



PHOTO BY HAFM JANSEN

13



Analyzing Qualitative Data

Topics covered in this chapter:

Approaches to qualitative analysis
Data coding
Using a computer for coding and categorizing
Data reduction and data displays
Ensuring rigor in qualitative research

In this chapter, we concentrate on making sense of textual data in the form of expanded field notes, transcripts of recorded interviews, or naturally occurring texts such as newspaper accounts, popular literature, or policy documents.

APPROACHES TO QUALITATIVE ANALYSIS

There are as many different ways to analyze qualitative data as there are means to collect it. Generally, however, all forms of analysis involve organizing the data according to specific criteria, reducing it to a more manageable form, displaying it in a form to aid analysis, and interpreting it.

Analyzing reams of written text is as complex and time-consuming as analyzing quantitative data—if not more so. Good qualitative analysis is part art and part science, and hundreds of books have been written on the topic. This chapter is not intended to substitute for these texts or to “teach” qualitative data analysis. Rather it is

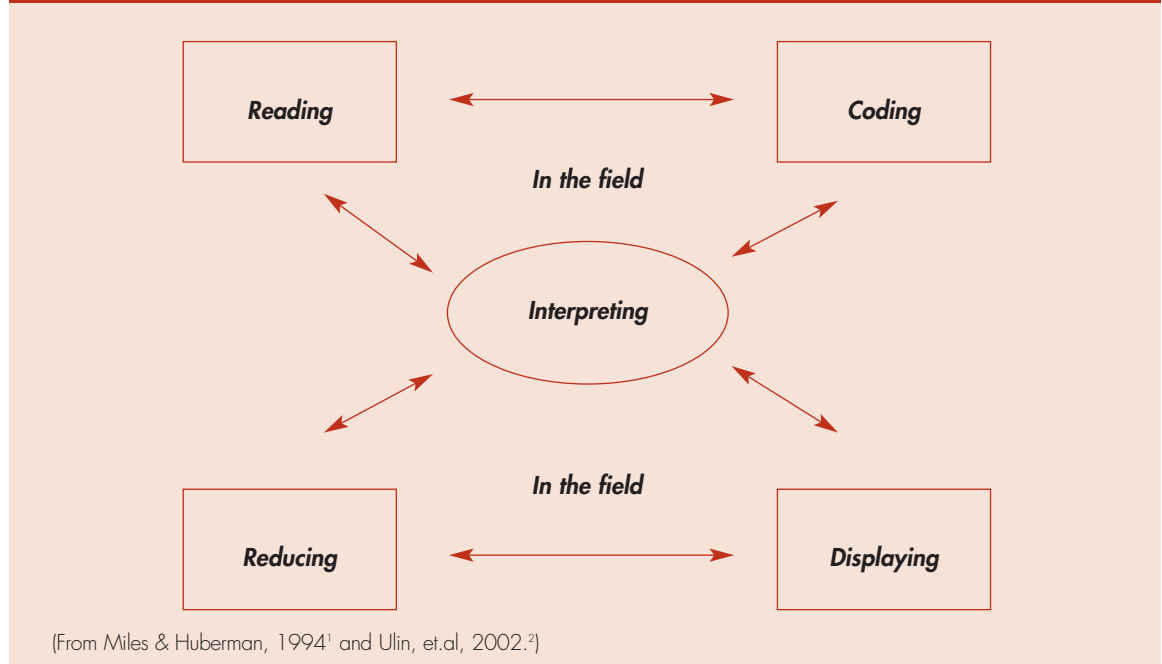
meant to acquaint the reader with the process and to direct interested parties to useful texts in the field. Generally, projects involving qualitative data analysis should proceed under the direction of someone trained in qualitative research.

We have found the framework of researchers Miles and Huberman helpful in understanding qualitative analysis.¹ Analysis, according to them, consists of “three concurrent flows of activity: data reduction, data display, and conclusion drawing or verification.” Ulin and colleagues offer a slightly revised version of the Miles and Huberman framework that identifies four separate flows of activity—data reading, coding, reducing, and displaying—all informing a constant process of interpretation.² As emphasized in Figure 13.1, this constant interplay of activity is an iterative process that begins in the field and continues at your desk once all the data are collected.

Below we describe each of these processes briefly before exploring data



FIGURE 13.1 THE PROCESS OF QUALITATIVE DATA ANALYSIS



coding and interpretation in greater depth.

Data immersion is the process of reading and rereading each set of notes or transcripts until you are intimately familiar with the content. Early readings of the data can help guide future data collection. Are you getting the kinds of data that you expected? Does your reading suggest new avenues of inquiry? Repeated readings are also the first step toward identifying emerging themes, possible relationships among themes or categories, and unusual or contradictory responses.

Data coding means attaching labels or **codes** to different segments of text that are associated with different issues. Coding is a way to help separate information into categories or themes so that information from different sources can be easily sorted and compared. With key themes coded in this way, you can later search and retrieve interesting segments and look at them as separate files. Data can either be coded by hand (by writing codes into the margins of the text) or with the aid of a computer. Computers have helped revolutionize qualitative data analysis by greatly easing

the task of sorting text segments by thematic code.

Data display is defined as an organized assembly of information that allows conclusions to be drawn and actions to be taken.¹ Most frequently qualitative data is displayed as narrative text, which tends to overload people's information-processing capabilities. Matrices, graphs, networks, and charts can present information in more compact form that make the data more accessible. Many of these techniques are presented later in this chapter.

Data reduction involves selecting, focusing, simplifying, abstracting, and transforming the "raw" data of field notes or transcriptions into typed summaries organized around themes or patterns based on the original objectives of the research. Data reduction continues until the final report is written.

Interpretation/conclusion drawing refers to the process of deciding what things mean, noting themes, regularities, patterns, and explanations. The researcher may begin to draw conclusions in draft form throughout the entire data collection



BOX 13.1 WHERE TO START: AN APPROACH TO QUALITATIVE DATA ANALYSIS

Transcribe and organize the information into files—one for each interview, focus group, or observation session.

Immerse yourself in the data. Get a sense of the whole. Read through all of the transcripts several times. Perhaps jot down some ideas as they come to mind.

Pick one transcript or document—the most interesting, the shortest, the one on the top of the pile. Go through it, asking yourself: What is this about? Read both for content and to identify emergent themes and insights into underlying meaning. Write thoughts in the margin or in your field notebook.

When you have completed this task for several documents, make a list of all the topics you encountered. Follow the trail of issues, intuitions, and ideas that arise from reading the data. Cluster together similar topics. Form these topics into columns that might be arrayed as major topics, unique topics, and leftovers.

1. Develop preliminary coding categories. Start by listing all the issues, perceptions, attitudes, beliefs, and other important aspects identified during the preceding stage. Once the principal coding categories have been identified, review the list and see if any categories overlap or are repeated. Then assign a code and, if so desired, code by color.
2. Now take this list and go back to your data. Write the codes next to the appropriate segments of the text. Try out this preliminary organizing scheme on several transcripts to see whether new categories and codes emerge.
3. Make a final decision on the abbreviation for each category and alphabetize these codes.
4. In the margin of each transcript, write the code assigned to each category found in the text (or use a computer to assign codes to particular segments of text).
5. Assemble the data belonging to each category in one place and perform a preliminary analysis.
6. Look for the most important relationships between the data. Also look for relationships between different analysis categories. In other words, look for the most important relationships between people, events, ideas, perceptions, behavior patterns, and other aspects of the data.
7. Discovering what is significant should lead to a search for similarities and differences. Look for the most common and recurrent norms and patterns in behavior, ideas, perceptions, attitudes, and expressions. Also look for the most salient differences.
8. Consider using one or more data display techniques to help make sense of your data—for example, matrices, decision trees, or flow charts.

(Adapted from Creswell, 1994³ and Shrader and Sagot, 2000.⁴)

exercise, but eventually these become more explicit and firm at the point when the “final report” is written. Conclusions must also be verified as analysis proceeds. As researchers try to explain what the data mean, they should continually examine their explanations for plausibility and validity—do their explanations make sense within the context of the study? It is often possible to test the validity of conclusions by presenting preliminary findings and interpretations back to project stakeholders and/or members of the population being researched. Researchers can then incorporate this feedback into the final version of their written report.

DATA CODING

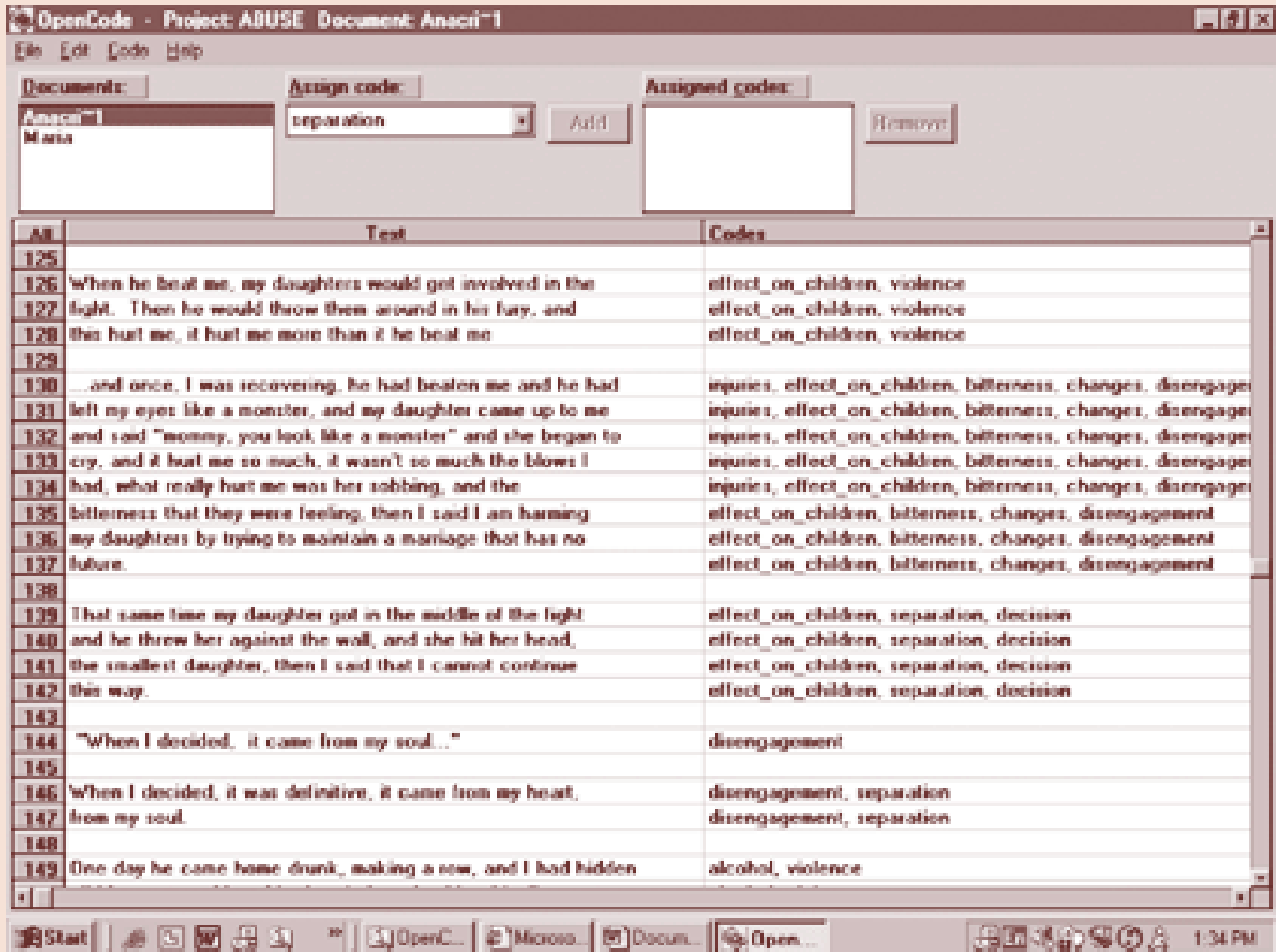
Although most qualitative research uses some process of coding to organize data, there are no standard rules about how to do so. As Ulin and colleagues observe:

Researchers differ on how to derive codes, when to start and stop coding, and the level of detail they want. Some researchers develop codes that closely match the ideas or language found in the textual data. They want to avoid imposing words or concepts that might prevent them from seeing their data in a new way. Others borrow terms from the social science literature that represent more abstract concepts



BOX 13.2 AN EXAMPLE OF DATA CODING USING OPENCODE

The following is an example of data coding and organization using the software OpenCode developed by Umeå University, Sweden. OpenCode allows the researcher to apply an unlimited number of codes to specific pieces of text, which may then be reorganized according to specific themes, or sets of codes. In this example, we used two interviews from Nicaraguan women, Ana Cristina and Maria, both of whom had been abused by their former husbands. In the interviews, both women described the characteristics of the violence, its effect on them and their families, different strategies they had used for coping with the violence, different sources of family and social support (or lack thereof), and finally how they made the decision to separate from their abusive partners.



Each interview was coded separately, and more than 100 different codes were generated at this stage. Some of the codes were related to specific language or feelings that the respondents used (e.g., "bitterness," "sadness") whereas others were thematic categories such as "effects on children" or "disengagement." After both interviews were coded, we organized the data into common themes, using the program's "search" option.

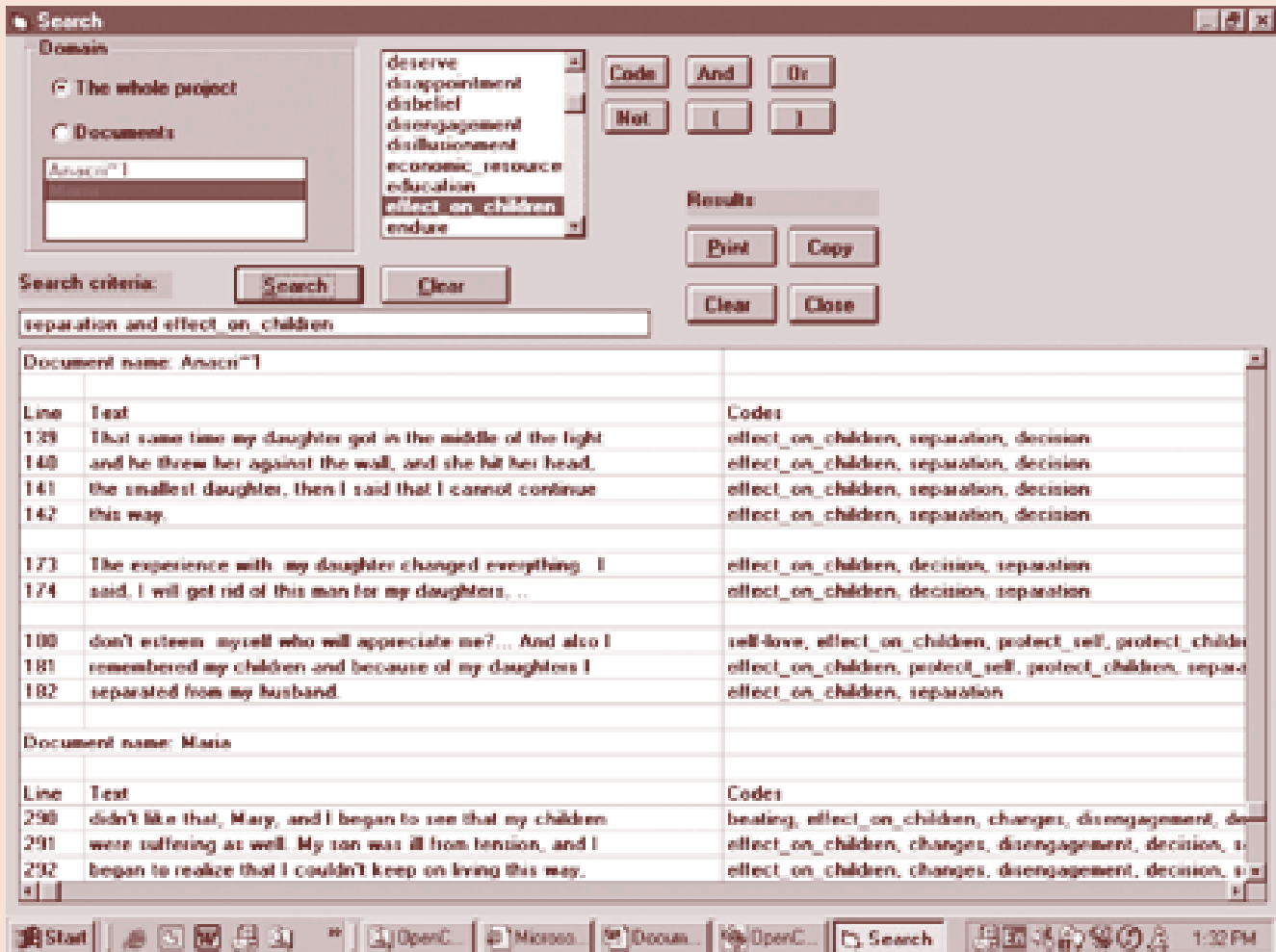
The second figure shows a piece of text where Ana Cristina describes how her children were affected by her husband's violence, as well as how their suffering influenced her own views towards the marriage. It appeared that the desire to protect her children was a significant factor in Ana Cristina's decision to leave her husband. To see whether similar themes might also be relevant for Maria, we conducted a search of both interviews using the criteria "effect_on_children" AND "separation."

The text that appears in the next figure shows that both women felt that their children's suffering was a major influence on their decision to leave.

In further analysis, we looked for commonalities and differences in other types of motivations mentioned for leaving (for example, the existence or lack of family support, economic resources, self-esteem) The same results could certainly have been accomplished by cutting and pasting text, but the use of the computer software was particularly useful for looking at a single piece of text from different perspectives, and in different combinations of codes.

BOX 13.2 AN EXAMPLE OF DATA CODING USING OPENCODE

Using computer software (OpenCode) to organize text by common themes.



important to their field. These have the advantage of being understood by a wider audience. Whether borrowed or emergent, labels allow you to assemble under one concept many seemingly disparate pieces of text and to search for connections.²

In part, these different approaches to coding evolve from the theoretical underpinnings of different types of qualitative research. **Grounded theory**, an approach to qualitative research pioneered in the late 1960s by Glaser and Strauss, emphasizes coding that sticks close to the text and then constructs increasingly abstract categories and domains by analyzing the

relationships among codes. Because this approach allows conceptual categories to emerge from the text rather than from the mind of the researcher, it is especially well suited for theory building.⁵

For more applied research projects, it is possible to code segments using categories that correspond to well-recognized phenomena or concepts. For example, when coding an interview from a woman living with abuse, it is easy to imagine codes emerging that correspond to different feelings, coping strategies, responses to abuse, and reactions of others. Even when working with more categorical



TABLE 13.1 SAMPLE MATRIX: FACTORS THAT MOTIVATE AN ABUSED WOMAN TO SEEK HELP OR INHIBIT HER FROM SEEKING HELP

<i>Forms of Violence</i>	<i>Motivating Factors</i>		<i>Inhibiting Factors</i>	
	<i>Internal</i>	<i>External</i>	<i>Internal</i>	<i>External</i>
Physical	<ul style="list-style-type: none"> ■ Fear that her daughter will be taken away. ■ Search for protection, she is beaten because of another woman. 	<ul style="list-style-type: none"> ■ Beatings, biting, miscarriages. ■ Loss of teeth. ■ Referred by a service provider; taken by her family. 	<ul style="list-style-type: none"> ■ He is the father of her children. ■ She does not believe in the system, fears more abuse. 	<ul style="list-style-type: none"> ■ The abuser is released from jail. ■ He threatens to kill her.
Psychological	<ul style="list-style-type: none"> ■ Depression; takes drugs to feel better. ■ Seeks support because of stress; wants to vent her feelings; wants to be heard. ■ Wants someone to talk with her husband. 	<ul style="list-style-type: none"> ■ He leaves the home. ■ He refuses to marry her. ■ The children have problems at school. 	<ul style="list-style-type: none"> ■ She doesn't think this a matter to be handled by the authorities. ■ She thinks it is her fault. ■ She thinks he will change. ■ She thinks that her jealousy causes his infidelity. 	<ul style="list-style-type: none"> ■ Her mother tells her she should put up with it. ■ The church tells her to forgive him. ■ He asks forgiveness. ■ He leaves the other woman.
Sexual	<ul style="list-style-type: none"> ■ Fear of dying of an infection because he does not get treated. ■ Loss of libido. 	<ul style="list-style-type: none"> ■ Sexual rejection from him. ■ She has been told that this is sexual aggression. 	<ul style="list-style-type: none"> ■ She is ashamed to have others find out about her problem. ■ She thinks no one will believe her. 	<ul style="list-style-type: none"> ■ She thinks it is her duty to accept her relationship with her husband.
Economic	<ul style="list-style-type: none"> ■ She thinks he should support and help her. ■ She doesn't think her status should be reduced unfairly. 	<ul style="list-style-type: none"> ■ He won't give her money. ■ Her mother-in-law stays she should turn him in to the authorities. 	<ul style="list-style-type: none"> ■ Fear. ■ She thinks she does not need money or prefers not to insist on it in order to keep the peace. 	<ul style="list-style-type: none"> ■ The authorities do not force him to pay. ■ They do not monitor his support payments.

(From Shrader & Sagot, 2000.⁴)

labels, one's coding system will necessarily evolve over time as new insights and ways to organize the data emerge.

Figure 13.1 describes one approach to tackling qualitative analysis, following the conceptual approach outlined by Miles and Huberman and Ulin and colleagues. It begins with data immersion and includes developing and applying more categorical labels to segments of text. For

very simple projects, it may be possible to organize segments of text according to the various research questions of interest and then begin to look for similarities, differences, themes, and relationships. By its very nature, one's approach to qualitative analysis will vary based on complexity of the research question being asked and the resources and time available for analysis.



USING A COMPUTER FOR CODING AND CATEGORIZING

Sometimes a more complex classification scheme is necessary than can be managed and interpreted by hand. Here is where computers can come in handy. When using a computer for text-based analysis, the investigator still must “code” the data—that is, associate a code word with pieces of the text that represent a certain theme. But the computer allows the investigator to search and organize large volumes of data according to predefined codes and combinations of codes.

Many computer programs are available for use in coding and analyzing data, and all are different. Examples of some of the more common programs used for qualitative data analysis include HyperQual, NUDIST, Ethnograph, and QUALPRO. Generally these programs allow the investigator to attach one or more codes to different text segments. Data can then be sorted and resorted according to different themes and coding categories. A number of popular qualitative software programs are available commercially through Scolari Software at: <http://12.22.103.11/>.

We especially recommend considering one of three free “share-ware” programs that have been explicitly developed to assist researchers in the field of public health. The first, known as OpenCode, is a versatile, but highly accessible, program that can be downloaded from the web site of Umeå University, Sweden, at http://www.umu.se/phmed/epidemi/forskn ing/open_code.html. The second is CDC EZ Text, a simple program designed by the Centers for Disease Control and Prevention (CDC) to facilitate analysis of open-ended responses to structured questionnaires. CDC has also developed a third program,

called AnSWR, which is designed for more complex qualitative projects. Both CDC programs can be downloaded from <http://www.cdc.gov/hiv/software>.

There are also potential drawbacks to using computers, and the time and cost of using computers must be weighed against the potential benefits. For example, it is time-consuming to set up and code data by computer and the investigator must invest time, money, and energy into acquiring and learning a new software program. Clearly, for many small projects, it is best to conduct the analysis manually using highlighters and colored pens.

DATA REDUCTION AND DATA DISPLAYS*

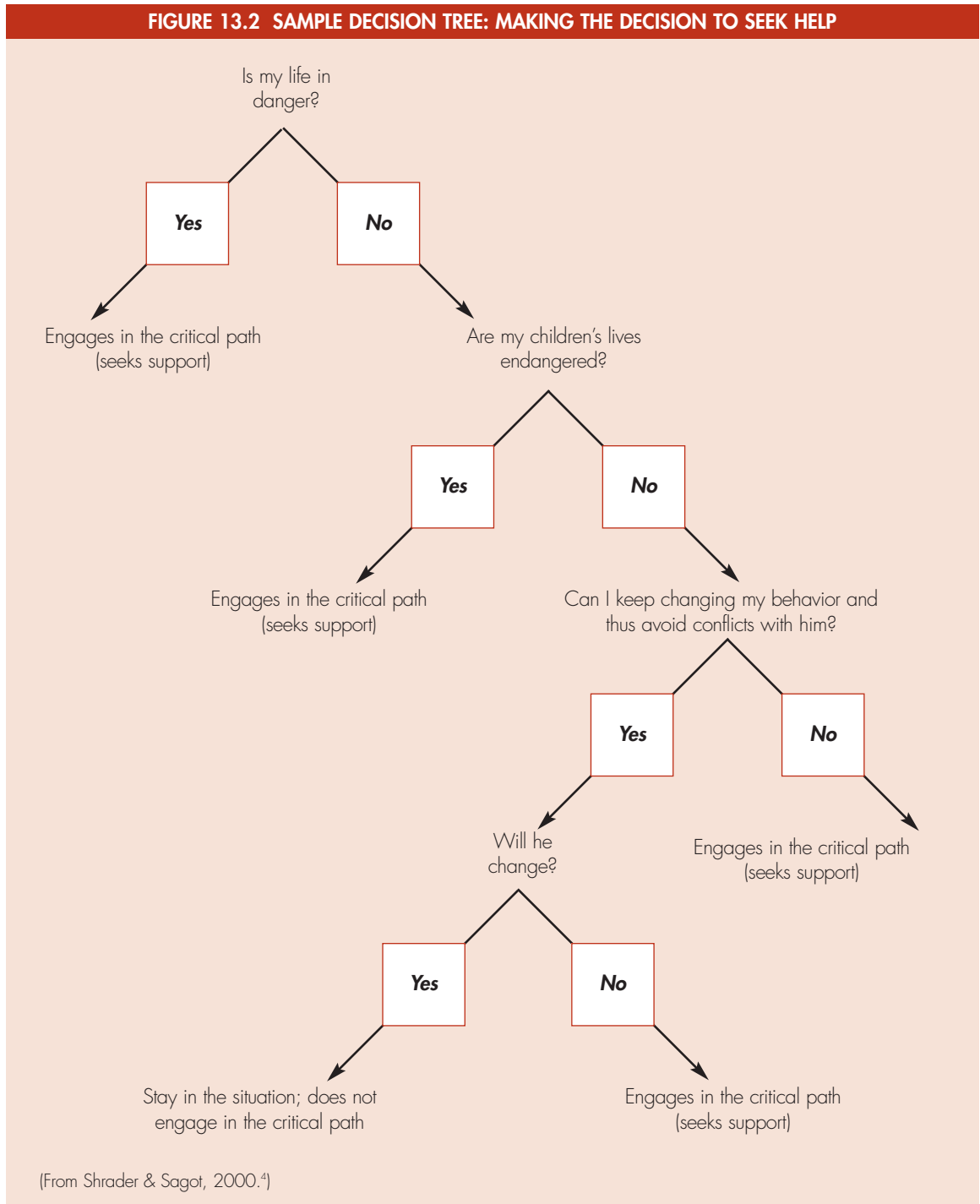
Data reduction means distilling the information to make visible the most essential concepts and relationships. The reduction process usually begins after all the data have been collected, coded, and the researcher is thoroughly familiar with the material. The goal of data reduction is to get the bigger picture from the data and begin to discern distinctions between primary and secondary themes.² While coding helps break down the data into smaller parts, data reduction is the process of abstracting back out from the particular to the conceptual.

To get a wider perspective on the data, it is often helpful to use one or more visual devices known as data displays. Data display is a nonnarrative way of analyzing and presenting qualitative data. The analysis may be intra-case (that is, analysis of the data related to a single respondent) or inter-case (comparative analysis of the data from two or more respondents). As with any other aspect of qualitative analysis, it is necessary to document the processes followed, the

* Information in this section is reprinted with permission from Shrader and Sagot, 2000.⁴



FIGURE 13.2 SAMPLE DECISION TREE: MAKING THE DECISION TO SEEK HELP



conclusions generated, and the way in which the analysis influenced the subsequent stages of the fieldwork.

Matrices, decision trees, and taxonomies are examples of different data display tools that can be useful for aiding analysis as well as in presenting results. Following is a presentation of these three

methods using examples from the “Ruta Crítica” study, a multi-country qualitative study on women’s responses to violence performed by the Pan American Health Organization in ten Latin American countries. (See Box 5.4 for a description of the study.) Keep in mind that it is not necessary to display data in every way



suggested, only in those that can elucidate some aspect of the analysis.

Matrices

A matrix is a table that displays the intersection of two or more concepts. It consists of cells, which may be filled with summary text, graphs, or other explanatory information. The matrix is very useful for summarizing a great deal of information or presenting comparative information on a single sheet. Table 13.1, for example, represents a data matrix used to summarize the factors that emerged as critical to either motivating a woman affected by violence to seek outside help or inhibiting her from seeking help. This particular matrix is taken from the Panama country report of the Ruta Crítica study. It illustrates the complexity and detail that can be included in a single table.

Decision trees

A decision tree is a graphic representation that depicts the sequence of decisions that leads to an action. In general, the decisions are formulated as a dichotomous question, i.e., one that is answered with “yes” or “no.” Arrows indicate the consequent question, depending on the answer given. A simple decision tree shows the sequence of decisions made by a person; a complex decision tree shows the combined decisions of one or more persons.

Figure 13.2 illustrates a simple decision tree derived from Ecuador data of PAHO’s Ruta Crítica study. The figure illustrates the thinking and decisions characteristic of an abused woman’s process of deciding whether to seek outside help for the violence.

Flow charts

Also known as an explanatory network, flow charts show the relationship between two or more elements, emphasizing directionality, causality, or temporary association. Geometric figures and arrows usually

are used to indicate temporary or causal relationships between the concepts. Figure 13.3 presents a sample flow chart showing the path followed by a woman who seeks to escape from a violent relationship. The chart includes an analysis of the time elapsed between one violent incident and the next one, as well as the time it takes to carry out the trial. It reveals that this respondent was the victim of three serious incidents of violence before finally obtaining the desired result—for him to stop drinking and beating her.

Taxonomies

A taxonomy attempts to summarize concepts in a hierarchy that goes from the greatest level of generalization to the greatest level of specificity. Taxonomies are useful for showing the relationship between many separate concepts, joining concepts that share one or more characteristics in cumulative stages.

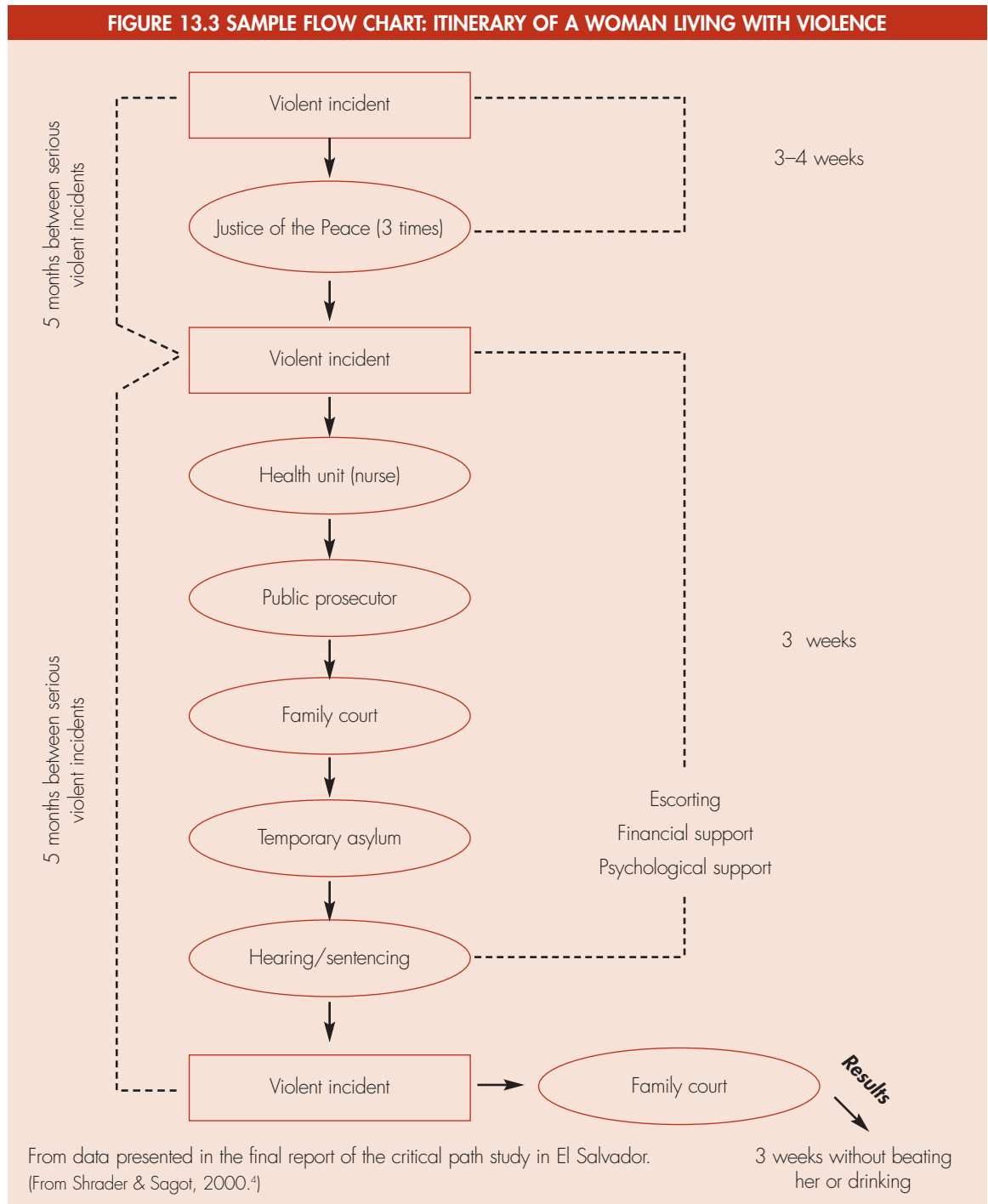
Figure 13.4 presents an example of a taxonomy developed by the U.S. Centers for Disease Control to describe several forms of interpersonal violence.

ENSURING RIGOR IN QUALITATIVE RESEARCH

Ensuring rigor, while equally important in qualitative as in quantitative research, is addressed somewhat differently by the two traditions. Lincoln and Guba suggest that the four issues that must be addressed in any systematic enquiry into humans and human behavior are truth value, applicability, consistency, and neutrality.⁶ (See Table 13.2.) Whereas quantitative research tends to use such criteria as external and internal validity, reliability, and objectivity to assess whether research has been carried out and interpreted using accepted scientific standards of rigor, qualitative research tends to assess the trustworthiness of research in terms of credibility,



FIGURE 13.3 SAMPLE FLOW CHART: ITINERARY OF A WOMAN LIVING WITH VIOLENCE

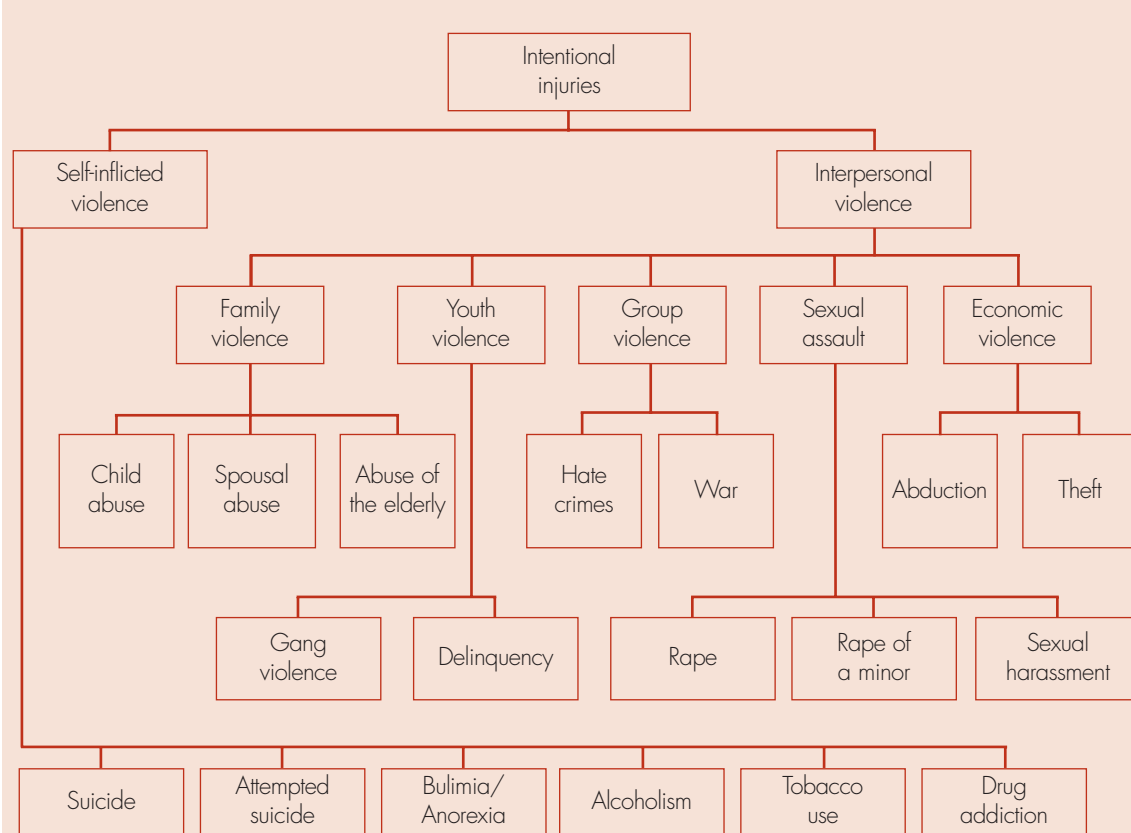


transferability, dependability, and confirmability. The naturalistic research tradition does not assume that there is a single truth that may be revealed by inquiry. Instead it affirms that many truths may exist at the same time, and these may be discovered by research, and at the same time modified by the research endeavor

itself. Different standards are therefore needed to assess the rigor of qualitative research. Table 13.2 presents a comparison of how different aspects of rigor are addressed in the different traditions.

Lincoln and Guba also propose alternative criteria and techniques for establishing trustworthiness within a naturalistic

FIGURE 13.4 SAMPLE TAXONOMY: PUBLIC HEALTH PERSPECTIVES TOWARD VIOLENCE



This taxonomy was adapted from a typology used by Rosenberg et al. of the United States Centers for Disease Control and Prevention (CDC), Center for Injury Prevention and Control, Atlanta, Georgia. (Cited in Shrader and Sagot, 2000.)⁴

framework. In place of validity, credibility is suggested as a criterion for determining the truth value of a qualitative study.

According to Sandelowski,

*A qualitative study is credible when it presents such faithful descriptions or interpretations of a human experience that the people having that experience would immediately recognize it from those descriptions or interpretations as their own. A study is also credible when other people (other researchers or readers) can recognize the experience when confronted with it after having only read about it in a study.*⁸

There are a variety of techniques that may enhance credibility of a study. For example:

Prolonged engagement in the field.

The researcher should be involved in a community or study site long enough to be

thoroughly oriented to the cultural and historical context, and to be able to detect and take into account possible distortions due to misinformation or his or her own subjective responses to the data.

Triangulation. The use of multiple sources, methods, and investigators to explore the same topic can increase credibility. For example, if the goal of a study is to assess the quality of care given to survivors of abuse in a community clinic, it might be useful to compare the views of health care providers with women who have used the services. The use of multiple sources might reveal important differences, not only as to the overall level of client satisfaction, but also in the way that clients and providers define quality of care. A mix of methods, for example, combining an



TABLE 13.2 CRITERIA ASSESSING THE TRUSTWORTHINESS OF RESEARCH FINDINGS ACCORDING TO QUANTITATIVE AND QUALITATIVE RESEARCH TRADITIONS

Aspect	Qualitative criteria	Quantitative criteria	Question asked
Truth value	Credibility	Internal validity	Have we really measured what we set out to measure?
Applicability	Transferability	External validity	How applicable are our results to other subjects and other contexts?
Consistency	Dependability	Reliability	Would our findings be repeated if our research were replicated in the same context with the same subjects?
Neutrality	Confirmability	Objectivity	To what extent are our findings affected by personal interests and biases?

(From Dalgren, et al., 2003.)

exit survey of clients with in-depth interviews, focus groups, or reviews of medical charts, might provide additional insight. Having two different researchers code interview transcripts and compare results afterwards is another form of triangulation that may increase the credibility of the findings.

Negative case analysis. This refers to seeking cases actively that appear to be exceptions to the general pattern, as a way of testing and refining hypotheses. The method relies on the analysis of “outliers” to improve the credibility of the findings.

Member checking. Member checking refers to the process by which data, conclusions, and interpretations are presented to members of the groups from which the data were collected. It is a powerful technique for establishing credibility. Member checking may be performed throughout the research process, through both formal and informal sessions. It allows stakeholders the

opportunity to react to the data and to correct errors in the collection and interpretation of data. Most importantly, it helps to establish the meaningfulness of the findings and interpretations. In a review of Central American health programs addressing gender-based violence, the preliminary results were presented to representatives from the different countries that participated in the original focus group discussions. The purpose of the session was to correct factual errors, as well as to hear reactions to the overall conclusions of the study. Not only did the discussion strengthen the interpretation of the findings, but it also created ownership of the findings among the participants. The study results were disseminated and discussed widely among the programs, and many of the recommendations were implemented.

Qualitative research refers to the transferability of findings instead of external validity or generalizability. In survey research, the ability to generalize findings to a broader population depends largely on how representative the sample is with regard to demographic characteristics and other key variables. Since qualitative research relies on purposive rather than random sampling, there is no expectation that findings from one small group of study participants can be statistically representative of a larger population. Theory generated from qualitative research may well be generalizable to a broader population. As Dalgren and colleagues argue, qualitative researchers aim for analytical generalizations, taking into account that,

Each subject has been selected to contribute to the theory that is being developed, and the knowledge gained from this theory should fit all scenarios that may be identified in a larger population. Hence the theory is applicable also beyond the study sample to all similar situations, questions and problems, regardless of demographic characteristics. In this way, qualitative researchers are



*no less able to generalize beyond their sample than are epidemiologists.*⁷

In quantitative research, consistency or reliability of results is a key aspect of trustworthiness. If another researcher conducted the same study using the same or similar methods on the same or similar population, would the same results be obtained? Reliability is enhanced in quantitative research by such measures as accurate instruments and equipment, well-trained fieldworkers, and uniform data collection procedures. In contrast, qualitative research is based on the premise that researchers and study participants are intimately related, and their interaction inevitably shapes the research process. Therefore, there is no pretense that the same results could be achieved with different researchers and participants.

Accordingly, qualitative research is assessed by its dependability, or the degree to which the researcher can describe and account for the way data were collected, interpreted, and how changing conditions were addressed throughout the process. The research process may be documented by means of an **audit trail**, including field notes, raw data (such as transcripts and diagrams) and process notes (procedures, designs, and strategies.) An audit trail should allow another to follow the “decision trail” used by the researcher.⁶

Finally, quantitative and qualitative traditions diverge on the issue of neutrality. Quantitative researchers aspire to maintain an objective attitude during the research process, and to avoid letting their feelings or subjective views influence the data collection or statistical analysis. Qualitative researchers argue that subjectivity is unavoidable, and that the interaction between participants and researchers actually enriches qualitative research. The neutrality of the data rather than that of the researcher is sought, using the criteria of confirmability. Audit

trials also help establish confirmability. If the research can be confirmed the auditor should be able to find conclusions grounded in the data.⁶

1. Miles M, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Thousand Oaks, California: Sage Publications; 1994.
2. Ulin P, Robinson E, Tolley E, McNeill E. *Qualitative Methods: A Field Guide for Applied Research in Sexual and Reproductive Health*. Research Triangle Park: Family Health International; 2002.
3. Creswell JW. *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks, California: Sage Publications; 1994.
4. Shrader E, Sagot M. *Domestic Violence: Women's Way Out*. Washington, DC: Pan American Health Organization; 2000.
5. Strauss A, Corbin J. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, California: Sage Publications; 1990.
6. Lincoln Y, Guba E. *Naturalistic Inquiry*. London, UK: Sage; 1985.
7. Dahlgren L, Emmelin M, Winqvist A. *Qualitative Approaches for International Public Health*. Umeå, Sweden: Umeå University; 2003.
8. Sandelowski M. The problem of rigor in qualitative research. *Advances in Nursing Science*. 1986;8(3):27-37.