a causal explanation? The simplest way is to examine a relationship that should be true if the theory is correct, and then make a statistical test to

see if the predicted relationship actually occurs significantly more often than would be expected by chance. In our own study of matrilocal versus patrilocal residence, we derived the following prediction from the "division of labor" theory: if females did relatively more work than males, residence would tend to be matrilocal; if males did relatively more subsistence work, residence would tend to be patrilocal. Notice that although the prediction (or more formally the hypothesis) has almost the same form as the theory we stated above, it differs in a fundamental way — the hypothesis simply predicts an *association* between two variables and says nothing about causality. Still, if two things are causally related, they should be statistically associated.

In our case, when we examined the association between division of labor and residence in a worldwide sample of societies, the predicted association *was not found*. This led us (and later Divale [1974]) to reject the theory that division of labor largely determines residence. After rejecting the "division of labor" explanation (at least as a major cause) we went on to test other explanations. Eventually we ended up developing a new theory that internal warfare (warfare within the society) would produce patrilocal residence, and purely external warfare (particularly if women do a great deal of subsistence work) would produce matrilocal residence. Note that division of labor remains a partial cause in our explanations. Note too that even if a predicted relationship is supported, it may still be open to different interpretations. Indeed, Divale (1974) offers a very different explanation for the obtained relationship between type of warfare and residence.

The study we just discussed illustrates the fundamental assumption of worldwide cross-cultural (or holocultural) research; if a theory has merit, the presumed causes and effect should generally be associated synchronically (see J. W. M. Whiting 1954; K. F. Otterbein 1969; R. Naroll, Michik, and F. Naroll 1976). A synchronic association is one that involves data (for each sample case) from more or less the same point in time, as if we were examining a large number of "ethnographic snapshots," each one capturing a society at a single point in time. The cross-cultural method therefore provides a way of eliminating theories that have no predictive value. Theories that postulate causes, consequences, or relationships are tested in the same way — that is, by looking to see if predicted associations obtain.

Sampling

Cross-cultural researchers must decide what societies to examine. No one can examine all cultures; even if one could, the labor and time costs involved would not justify doing so. The most important operating principles in a scientific test of a hypothesis are: 1) to choose a sample that is representative of some universe of societies the researcher wants to generalize the results to; and 2) to use a large enough sample such that the results are likely to be true for the larger universe of cases. As yet, there is no complete list of the world's cultures to sample from, so researchers cannot do what is ideal, which is to sample randomly from a complete list. Instead, cross-cultural researchers usually sample from one of a number of published cross-cultural samples. (These lists can be thought of as "sampling frames.") The most commonly used are (from larger to smaller): the

"summary" Ethnographic Atlas (Murdock 1967); the Atlas of World Cultures (Murdock 1981); the Human Relations Area Files (HRAF) Collection of Ethnography (annually distributed by the Human Relations Area Files); the Standard Cross-Cultural Sample (Murdock and White 1969); and the HRAF Probability Sample Files (Naroll 1967; Lagacé 1979), which is a subset of the entire HRAF Collection of Ethnography. While none of these samples is perfect, the important point about all of these lists is that they were not designed to support any researcher's pet idea or theory. In contrast, a set of cases chosen from a researcher's own personal library would be scientifically suspect.

Why use the HRAF Collection of Ethnography?

Most of the samples mentioned above contain bibliography (or pointers to bibliography) and at least some coded information on traits of interest to a variety of researchers. The HRAF Collection of Ethnography is different in that it contains no precoded data, but full texts indexed by subject matter and grouped by culture for the rapid retrieval of particular kinds of information. If you want to read about a particular aspect of culture and make your own coding decisions on a sample of societies, the HRAF collection is ideal because you do not have to collect all the books and articles on each of the cultures and then search for a particular subject through all the texts. HRAF's subject index, the *Outline of Cultural Materials* (Murdock et al., 2008; online version at <http://www.yale.edu/hraf/outline.htm>), can be used to identify particular subject categories to look at to find the information of interest to you. If you are working from the print version of the *Outline of Cultural Materials* (OCM for short), the easiest way to find a subject category is by using the extensive index in the back of the OCM. This index will point you toward a number of possible numbered subject categories. When you read about these subjects in more detail, you will find out if the subject categories are appropriate. The OCM system is mostly hierarchical in that the first two digits usually reflect the major subject category. So, for example, all the "59s" (591-597) refer to the major subject labelled "Family." The last digit is a subcategory (e.g., 596 is "extended families"). If you are working on eHRAF, the A-Z Index can be found under Browse/Subjects. There is also a list of OCMs organized by Major Subject and a list in OCM or numerical order. In addition, in Lookup Search, you can type in a word in the Lookup box (e.g., diet) to see what subjects are suggested. Often researchers will need to search for more than one subject category to ensure that they will find what they are looking for. Keep in mind that not all ethnographers discuss all topics, so some categories will be empty for some cultures.

Sampling within the HRAF Collection

It is rarely necessary to use the entire HRAF collection for comparative studies. The only reason it might be necessary to examine all the cases is if some trait or custom occurs rarely or is only rarely described. In that case, researchers might have to scan all the societies to find enough cases of a particular type. Examples of relatively rare traits are age-set systems, cannibalism, and woman-woman marriages.

Researchers use a variety of strategies to sample the collection. If researchers want to use some already coded data (coded previously by themselves or other researchers) for their study, they usually choose to limit themselves to those sample cases for which the desired precoded data are available. Some researchers find that the HRAF Collection of Ethnography speeds up their data retrieval so much that they use it for as many cultures as they can and then look up books and articles for the remaining cultures. Others choose the overlap between the HRAF sample and another sample. The important thing to keep in mind in using information from two different samples is that the information in the different samples may pertain to different time periods and different communities. Since cultures change over time and vary from community to community, it is extremely important to make sure that the same-named cases in the overlapping samples actually are the same in time and place. Otherwise, the researcher is introducing error.

For example, suppose one is examining the possible relationship between male mortality rates in warfare and frequency of polygyny (see M. Ember 1974 for a test of the hypothesis that high male mortality in warfare should be associated with appreciable polygyny). For information on a given society with regard to male mortality in warfare, one would look in categories Mortality (165), Instigation of War (721), and Aftermath of Combat (727) and might find ethnographic material from 1890 indicating that many men died in warfare. For information on extent of polygyny (category 595) the researcher may find the best information to be from 1950. If you used these two pieces of information (one from 1890 and the other from 1950) you might very well have a case that looks like it does not support your hypothesis. This would be an error if the society had appreciable polygyny in 1890 but did not have much in 1950. Pacification by external authorities might have eliminated war, thus evening out the sex-ratio, and thereby lessening polygyny. In this instance, the data from either 1890 or 1950 might support the hypothesis (high male mortality/high polygyny or low male mortality/low polygyny), but mixing data from different time periods would have created an "error."

The *Computerized Cross-Cultural Concordance* (C. R. Ember 1992) was developed to help researchers see if times and places match across different samples. One of the most useful aspects of this concordance is that it gives the researcher the appropriate sources to look at in the HRAF Collection of Ethnography if she or he wants to match cases in another sample. HRAF processes an extensive set of sources for each society included in the archive. Usually there are multiple time and place foci, so it is important that a researcher attend to the need to choose exactly the right focus. Researchers who want to use data already available from other samples commonly use the Ethnographic Atlas summary (Murdock 1967) or the Standard Cross-Cultural Sample (Murdock and White 1969). The codes for the Standard Sample (published by many different authors) have appeared in two journals (*Ethnology and Cross-Cultural Research* [formerly *Behavior Science Research*]). Many have been reprinted in Barry and Schegel's (1980) *Cross-Cultural Codes and Samples*, and have been put into computer format for the *World Cultures* electronic journal.

If the researcher does not need information from outside HRAF, sampling from the HRAF Collection of Ethnography can proceed differently. The HRAF Collection itself

can be used as a sampling frame (a list to sample from) and researchers can randomly choose cases from that list by using a table of random numbers. (Your library might not have the complete collection—check with your librarian for the cultures found in your library.) The subset of HRAF known as the Probability Sample Files is a special kind of random sample called a stratified random sample. The world was divided into 60 culture areas (strata) and one case from each area was randomly chosen from a list of societies that met certain criteria (such as whether one of the ethnographers stayed for more than a year). See Naroll (1967) for the rules used in selecting the sample cases. In the *complete collection* list at http://www.yale.edu/hraf/collections_body_ethnoallformats.htm, look for the column labeled "PSF."

What is a large enough sample? Statisticians have worked out formulas for calculating the size of the representative (random) sample that is needed to obtain a significant result (one likely to be true). The samples needed are usually much smaller than you might imagine. If a relationship is strong, a random sample of 20-30 is sufficient. (Weak associations can be significant only in large samples.) By using the random sampling strategy, researchers can always add cases randomly to increase sample size. Random sampling also enables researchers to estimate whether a phenomenon of interest occurs frequently enough to be studied and whether the measures adopted are usable on the available data.

While most people assume that "bigger is better," bigger samples require much more time and effort and expense. And they may not yield much more information or accuracy than a smaller random sample. Political opinion polls are a case in point. Samples of a few hundred to a few thousand people in the entire United States can often yield quite accurate predictions of elections.

Measures

The concepts in a predicted association (hypothesis) can be fairly specific, such as whether or not a culture has a ceremony for naming a newborn child, or they may be quite abstract, such as whether the community is harmonious. But whether the concept is fairly specific or not, no concept is ever measured directly. This is true in physical as well as in the social sciences. We are so used to a thermometer measuring heat that we may forget that heat is an abstract concept that refers to the energy generated when molecules are moving. A thermometer reflects the principle that as molecules move more, a substance in a confined space (alcohol, mercury) will expand. We do not see heat; we see the movement of the substance in the confined space.

The three most important principles in designing a measure are: 1) try to be as specific as possible in deciding how to measure the concept; 2) try to measure the concept as directly as possible; and 3) if possible, try to measure the concept in a number of different ways. The first principle recognizes that science depends upon replication; it is essential for other researchers to try to duplicate the findings of previous researchers, so researchers have to be quite explicit about what they mean and exactly how they measure the concept. The second recognizes that although all measurement is indirect, some measures

are more direct than others. If you want to know how "rainy" an area is, you could count the number of days that it rains during December, but a better measure would be the number of rainy days on average over a number of years. The third principle is that since no measure exactly measures what it is supposed to measure, it is better, if possible, to have more than one way to tap the concept of interest.

Measures have to be specified for each variable in the hypothesis. Devising a measure involves at least four steps; 1) theoretically defining the variable of interest (in words or mathematically); 2) operationally defining the variable, which means spelling out the "scale" that the researcher has devised for measuring it; 3) telling the coder where to find the required information (in the case of research using the HRAF Collection of Ethnography, this means specifying which subject categories (OCMs) the coder should look at; in the electronic version one can also specify what words or combinations of words to look for and 4) pre-testing the measure to see if it can be applied generally to most cases. Designing a measure requires some trial-and-error. If the scale is too confusing or too hard to apply (because the required information is lacking), the measure needs to be rethought.

To illustrate the procedure, let us consider a variable that seems rather straightforward—the degree to which a society has extended family households. Although this concept may appear straightforward, it still needs to be defined. The researcher needs to state what an extended family means, what a household means, and how she or he will decide the "degree" to which a sample society has extended family households. The first thing would be to decide on what is meant by an "extended family." The researcher may choose to define a family as a social and economic unit consisting minimally of at least one or more parents and their children; an extended family as consisting of two or more constituent families united by a blood tie; and an extended family household as an extended family living co-residentially—in one house, neighboring apartments, or in a separate compound. Having defined the concepts, the researcher must now specify how to measure the degree to which a society has extended family households.

Definitions are not so hard to arrive at. What requires work is evaluating whether an operational definition is useful or easily applied. For example, suppose by "degree" (of extended familiness) we operationally mean the percentage of households in a focal community that contain extended families. The range of possible scale scores is from 0 to 100 percent. Suppose further that we instruct our coders to rate a case only if the ethnographer specifies a percentage or we can calculate a percentage from a household census. If we are also using information from another study, we tell our coders to look at the Household (592) and Extended Family (596) for the same community specified in the other study and at the time specified in the other study (same time). (If we are not taking data from another study, we can ask our coders to pick a community and a time which is most thoroughly described with regard to household form.) If we did a pretest, we would find out that very few ethnographers tell us the percentage of extended family households. Rather they usually say things like, "Extended family households are the norm." Or, "Extended families are typical, but younger people are beginning to live in independent households." So our operational definition of percentage of extended family

households, although perfectly worthy, may not be that useful if we cannot find enough societies with household censuses.

What can we do? There are three choices. We can stick to our insistence on the best measure and study only those societies for which a percentage is given. We may have to expand our search (enlarge our sample) to find enough cases that have such precise information. Or, we can redesign our measure to incorporate descriptions merely in words (no census material is available). Or, we can choose not to do the study because we can't measure the concept exactly how we want to. Faced with these three choices, most cross-cultural researchers opt to redesign the measure so as to incorporate word descriptions. Word descriptions do convey information about degree, but not as precisely. If an ethnographer says "extended family households are typical," we do not know if that means 50% or 100%, but we are very confident it does not mean 0-40%. And we can be fairly sure it does not mean 40-49%. If the relative frequency of extended families is related to something else, we should be able to see the relationship whether we measure in percentages or words.

A newly designed measure might read something like this: Code extended family households as

4) *Very high* in frequency if the ethnographer describes this type of household as the norm or typical in the absence of any indication of another common type of household. Phrases like "almost all households are extended" are clear indicators. Do not use discussions of the "ideal" household to measure relative frequency, unless there are indications that the ideal is also practiced. If there is a developmental cycle, such as the household splitting up when the third generation reaches a certain age, do not use this category. Use category #3 if the extended family household remains together for a substantial portion of the life-cycle and #2 if the household remains together briefly.

3) *Moderately high* in frequency if the ethnographer describes another fairly frequent household pattern but indicates that extended family households are still the most common.

2) *Moderately low* in frequency if the ethnographer describes extended family households as alternative or a second choice (another form of household is said to be typical).

1) *Infrequent or Rare* if another form of household is the only form of household mentioned and if the extended family form is mentioned as absent or an unusual choice. Do not infer absence of extended families from the absence of any discussion of family and household type.

don't know if there is no information in the appropriate subject categories, or there is contradictory information for the same time and place from different sources.

The next step is to pre-test this measure. It may turn out that four distinctions are too difficult to apply, so a researcher might want to collapse the scale a little. If we decide to use the scale described above, what do we do when we do get numbers or percentages from the ethnographers for some cases? Most of the time, we can fit those numbers into the word scale. So, for instance, if 70% of the households have extended families, and 30% are independent, we would choose scale position 3. But we might decide to use two scales: a precise one based on numerical measurement (percentages), the second a vaguer one based on words (C. Ember et al. 1991 recommend that we use both types of scale when we can). The advantage of using two scales is that the more precise (quantitative scale) should be more strongly related to other variables than the less precise scale, which result would increase confidence in the relationships found.

Measuring a concept like the degree to which a society has extended families may not be easy. But it is not that difficult either, because ethnographers usually attend to basic economic, social, and political features of a society. We can think of these things as "standard cultural observables." Of course, there are concepts which are much more difficult to operationalize using ethnographic data, because ethnographers do not conventionally attend to these subjects. For instance, few ethnographies contain information that would allow construction of an indicator of rainfall variability, pH of the soil, or number of minutes per day adults spend in housework. For these types of information, researchers may decide to alter their operational definitions to make use of the data that are available. A better research strategy may be to use other kinds of data outside of the HRAF Collection of Ethnography. Some libraries have worldwide climate records. This information can often be linked to ethnography by looking up the nearest weather station (in subject category Research and Development, 654) or longitude and latitude in subject category Location (131) of the society.

Concepts may be difficult to operationalize for other reasons. They may be quite abstract, like the concepts of community solidarity or the relative status of women. These two are not only abstract, but they deal with information which is not usually discussed in conventional ethnographic topics. Information relevant to status might be found under discussions of kin group decisions, political decision-making, relationships of people within the household, sexual rights and obligations, how marriages are arranged, etc.

Research by Martin Whyte (1978b) suggests that it is preferable to avoid rating very abstract variables such as "the status of women." Rather researchers should probably confine ratings to more specific variables, as Whyte himself did. Whyte chose 52 very specific variables to assess the status of women. These variables included the degree to which women had political roles, the importance of female gods, how easily women could get divorced, etc. Whyte found that the various aspects of status did not relate to each other. He concluded that if a researcher wants to discuss status it would be preferable to discuss at least 10 different (and independent) dimensions of status. Furthermore, when he tested for the possible bias in reporting by male versus female ethnographers (Whyte 1978a), he found that whatever bias may exist is more likely to be found in the reporting of more abstract (versus more specific) matters. This suggests that codes should be designed to tap very specific aspects of a phenomenon.

Researchers can always use a variety of scaling procedures to make specific measures into combined or more general measures, as many have done to measure degree of cultural complexity (combining ratings of specific features such as type of subsistence, average size of communities, level of political integration, etc.)

Analysis of Results

When the researcher has measured the variables of interest for all sample cases, he or she is ready to see if the predicted relationship actually exists in the data. After all, there are likely to be exceptions to the predicted relationship. Do the exceptions invalidate the prediction? How many exceptions would compel a rejection of the hypothesis? It is precisely here that cross-cultural researchers usually resort to statistical tests of significance.

Statisticians have devised various tests that tell us how "perfect" a result has to be for us to believe that there is probably an association between the variables of interest, that one variable generally predicts the other. Essentially, every statistical result is evaluated in the same objective way. The question is asked: What is the chance that this result is purely accidental, that there is really no association at all between the two variables? Although some of the mathematical ways of answering this question are rather complicated, the answer always involves a probability value (or p-value), the likelihood that the observed result or a stronger one could have occurred by chance. So, if a result has a p-value of less than .01, this indicates that there is less than one chance in one hundred that the relationship observed is purely accidental. A p-value of less than .01 is a fairly low probability; most social scientists conventionally agree to call any result with a p-value of .05 or less (five or fewer chances in one hundred) a *statistically significant* or probably true result.

In a study we did with Burton Pasternak on extended family households (Pasternak, C.R. Ember and M. Ember 1976), we tested the hypothesis that incompatibility of activity requirements would generally explain why people may choose to live in extended family households. By incompatibility of activity requirements we meant that an adult in the household was required to perform two activities in different places at the same time. A common example for women is childtending and agricultural work in the fields. An example for men is working away from home for wages and having to plow the fields. If the household includes two or more families, i.e., if there is an extended family household, there will likely be two adults of each gender to perform the required tasks. We decided to read and code ethnography to measure incompatibility requirements first, before we knew what the household form was, and then we subsequently looked up previously published coded data on the presence or absence of extended family households. We decided not to code both variables (incompatibility of activity requirements and extended family households) ourselves because we did not want our hypothesis to influence our judgments. The sample investigated was chosen by randomly sampling 60 cultures from the overlap between the HRAF Collection of Ethnography and

the *Ethnographic Atlas* (Murdock 1967). Even though we were only able to code 23 of the sample societies, the statistical test of the relationship between incompatibility of activity requirements and extended family households was statistically significant. The p value was .003, which meant that the result was likely to occur by chance just 3 out of 1000 times. We were able to predict 11 out of 13 of the societies with extended family households and 8 of the 10 of the societies with independent family households.

Why should a probably true relationship have any exceptions? If a theory or hypothesis is really correct, one would presume that all the cases fit. There are many reasons that one cannot ever expect a perfect result. First, even if a theory is correct about a major cause of what one is attempting to explain, there may still be other causes that have not been investigated. Exceptions to the predicted relationship might also occur because of what has been called "cultural lag." Cultural lag occurs when change in one aspect of culture takes time to produce change in another aspect. A sample society might be an exception to the predicted relationship, but it might fit the theory if the variables could be measured for a later time period. Measurement inaccuracy is another source of exceptions, because measurement error is usually random error and random error usually weakens statistical relationships. For example, if some cases in a straight-line relationship are inaccurately measured (either too high or too low) on even just one variable, those cases will not be located on the line of the relationship.

In addition to its statistical significance, a cross-cultural relationship should also be evaluated with regard to its strength, or the degree to which the dependent variable is predicted statistically. After all, the goal in research is to find strong predictors, not just statistically significant ones.

If confidence in an explanation is required, a single cross-cultural test is not enough. Replications by other researchers using other samples, tests against alternative explanations, and tests using other research strategies are also needed. This may seem tiresome, but good research always gives a cherished theory many chances to fail.

More Advanced Reading

For more advanced treatments of these topics, the reader is urged to peruse the articles in "Cross-Cultural and Comparative Research: Theory and Method. Special issue," 1991. *Behavior Science Research* 25:1-270; and Carol R. Ember and Melvin Ember, *Cross-Cultural Research Methods*, Walnut Creek, CA: AltaMira Press, 2001.

References

Barry, Herbert III, and Alice Schlegel. 1980. *Cross-Cultural Samples and Codes*. Pittsburgh: University of Pittsburgh Press.

Divale, William T. 1974. "Migration, External Warfare, and Matrilocal Residence." *Behavior Science Research* 9:75-133.

Ember, Carol R. with the assistance of Hugh Page, Jr., Timothy O'Leary, and M. Marlene Martin. 1992. *Computerized Concordance of Cross-cultural Samples*. New Haven: Human Relations Area Files.

Ember, Carol R., Marc Howard Ross, Michael Burton, and Candice Bradley. 1991. "Problems of Measurement in Cross-Cultural Research Using Secondary Data." *Behavior Science Research* 25:187-216.

Ember, Melvin. 1974. "Warfare, Sex Ratio, and Polygyny." *Ethnology* 13:197-206. Reprinted with afterthoughts in *Marriage, Family, and Kinship: Comparative Studies of Social Organization*. Melvin Ember and Carol R. Ember. 1983. New Haven: HRAF Press, pp. 109-124.

Ember, Melvin, and Carol R. Ember. 1971. "The Conditions Favoring Matrilocal Versus Patrilocal Residence." *American Anthropologist* 73:571-94. Reprinted with afterthoughts in *Marriage, Family, and Kinship: Comparative Studies of Social Organization*. Melvin Ember and Carol R. Ember. 1983. New Haven: HRAF Press, pp. 151-198.

Lagace, Robert O. 1979. The HRAF Probability Sample: Retrospect and Prospect. *Behavior Science Research*, 14:211-229.

Linton, Ralph. 1936. *The Study of Man*. New York: Appleton-Century.

Lippert, Julius. 1931. *The Evolution of Culture*. George P. Murdock, trans. and ed. New York: Macmillan.

Murdock, George P. 1949. *Social Structure*. New York: Macmillan.

Murdock, George P. 1967. *Ethnographic Atlas*. University of Pittsburgh Press. Also *Ethnology* 6:109-236.

Murdock, George P. 1981. *Atlas of World Cultures*. Pittsburgh: University of Pittsburgh Press.

Murdock, George P. 1983. *Outline of World Cultures*, 6th ed. New Haven: Human Relations Area Files.

Murdock, George P., Clellan S. Ford, Alfred E. Hudson, Raymond Kennedy, Leo W. Simmons, John W. M. Whiting. 6th revised edition with modifications 2008. *Outline of Cultural Materials*. New Haven: Human Relations Area Files.

Murdock, George P., and Douglas R. White. 1969. "Standard Cross-Cultural Sample." *Ethnology* 8:329-369.

Naroll, Raoul. 1967. "The Proposed HRAF Probability Sample." *Behavior Science Notes* 2:70-80.

Naroll, Raoul, Gary Michik, and Frada Naroll. 1976. *Worldwide Theory Testing*. New Haven: Human Relations Area Files.

Otterbein, Keith F. 1969. "Basic Steps in Conducting a Cross-Cultural Study." *Behavior Science Notes* 4:221-236.

Pasternak, Burton, Carol R. Ember, and Melvin Ember. 1976. On the Conditions Favoring Extended Family Households. *Journal of Anthropological Research* 32: 109-23. Reprinted with afterthoughts in *Marriage, Family, and Kinship: Comparative Studies of Social Organization*. Melvin Ember and Carol R. Ember, 1983. New Haven: HRAF Press, pp. 109-124.

Whiting, John W. M. (1954) "The Cross-Cultural Method." In *Handbook of Social Psychology*, Vol. 1. Gardner Lindzey, ed. Cambridge, Mass.: Addison-Wesley, pp. 523-31. Reprinted in *Readings in Cross-Cultural Methodology*. Frank W. Moore, ed. New Haven: HRAF Press, pp. 287-300.

Whyte, Martin K. 1978a. "Cross-Cultural Studies of Women and the Male Bias Problem." *Behavior Science Research* 13:65-80.

Whyte, Martin K. 1978b. *The Status of Women in Pre-industrial Societies*. Princeton: Princeton University Press.