methods as somewhat independent of the kind of data produced. This book is organized by method and the different "types" of data that can be obtained from each method are discussed in each chapter. Data types that are referred to include: similarity data, ordered data, and test performance data. Similarity data consist of informants' judged similarity estimates among study items. Ordered data consist of an ordinal ranking (or rating) of items on a single conceptual scale. Performance data refer to responses that can be graded as "correct" or "incorrect."

Plan of the Book

In the following chapters we describe the use of various formats for the collection of data in structured interviewing and paper-and-pencil questionnaires. In each chapter we discuss the method in general and then give detailed step-by-step directions for its use. We also give examples of the use of each of the methods as well as summarize their strengths and weaknesses.

We begin, in Chapter 2, by describing the procedures for finding how respondents define the boundaries of a given subject matter or domain of interest. Chapters 3 and 4 discuss the use of the pile sort as a useful format for obtaining judged similarity among a large number of objects. The pile sort is a format in which the subject is asked to sort items into piles so that objects within piles are more similar to each other than to objects in the other piles. Chapter 5 presents the method of triads. In the triad task the informant is presented with three items at a time and typically picks the one most different from the other two. Triads may be used to collect either similarity or ordered data. Rating scales, one of the most varied and widely used methods in the social sciences, are discussed in Chapter 6. Rating scales may be used to collect ordered data or to collect similarity data on pairs of items. Rank order tasks are described in Chapter 7. Chapter 8 provides a complete description and explanation of balanced incomplete block designs that may be used to shorten and simplify triad, paired comparison, and ranking tasks. We devote Chapter 9 to a discussion of the use of sentence frames as a systematic interviewing technique that is especially useful where the items in a domain are characterized by a variety of features. In Chapter 10 we discuss a variety of other data collection formats. These include true-false questions, multiple-choice questions, fill-in-the-blank questions, matching designs, direct estimation, and "pick n" items from an array. In the last two chapters we discuss reliability and validity. We illustrate various ways in which confidence in the results from interviewing can be increased by using such techniques as replication and the convergence of results from different methods.

Although this book is addressed primarily to social scientists, it is appropriate for anyone who wants to study attitudes and beliefs. While a political scientist may want to use a technique to study how particular political candidates are perceived, a market researcher may use the same techniques to study preferences for different brands of cigarettes. Similarly, one sociologist may use structured interviewing techniques to study perceptions of occupational prestige, while another may study social networks. Anthropologists working with nonliterates may be particularly interested in the oral data collection methods. While we cannot review all the possible applications, we would like to note that the interviewing and data collection tasks contained in this volume are as appropriate for use in such exotic settings as the highlands of New Guinea as they are in the corporate offices on Wall Street. It is our expectation that the book will provide the researcher with a larger choice of data collection techniques than has been heretofore available in a single place.

2. DEFINING A DOMAIN AND FREE LISTING

The first step in any study is to obtain a clear understanding of the definition and boundaries of what is being studied. Since this book is about interviewing there is an implicit assumption that the researcher is interested in what the respondents think about "something." For convenience we call the "something" a semantic or cultural domain. The semantic or cultural domain is simply the subject matter of interest, a set of related items. Examples of domains that have been studied include color terms, kinship terms, diseases, plant terms, animal terms, airplane piloting errors, kinds of pain, and characteristics of infant feeding methods. The concept of a domain is a very general one and may include almost any coherently defined subject matter.

A domain may be defined as an organized set of words, concepts, or sentences, all on the same level of contrast, that jointly refer to a single conceptual sphere. The items in a domain derive their meanings, in part, from their position in a mutually interdependent system reflecting the way in which a given language or culture classifies the relevant conceptual sphere. For example, the concept of "shape," may have category members such as "round," "square," "rectangular," and so on. Each of these is a kind of shape and each says something different about shape. We refer to "shape" as the generic name of the category and the

words "round" and the like as items or objects in the domain.

The overall success of any study depends in part on giving careful attention to the definition of the domain as the first step of the research. Generally the domain should be defined by the informants, in their language, and not by the investigator. It is easy to assume falsely that the investigator knows the domain and can therefore define what items belong in the domain and what items do not belong in the domain. If researchers want to study beliefs about discipline in different ethnic groups, it is only after disciplinary actions and punishments are explicitly listed that it is appropriate to worry about the format and design of the interview instrument. The immediate problem is to specify which disciplinary actions and punishments should be included in the interview.

There are many ways to compile a list of items to define the domain of study items. On rare occasions there may be an absolute definition. The alphabet, the Morse code, the names of the months, the days of the week, presidents of the United States, and the states of the United States are all examples of a priori defined domains. Usually, however, the investigator does not know the boundaries of the domain and needs some sort of elicitation procedure to ensure that the investigator's definition of the domain corresponds to that of the informants.

The most useful general technique for isolating and defining a domain is the free listing task. The use of free listing as an elicitation technique has several useful applications. Perhaps its most important use is to ensure that one is dealing with culturally relevant items and to delineate the boundaries of a semantic or cultural domain. The free listing task can also be used to study or make inferences about informants' cognitive structure from the order of recall, the frequency of recall, and the use of modifiers (Romney and D'Andrade, 1964).

Requests such as "Name all the x's that you know" or "What kinds of x's are there" can elicit domain items. For example, in a comparative study of disease concepts a free list task was used by Weller (1984a: 342):

> To ensure that culturally relevant items would be used, 20 women in each country were asked to name all the illnesses they could think of and to describe each. Using the most frequently mentioned items, a domain of 29 English and 27 Spanish disease terms was selected . . . and used in the subsequent data collection tasks.

There are several things that can be observed and inferred from such lists. First, some items are more "salient," "better known," "important," or "familiar" than other items, and such items occur earlier or higher up in an individual's list than those that lack such characteristics. Second,

there is usually a great range in the number of people that mentioned each item. Thus we can think of two different indices of "saliency." The first is the position of an item on a list and the second is the proportion of the lists on which the item appears. These two indices tend to be highly correlated. For example, Romney and D'Andrade (1964: 155) report a correlation of .83 between these two indices for 105 free lists of "kinds of relatives." These indicators of salience also are closely related to the frequency of usage of each item in ordinary language as measured in the Thorndike-Lorge (1944) word list. Free lists, when drawn from fairly large samples (such as 100 or more) provide similar information.

Many domains contain far more elements than are practical to include in a single interview instrument. This adds the complication of sampling items from the domain for inclusion in the interview. The free listing task allows us to find the most salient items with minimal effort. Items that do not appear on the lists are probably not as common or salient as items that do appear. The list is not to be taken as definitive and complete. If we interviewed additional informants we might increase the number of items. However, as the number of informants increases, the list becomes stable and the order of items tends not to change as few new items are added by each new person.

How to Do It

We would like to emphasize the importance of defining the domain of items for study as a first step prior to the use of other systematic data collection techniques. Domain definition, whether done with a free list or variation on the free list, is extremely important and assures that the domain is defined by the informants in their language. Without free listing the items may reflect the ideas of the researcher rather than the informants. This step is so important that we suggest that it not be omitted or delegated.

What to ask. The first step in obtaining items for study is to decide upon a domain, for example, illnesses, cars, colors, kin terms. The next step is to decide on how to ask informants to list the items. Getting the right question can be difficult, although on occasion it is easy.

To find out if the question is productive it is necessary to pretest the wording. By trying the wording on four or five people it is generally possible to tell if the question is appropriate. When one has the right wording, the interview flows smoothly, the informant understands the question and produces reasonable items at reasonable frequencies. Sometimes informants respond with too few items. In such cases, the question wording might need to be changed. Sometimes informants are not used to responding to questions with lists of items and probes may

be useful. For example, "You said that _____ and _____ are ways to treat malaria. What other ways are there for treating malaria?" Do not ask *if* there are any other treatments, ask *what* other treatments there are. Open ended questions that can be answered with a yes or no are too often answered with a no.

If the problem lies with the question you are asking, you need to try asking the question in new ways. It is possible that what the researcher thinks is a domain is not a domain for the informants. Rephrase the question and talk with the informants until a question is found that makes sense. In some cases it helps to elicit statements rather than words. Instead of asking "List kinds of health care" it is possible to ask "What did you like about the health care you received? What were the things that you did not like? What things do you wish you had gotten?" Responses should be recorded verbatim.

It is important to encourage the informants to clarify their responses. Do not assume that you know what the respondent means, ask for further explanation. For example, when Weller and Dungy (1986) asked about the advantages of breast feeding, one woman responded that it was convenient. That statement sounds clear. We think we know what it means, but what exactly is meant by convenient? As it turns out, "convenient" means something different to breast and bottle feeders. To the breast feeders it means that they can feed their baby without the hassle of preparing bottles. To bottle feeders it means that they can feed their baby anywhere without embarrassment.

Variations on what to ask. The choice of which questions to use in eliciting the free lists is sometimes fairly straightforward. However, in many cases there is no obvious single question to elicit domain items and variations in approach are useful. Sometimes just asking for a free listing does not elicit very long lists. In some of these situations a modified type of free listing task can be used to elicit descriptive phrases for further study. For example, Romney et al. (1979), in attempting to identify and measure attributes of success and failure, asked respondents to list five friends and characterize the ways in which each was successful and the ways in which each might be considered a failure. Respondents were encouraged to use their own definitions of success and failure.

Combining more than one kind of list is another possibility. For example, in seeking the reasons why women choose to breast or bottle feed, Weller and Dungy (1986) asked each woman to list (1) the advantages of breastfeeding, (2) the disadvantages of breastfeeding, (3) the advantages of bottlefeeding, and (4) the disadvantages of bottle-feeding. All four lists were combined to form the domain of "the characteristics of feeding methods."

Another kind of variation on free listing is the use of "contrasting

questions." Young (1980) wanted to elicit reasons given by Mexican Indian villagers for choosing one source of medical care over another. He first obtained a list of all sources of health care in the village. He then presented the informants with *pairs* of alternative sources of health care. The format of his questions was "Why/when would you go to _____ instead of _____?" In this way he was able to construct a list of reasons for choosing one medical care over another. Krackhardt and Kilduff (1987) presented items in sets of three, asked informants to choose the most similar pair, and asked informants to explain why those items were similar.

Sometimes it is useful to work with two related lists, for example, to elicit disease terms and then elicit symptoms and causes of each disease. This technique was used by Weller et al. (1987). In initial interviews, adolescents were asked to list things they could "do wrong." Probes included asking about things that their friends or anyone else had done wrong. Then after listing all the things possible, each person was asked, for each of the misbehaviors, to name all of the things that their parents or some other adult might do in response as punishment. Thus for each misbehavior an additional list was obtained, namely, a list of possible parental punishments.

Free lists also can be combined with information from other sources. For example, in the study of "discipline" discussed above the authors wanted to include some "abusive" forms of punishments to see if there were differences between Anglo and Mexican American perceptions of the acceptability of those punishments. Not surprisingly, the free lists did not include extremely harsh parental responses. To ensure that the final list did include some harsh punishments, the researchers obtained a list of the most frequently reported cases of physical abuse seen in the Pediatric Emergency Room at the University Hospital. They added these to round out the final list.

Berlin and Romney (1964), in attempting to find all the numeral classifiers in Tzeltal, used a clever way of eliciting all members of a semantic domain. By depending upon informant recall and other informal methods they isolated a few dozen numerical classifiers. Since all of them had the form CVC (Consonant, Vowel, Consonant) they were then able to generate, with the help of a computer, all possible combinations of Tzeltal consonants and vowels of the appropriate form. There were 4,410 such possible combinations. By presenting each of the forms independently to two informants they were able to isolate 557 numeral classifiers: "This number is considerably greater than any other inventory thus far published. Without the systematic eliciting procedures described above, less than a tenth of the classifiers in Tzeltal would have been discovered, for their actual occurrence in textual material is

infrequent. The list may not be complete, but it approaches the full inventory" (1964: 81).

Group interviews (nominal group process, Delphi groups, focused group interviews) are also sometimes used to establish a list of items of interest. Nominal and Delphi groups usually begin with a group leader requesting that each participant write down all of their feelings, reasons, or concerns regarding the topic of discussion. For example, if the group is assembled to discuss student evaluations, then they might be asked to list all of the ways in which a student might be properly evaluated. If the group has been assembled by the vice president of a major car manufacturing company, they might be asked to list all the features that the next new model car should have. The group process begins as each person is asked to contribute one or two items to a master "group" list. When the master list contains most or all of the items from the initial lists, then participants are asked to choose the most important items. The group process continues with individuals discussing the listed options with the goal of obtaining a single consensual list.

Data can be used from nominal or Delphi groups if the following things are kept in mind. First, the sample size is not the number of participants in a group, rather it is the number of groups. If the initial lists are collected prior to discussion, the number of respondents would be equal to the number of participants in the group. After discussion and interaction, however, the data from individual lists are no longer independent and the group, in effect, becomes an individual in terms of generating items.

Minimum number of informants needed. Usually with a coherent domain, 20 to 30 informants are sufficient. Larger or smaller numbers of informants are necessary depending upon the amount of agreement in the responses. Of course 10 informants at this stage are better than no informants. If one keeps track of the frequencies in a sequential way it is possible to tell when stability in order is reached and use this as a guide for how many informants are necessary. To illustrate, Weller (1980) cumulatively tabulated lists and examined the order and frequency of items. Since the relative order of frequencies of items (see Tabulation below) did not change, and few new items were added by increasing the sample size from 10 to 15 and from 15 to 20, a sample size of 20 was assumed to be adequate.

Tabulation

Responses are tabulated by counting the number of respondents that mentioned each item. Items are then ordered in terms of frequency of response as in Tables 2.1 and 2.2. The tabulated list of items is a

frequency distribution of the number of respondents that mentioned each item. Frequencies can be re-expressed as percentages where desirable. Frequencies or percentages may be used as estimates of how salient or important each item is to the sample of informants.

It is important to note that when multiple questions are asked of a single informant, responses should be tabulated by the number of persons mentioning each item and not the total number of times that an item is mentioned. With multiple-related questions an informant may give the same response more than once. When this occurs it is important to count that response only once for that informant. The frequency distribution of the tabulated items should always reflect the number of persons that mentioned the item.

The final tabulated list of items can sometimes be diagnostic of whether or not the researcher asked a meaningful question. If items are arranged in order of their frequency of mention, with the most frequently mentioned items at the top of the list, the top item probably will have been mentioned by a majority of the sample (say 75%). Frequencies should then descend slowly, dribbling down to the lowest frequency (say, to the twentieth or fiftieth item that was mentioned only once). If instead, only the top two or three items were mentioned by a majority and the frequencies drop off sharply, that indicates (1) the domain is small (i.e., has only a few members) or, more likely, (2) the researcher did not use adequate probes and encouragements with informants, or (3) the researcher did not ask about a coherent domain.

Coding and standardization of responses. When the lists consist of phrases or statements rather than words it is common for each list to contain different phrasings of the same concept. Here, in tabulating the lists the researcher will have to use judgment as to which statements refer to the "same" concepts. The researcher also has to use judgment as to the best phrasing of each concept. The goal is to collect and tabulate verbatim responses, not to try and infer categories. When making categories for responses, the researcher runs the risk that the categories will reflect his or her own preconceived notions and biases and not the ideas of the informants. (To collect data on the categories used by your informants, refer to Chapters 3 and 5.) In difficult cases it would be desirable to seek the aid of informants in deciding when two different phrases represent a single concept.

The following criteria for standardizing statements are suggested as generally useful: use correct grammar and speech; make each statement autonomous and clear (e.g., the sentence "He works hard at it," should be rewritten as "He works hard at his job."); use present tense; avoid specific words and phrases where possible (e.g., use the word *car* rather than *Ford,* unless it changes the meaning).

What items should be included in the final study. The decision of how many items to use is determined by a number of considerations including: the purpose of the study, the number and frequency of the items elicited in the free listing, and the type of formal data collection format to be used. Clearly, the most frequently named items should be given top priority. Where to draw the cutoff point is more difficult to specify. Generally the domain is too large to allow consideration of every item and therefore the items to be studied are a sample of the whole. Sometimes low-frequency items are included to ensure variety of objects. There are no absolute rules for inclusion and exclusion of items. For many purposes a couple dozen items seems reasonable while for others larger samples of items may be required. The researcher should be sure that items included in the study domain *are known* by the vast majority of the informants. Otherwise the results can be severely biased by informants who are not acquainted with the items.

Strengths and Weaknesses of Free Listing

We recommend free listing as the first step in all research involving the definition of new domains. Free listing is the best way to ensure that the concepts and the domain are culturally relevant. It provides a strong source of cognitive data in terms of frequencies and the order properties of the individual lists. Informants can usually do the task in an easy and natural way. Free listing helps prevent researchers from using inappropriate items.

Free listing, however, is not a perfect tool. Sometimes it is difficult to find appropriate generic terms to start the listing process or lists may not be productive (lists are too sparse). Finally, there are no generally recognized ways to check the statistical reliability of the free listing task.

An Example

In order to illustrate the use of the free listing task, we present the results of a classroom exercise using the domains of "fruits" and "vegetables" to illustrate problems of defining boundaries between domains. The question arose as to whether the two domains were absolutely separate or whether they might overlap somewhat in membership. In order to get a preliminary idea of the relation of the two domains we asked 40 students to "List the names of all the fruits that you can." We asked a different set of 60 students to "List the names of all the vegetables that you can."

The results of the free listing tasks are shown in Tables 2.1 and 2.2. Table 2.1 shows the number of times each fruit was mentioned. The fruits are ordered from the most frequently mentioned to those that were mentioned only once. Table 2.2 shows the same data for vegetables. Note that there is a great range in the number of people that mentioned each fruit. For example, "apple" appears on 37 of the lists (about 93% of the respondents) and "taro" only appears on one list. This is typical of results from the free listing task. The question arises as to whether or not big differences in the proportion of times items are mentioned has any significance. We feel that it has great significance.

Notice that if we asked additional informants to list all the fruits that they knew that we might increase the number of items on the list in Table 2.1. Therefore the list is not to be taken as definitive and complete. The fruits that are not on the list are probably not as common or salient in these students' conceptions as are the fruits on the list. Notice that as the number of subjects increases we expect the list to become more stable, the order would not change, and few items would be added by each new subject.

TABLE 2.1
Frequency of Mention of "Fruits" in Free List Task

Apple	37	Honeydew	9
Orange	35	*Avocado	8
Pear	34	Mango	8
Banana	33	Date	7
Grape	32	Fig	7
Peach	30	Prune	7
Tangerine	27	Gooseberry	6
Cherry	26	Raisin	5
Grapefruit	26	*Pumpkin	4
Pineapple	26	Casaba melon	3
Strawberry	22	Kumquat	3
Watermelon	21	Melon	3
Lemon	20	Breadfruit	2
*Tomato	19	Kiwi	2
Apricot	18	Passionfruit	2
Blueberry	18	Persimmon	2
Plum	18	Cranberry	1
Cantaloupe	17	Crenshaw melon	1
Lime	16	Currant	1
Nectarine	14	Elderberry	1
Papaya	14	Huckleberry	1
Raspberry	14	Loganberry	1
Blackberry	13	Manderine	1
Boisenberry	12	*Rhubarb	1
Tangello	11	Salmonberry	1
Guava	10	*Squash	1
Pomegranate	10	Taro	1
Coconut	9	Turnip	1

TABLE 2.2
Frequency Distribution of "Vegetables" Free Listing Task

Green beans	55	Chinese peas	6
Corn	50	Greens	6
Carrots	49	Okra	6
Peas	41	Summer squash	6
Lima beans	40	Blackeyed peas	5
Lettuce	38	Swiss chard	5
Broccoli	37	Wax beans	5
Califlower	36	Bamboo shoots	4
Brussels sprouts	35	Navy beans	4
*Tomatoes	32	Alfalfa sprouts	3
Onions	30	Chile peppers	3
Spinach	30	Endive	3
Asparagus	29	Kidney beans	3
*Squash	28	Leek	3
Cucumbers	26	Parsnips	3
Celery	25	*Pumpkin	3
Cabbage	24	Redleaf lettuce	3
Zucchini	24	*Rhubarb	3
*Turnips	23	Water chestnuts	3
Potatoes	20	Butterleaf lettuce	2
Artichokes	18	Green onions	2
Bell peppers	18	Kale	2
Radishes	18	Kolari	2
*Avocado	18	Red onions	2
Beets	13	Sauerkraut	2
Rutabaga	11	Butternut squash	1
Bean sprouts	10	Garlic	1
Eggplant	9	Hubbard squash	1
Mushrooms	8	Jicama	1
Parsley	8	Peapods	1
Pinto beans	8	Pickles	1
Yams	7	Soybeans	1

*Indicates items that appear on both "fruit" and "vegetable" lists.

We can see from the lists that the two domains are not mutually exclusive. Six words occur on both lists. "Tomato" and "Avocado" appear frequently on both lists, while "Squash" and "Turnip" are more often seen as vegetables. In order to decide which list the items belong to we would need more information for "Pumpkin" and "Rhubarb." We might suspect that "Squash" and "Turnip" were really vegetables and that one person may have "made a mistake."

Where should we draw a line between the two domains? This is a matter of judgment and is usually based on the free list results. In this case, tomato and avocado might be assigned to both fruits and vegetables. In order to give an intuitive and visual answer to this kind of

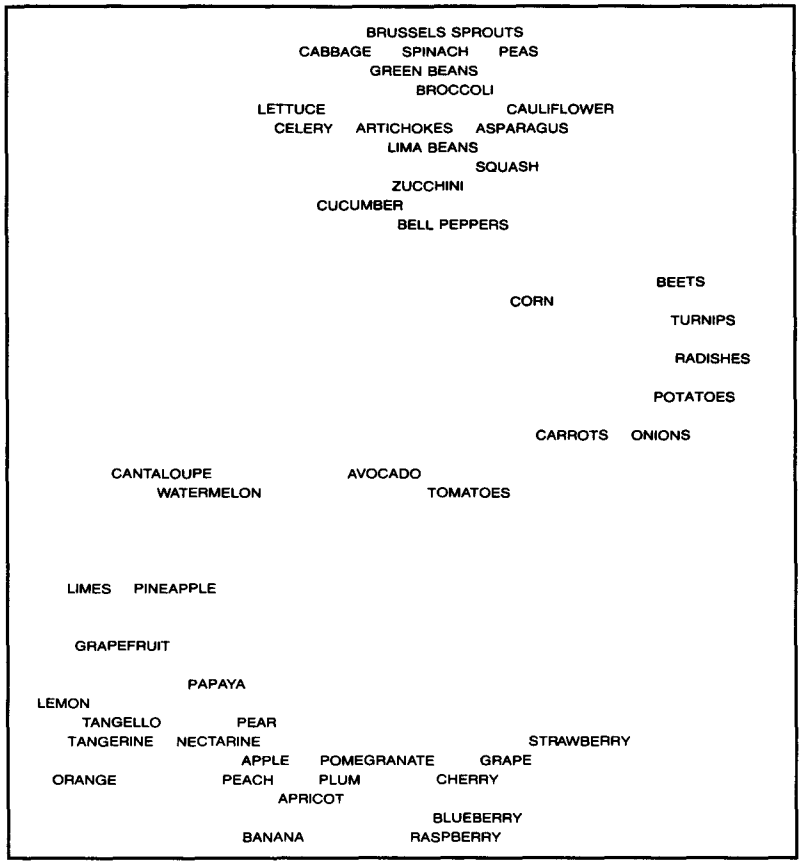

Figure 2.1: Judged Similarities Among Fruits and Vegetables Displayed in a Spatial Representation

situation we conducted further data collection and analysis and obtained a spatial representation of the more common fruits and vegetables in terms of their similarity. Although such an analysis is not necessary to define domain items, we present the results to illustrate the overlap between two domains and also to show kinds of results that can be obtained from structured interviewing. Students were asked to sort the names of fruits and vegetables into piles (see Chapter 3 for details). From the resulting judged similarity data we performed a multidimensional scaling analysis to obtain a spatial representation. Figure 2.1 shows the results. In this picture, items that were judged similar are close to each other while items judged as dissimilar are further apart. Closeness in the picture represents closeness in the similarity judgments.

Notice that the vegetables are all in the upper-right part of the picture and the fruits are in the lower part of the picture. Notice also that the root vegetables (beets, turnips, radishes, potatoes, onions, and carrots) are clustered together at the extreme right of the picture. Fruits that are berries are clustered at the lower right. A further cluster is formed by the citrus fruits (limes, grapefruit, lemon, tangello, tangerine, and orange) at the lower left. Thus the picture seems to represent similarities among the items in an interpretable way.

The main thing to notice from Figure 2.1 is that "Avocado" and "Tomato" are found exactly in the middle of the two groups. It can be seen that the two are intermediate between the two dense clusters that constitute "pure" fruits or "pure" vegetables. The cognitive boundary between fruits and vegetables is not a sharp all-or-none type distinction. Some items in each domain are more typical of the domain than others. The boundaries are somewhat arbitrary and may be defined differently for different purposes.

3. PILE SORT I: SINGLE SORTS

We turn now to more formal data collection methods that are appropriate when study items have been selected. We begin with sorting tasks. In a *pile sort* task informants are asked to sort cards, each containing the name of an item, into piles so that items in a pile are more similar to each other than they are to items in separate piles. In the unconstrained version of the task, subjects can make as few or as many piles as they wish. In the constrained version of the task, subjects are asked to create a specified number of piles. Subjects are generally asked to group items according to their similarity, without reference to specific criteria. The informant, rather than the researcher, decides what criteria are most salient and determine similarity. The pile sort is very easy to administer and allows for the collection of data among a large number of items.

Examples

The pile sorting task has been used extensively in field research. It is easy to use, respondents understand what is going on, and it can facilitate conversation. Miller and Johnson (1981) and Johnson and Miller (1983) used the pile sort to collect data on relations among fishermen. The task was quick, easy (some data was collected on the deck of a fishing boat while out at sea), and the men tended to talk about the social relations among the men that were being studied. When names

or descriptive phrases of objects are used on cards the task is limited to literate informants. In cases where it is possible to use the actual objects themselves, or reasonable pictures, the method can be extended to nonliterate informants.

In a series of ethnographic studies, Roberts and associates have used the pile sort to facilitate the description of relevant behavior events for a number of diverse culture patterns, such as eight ball pool (Roberts and Chick, 1979), pilot error (Roberts et al., 1980), women's trapshooting (Roberts and Nuttras, 1980), and tennis (Roberts et al., 1981). For example, using a combination of participant observation and structured interviewing, Roberts et al. (1980) explored kinds of errors P-3 pilots can make. Errors that were known to have occurred and that represented the major types of errors were selected for study. Pilots were then asked to sort 60 such errors into piles according to their similarity. After the sorting task, they were asked to rate the errors in terms of their severity. Hierarchical clustering and multidimensional scaling were used to uncover the categories and dimensions of error.

In another application of the pile sort, Freeman et al. (1981) and Romney et al. (1979) compared concepts of success and failure in the United States with those in Guatemala. They selected representative statements of characteristics of success and failure. Examples of some of the U.S. success statements were the following: "He's ambitious," "Everything works out for him and always for the good," "He's stable," and " Always there when he's supposed to be." Examples of failures were the following: "He drinks too much," "He seems to want things for free," and "He feels he has to cheat to get ahead." Judged similarity data was collected with a pile sorting task:

> Each statement was typed on a card and the respondents were handed randomly ordered stacks of cards. . . . They were asked to read through the stack of cards and then to sort them into piles, so that items in the same pile were more similar to each other than items in the other piles. The interviewer did not define the meaning of similarity and respondents used their own definitions [Romney et al., 1979: 307].

How to Do It

To use the pile sort procedure the items to be studied are usually presented on cards. Words or names may be written on index cards (for literates only) or pictures, drawings, or other visual stimuli may be presented (for literates and nonliterates). Cards are shuffled (randomized) before handing them to an informant. Informants are then asked to look through the cards and sort them into piles, so that similar items are in piles together. A subject may be allowed to make as many