Analysis of Qualitative Data

Much of the best work in sociology has been carried out using qualitative methods without statistical tests. This has been true of research areas ranging from organization and community studies to microstudies of face to face interaction and macrostudies of the world system. Nor should such work be regarded as weak or initial "exploratory" approaches to those topics.
—Randall Collins, "Statistics versus Words," p. 340

In field research, historical-comparative research, and a few other research areas, we collect a great deal of qualitative data to describe details about people, actions, and events in social life. The data are in the form of text from documents, observational notes, open-ended interview transcripts, physical artifacts, audio- or videotapes, and images or photos. It is not enough to collect the data; we also must analyze it. In qualitative approaches to research, analysis begins while gathering data, but such analysis tends to be tentative and incomplete.

To analyze data means systematically to organize, integrate, and examine; as we do this, we search for patterns and relationships among the specific details. To analyze, we connect particular data to concepts, advance generalizations, and identify broad trends or themes. Analysis allows us to improve understanding, expand theory, and advance knowledge.

The data used in quantitative studies are almost exclusively in the form of numbers. Compared to the vast volume, variety, and mutability of nebulous qualitative data, numbers are precise, uniform, standardized, and compact carriers of information. Applied mathematics has a large, highly developed area devoted to the analysis of numbers. Moreover, the statistics we use to analyze quantitative social science data are the same as those used across all quantitative science and many applied areas (e.g., business, education, medicine, agriculture, engineering, and so forth). As computer technology has advanced over the past 40 years, statisticians and computer scientists have developed a large array of sophisticated software and widely available programs to assist in quantitative data analysis.

Little of the vast statistical knowledge and related computer software is applicable for the analysis of qualitative data. Qualitative research allows us to be systematic and logically rigorous but often in different ways from statistical analysis.

Until about 20 years ago, qualitative researchers rarely explained how they analyzed data. In fact, a frequent criticism of qualitative research was that data analysis is not explicit or open to inspection, but its analysis has become more explicit and systematic.[1] We now have computer software for qualitative data analysis, some grounded in mathematical and other logical relations. Nonetheless, we use many approaches to qualitative data analysis.

This chapter has four parts. We first compare the similarities and differences between qualitative and quantitative data analysis. Next we discuss how to use coding and concept/theory building to assist in analyzing qualitative data. Third, we review some major analytic strategies that qualitative researchers have used and show how they link data to theory. We also examine the role that absence of direct, observable evidence can have in explanation. Lastly we review a few specific techniques available to examine patterns in the qualitative data.

COMPARISON OF METHODS OF DATA ANALYSIS

Qualitative and quantitative forms of data analysis have similarities and differences. In this section, we look at four similarities and four differences.

Similarities

First, in both types of data analysis, we infer from the empirical details of social life. To *infer* means to pass a judgment, to use reasoning, and to reach a conclusion based on evidence. In both forms of data analysis, we must carefully examine empirical information to reach a conclusion based on reasoning and simplifying the complexity in the data. This process requires some abstraction, or a moving back from the very specific details of concrete data, but how much this occurs varies. In all cases, we remain faithful to what is in the original, raw data.

Both forms of data analysis anchor statements made about the social world in an inquiry that has "adequacy." As Morse (1994:230) observed, "In qualitative research, *adequacy* refers to the amount of data collected, rather than to the number of subjects as in quantitative research. Adequacy is attained when sufficient data has been collected that saturation occurs."

A second similarity is that the analysis involves a public method or process. As we gather large amounts of data, we make our actions accessible to others. We describe the data and document the ways we collected and studied it, and we make how we did these things open to inspection by other members of the scientific community. The degree to which the method is standardized and visible varies. As King et al. (1994:118) noted, "Research designs in qualitative research are not always made explicit, but they are at least implicit in every piece of research."

Third, comparison is central in all data analysis, qualitative and quantitative. We compare the evidence we gathered internally or with other related evidence. We explore the data and identify multiple process, causes, properties, or mechanisms within it, looking for patterns: similarities and differences, aspects that are alike and unlike:

> *[Qualitative] researchers examine patterns of similarities and differences across cases and try to come to terms with their diversity. . . . Quantitative researchers also examine differences among cases, but with a different emphasis, the goal is to explain the covariation of one variable with another, usually across many cases. . . . The quantitative researcher typically has only broad familiarity with the cases. (Ragin, 1994a:107)*

Fourth, in both forms of data analysis, we strive to avoid errors, false conclusions, and misleading inferences. We are vigilant and alert for possible fallacies or illusions. As we sort through various explanations, discussions, and descriptions, and evaluate the merits of rival ways to describe and explain. We always seek the most authentic, valid, true, or worthy description and explanation among the alternatives.

Differences

Quantitative researchers can choose from a set of specialized, standardized data analysis techniques. Hypothesis testing and statistical methods are similar across the natural and social sciences. Quantitative analysis is highly developed and builds on a large body of applied mathematics. In contrast, qualitative data analysis is less standardized. The wide variety in qualitative research is matched by the many approaches to data analysis. An added complexity to having many approaches is that qualitative research is often inductive. We do not know the specifics of data analysis when we begin a project. Schatzman and Strauss (1973:108) remarked, "Qualitative analysts do not often enjoy

the operational advantages of their quantitative cousins in being able to predict their own analytic processes; consequently, they cannot refine and order their raw data by operations built initially into the design of research."

A second difference is that we do not begin data analysis in quantitative research until we have collected the data. Only then do we manipulate the numbers in seeking patterns or relationships. In qualitative research, we start looking for patterns or relationships while collecting data. We use results from early data analysis to guide subsequent data collection. Thus, analysis is less a distinct final stage of research than a dimension of research that stretches across all stages.

Another difference is the relation to social theory. Quantitative research involves manipulating numbers that represent empirical facts to test abstract hypotheses comprised of variable constructs. In contrast, qualitative research frequently creates new concepts and theory by blending empirical evidence with abstract concepts. Instead of testing a hypothesis, we may illustrate or color evidence to show that a theory, generalization, or interpretation is plausible.

The fourth difference is the degree of abstraction or distance from the details of social life (see Summary Review Box 1, Comparing Quantitative and Qualitative Data Analysis). In all data analysis, we place specific raw data into broader categories. We then examine and manipulate categories to identify patterns. In quantitative analysis, this process is clothed in statistics, hypotheses, and variables. We assume that we can capture or measure using numbers and then manipulate the numbers with statistics to reveal key features of social life.

In contrast, data in qualitative analysis are relatively imprecise, diffuse, and context based and can have more than one meaning. This is not always a disadvantage.

> *Words are not only more fundamental intellectually; one may also say that they are necessarily superior to mathematics in the social structure of the discipline. For words are a mode of expression with greater open-endedness, more capacity for connecting various realms of argument and experience, and more capacity for reaching intellectual audiences. (Collins, 1984:353)*

SUMMARY REVIEW BOX 1

Comparing Quantitative and Qualitative Data Analysis

SIMILARITIES	DIFFERENCES
Both infer from empirical data to abstract ideas	*Quantitative* uses a few shared, standardized techniques. *Qualitative* uses many diverse, nonstandard techniques.
Both use a public process and described in detail	*Quantitative* analyzes after all data have been collected. *Qualitative* begins data analysis while still collecting data.
Both make comparisons	*Quantitative* tests preexisting theories and hypotheses. *Qualitative* conceptualizes and builds a new theory.
Both avoid errors and false conclusions	*Quantitative* uses precise and compact abstract data. *Qualitative* uses imprecise, diffuse, relatively concrete data.

Explanations and Qualitative Data

We do not have to choose between a rigid idiographic/nomothetic dichotomy: that is, between describing specifics and verifying universal laws. When analyzing qualitative data, we develop explanations or generalizations that are close to concrete data and contexts, and we usually use less abstract theory. We may build new theory to create a realistic picture of social life and stimulate understanding more than to test causal hypotheses. The explanations tend to be rich in detail, sensitive to context, and capable of showing the complex processes or sequences of social life. They may or may not be causal. Our goal is to organize specific details into a coherent picture, model, or set of tightly interlocked concepts.

Qualitative explanations can be either highly unlikely or highly plausible. We provide supportive evidence to eliminate some theoretical explanations from consideration and to increase the plausibility of others. Qualitative analysis can eliminate an explanation by showing that a wide array of evidence contradicts it. The data might support more than one

explanation, but not all explanations will be consistent with it. In addition to eliminating less plausible explanations, we often want to verify a sequence of events or the steps of a process. This temporal ordering is the basis of finding associations among variables, and it supports causal arguments.

A few qualitative researchers are almost entirely descriptive and avoid theoretical analysis. In general, we always want to make theories and concepts explicit. Without an analytic interpretation or theory, the readers of qualitative research may use their own everyday, taken-for-granted ideas. Such ideas rarely advance general knowledge. Moreover, their commonsense framework will contain unexamined assumptions, biases, ethnocentrism, and ill-defined concepts taken from dominant cultural values.[2]

CODING AND CONCEPT FORMATION

Qualitative research often involves the use of general ideas, themes, or concepts as tools for making generalizations. Many are nonvariable concepts or simple nominal-level variables.

Conceptualization

When we perform quantitative research, we conceptualize variables and refine concepts as a step to measure variables. In contrast, when we perform qualitative research, we form new concepts or refine concepts that are grounded in the data. Concept formation is an integral part of data analysis and begins during data collection. Thus, conceptualization is a way to organize and make sense of data.

Those who conduct qualitative studies analyze by organizing data into categories based on themes, concepts, or similar features. While doing this, they may also develop new concepts, formulate conceptual definitions, and examine the relationships among concepts. Eventually, these researchers will link concepts to each other in terms of a sequence, as oppositional sets (*X* is the opposite of *Y*), or as sets of similar categories that are interwoven into theoretical statements.

You may begin to form concepts as you read through and ask critical questions of the data (e.g., field notes, historical documents, secondary sources). The questions can come from the abstract vocabulary of an academic field discipline such as sociology, for example: Is this a case of class conflict? Was role conflict present in that situation? Is this a social movement? Questions can also be logical, for example: What was the sequence of events? How does the way it happened here compare to the way over there? Are these the same or different, general or specific cases?[3]

Concept and evidence are mutually interdependent, particularly in case-study analysis. Cases are not given preestablished empirical units or theoretical categories apart from data; together, the data and theory define them. As you organize data and apply ideas, you create or specify a case. Making a case, called *casing*, occurs when you bring data and theory together. Determining what to treat as a case helps you resolve the tension between what you actually observe and your ideas about what you observe. "Casing viewed as a methodological step, can occur at any phase of the research process, but occurs especially at the beginning of the project and at the end" (Ragin, 1992b:218).

Coding Qualitative Data

When you code quantitative data, you arrange measures of variables into a machine-readable form for statistical analysis. Coding data has a different meaning in qualitative research than in quantitative research. In qualitative research you organize the raw data into conceptual categories and create themes or concepts. Instead of being a clerical task of data management, qualitative coding is an integral part of data analysis. Your research question provides a guide, but the process often leads to new questions. It frees you from entanglement in the details of the raw data and encourages you to think about them at a higher level, moving toward theory and generalizations:

> *Codes are tags or labels for assigning units of meaning to the descriptive or inferential information compiled during a study. Codes usually are attached to "chunks" of varying size—words, phrases, sentences or whole paragraphs, connected*

or unconnected to a specific setting. (Miles and Huberman, 1994:56)

Strauss (1987) defined three types of qualitative data coding and suggests you review the data on three occasions, using a different coding each time. He (p. 55) warned, "Coding is the most difficult operation for inexperienced researchers to understand and to master."[4]

1. *Open coding.* You preform **open coding** during a first pass through recently collected data. You locate themes and assign initial codes in your first attempt to condense the mass of data into categories. As you slowly read field notes, historical sources, or other data, you look for critical terms, central people, key events, or themes. Next you write a preliminary concept or label at the edge of a note card or computer record and highlight it with a different color or in some other distinctive way. You want to remain open to creating new themes and to changing these initial codes in subsequent analysis. A theoretical framework helps if you apply it in a flexible manner. When using open coding you bring themes to the surface from deep inside the data. The themes are at a low level of abstraction and come from your initial research question, concepts in the literature, terms used by members in the social setting, or new thoughts stimulated by an immersion in the data. As Schatzman and Strauss (1973:121) warned, you should see abstract concepts in concrete data and move back and forth between abstract concepts and specific details.

An example of moving between abstract concepts and details is found in LeMasters's (1975) field research study of a working-class tavern when he found that marriage came up in many conversations. If he open coded field notes, he might have coded a block of field notes with the theme "marriage." Following is an example of hypothetical field notes that can be open coded with this theme:

I wore a tie to the bar on Thursday because I had been at a late meeting. Sam noticed it immediately and said. "Damn it, Doc. I wore one of them things once—when I got married—and look what happened to me! By God, the undertaker will have to put the next one on." I ordered a beer, then asked him, "Why did you get married?" He replied, "What the hell you goin' to do? You just can't go on shacking up with girls all your life—I did plenty of that when I was single" with a smile and wink. He paused to order another beer and light a cigarette, then continued, "A man, sooner or later, likes to have a home of his own, and some kids, and to have that you have to get married. There's no way out of it—they got you hooked." I said, "Helen [his wife] seems like a nice person." He returned, "Oh, hell, she's not a bad kid, but she's a goddamn woman and they get under my skin. They piss me off. If you go to a party, just when you start having fun, the wife says 'let's go home.'" (Adapted from LeMasters, 1975:36–37)

Historical-comparative researchers also use open coding. For example, I studied the Knights of Labor, a nineteenth-century U.S. movement for economic and political reform. I read a secondary source about the activities of a local branch of the movement in a specific town. When reading and taking notes, I noticed that the Prohibition Party was important in local elections and that temperance was debated by members of the local branch. My primary interest was in the internal structure, ideology, and growth of the Knights movement. Temperance was a new and unexpected category. I coded the notes with the label "temperance" and included it as a possible theme (also see Expansion Box 1, Themes and Coding Qualitative Data).

In their qualitative content analysis interview data on twenty adults with type 1 diabetes, Graneheim and Lundman (2003) describe the open coding process. The interviews had asked about various aspects of living with type 1 diabetes. The researchers read the interview transcripts several times to obtain a sense of the whole. They then extracted text about the participants' experiences of having hyperglycemia and brought together the relevant passages into one text. This constituted the unit of analysis. They divided the text into meaning units (i.e., the constellation of words or statements that relate to the same central meaning) and then

Open coding The first coding of qualitative data that examines the data to condense them into preliminary analytic categories or codes.

EXPANSION BOX 1

Themes and Coding Qualitative Data

"A good thematic code is one that captures the qualitative richness of the phenomenon. It is usable in the analysis, the interpretation, and the presentation of research" (Boyatzis, 1998:31). To code data into themes, a researcher first needs to learn how "to see" or recognize themes in the data. Seeing themes rests on four abilities: (1) recognizing patterns in the data, (2) thinking in terms of systems and concepts, (3) having tacit knowledge or in-depth background knowledge (e.g., it helps to know Greek myths to understand Shakespeare's plays), and (4) possessing relevant information (e.g., one needs to know a lot about rock musicians and music to code themes about a rock music concert) (see Boyatzis, 1998:7–8).

Three errors to avoid when coding (see Schwandt, 1997:17) are (1) staying at a descriptive level only (not being analytic), (2) treating coding as a purely mechanical process, and (3) keeping codes fixed and inflexible. Codes have five parts: (1) a one- to three-word label or name, (2) a definition with a main characteristic, (3) a "flag" description of how to recognize the code in the data, (4) any exclusions or qualifications, and (5) an example.

ILLUSTRATION OF FIVE PARTS

Label. Gender-role disputes.

Definition. Interpersonal verbal disagreements are an example as are conflicts or disputes over what is proper or acceptable behavior for males and females in their interactions together or separately because he or she is male or female.

Flag. An example is making sarcastic remarks or jokes, or having disagreements (very mild to angry arguments) over what a male or female should do because he or she is male or female.

Qualifications. Only disputes among same gendered persons are considered. Any type of behavior (verbal or nonverbal) can be the target of a dispute. Interactions among overtly homosexual and transgendered persons are not included.

Example. Outside a classroom, Sara and Jessica, 16 years old, discuss their dates last night. Sara says, "We went out for pizza—of course he paid." Jessica remarks, "Of course? You mean you expect the guy to pay?" Sara answers, "Oh, forget it."

condensed them. They abstracted the condensed meaning units and labeled each with a code.

Although we can begin coding with a list of concepts, we usually generate most coding themes while reading data notes. Regardless of whether we begin with a list of themes, we list themes *after* finishing the open coding. Such a list serves three purposes:

1. It helps to see the emerging themes at a glance.
2. It stimulates us to find themes in future open coding.
3. We can use the list to build a universe of all themes in the study, which we reorganize, sort, combine, discard, or extend in further analysis.

Axial coding A second stage of coding of qualitative data during which the researcher organizes the codes, links them, and discovers key analytic categories.

We vary in how completely and in how much detail to code. Some researchers code every line or every few words; others code paragraphs or pages. Some of the data are not coded and are dross, or left over. The degree of detail in coding depends on the research question, the "richness" of the data, and the research purposes (see Expansion Box 2, The Process of Coding Qualitative Data).

Open-ended coding extends to analytic notes or memos that you write to yourself while collecting data. You should write memos on your codes (see the later discussion of analytic memo writing).

Axial Coding. This is a "second pass" through the data. During open coding, you focus on the actual data and assigning code labels for themes. You are little concerned about making connections among themes or elaborating the concepts that the themes represent. In contrast, you begin **axial coding** with an organized set of initial codes or preliminary

EXPANSION BOX 2

The Process of Coding Qualitative Data

Coding qualitative data, whether it is in the form of observational field notes, video or audio recordings, open-ended interviews, or detailed historical documents, is a challenge despite attempts by Strauss (1987) and others to systematize and simplify the process, making it appear as a fixed three-step sequence with open, axial, and selective coding. Some researchers rely on text-coding software programs (see discussion later in this chapter) that force them to create codes, but the software is just one tool in a larger coding process.

Weston et al. (2001) described their coding process in detail. Weston worked as part of a six-person research team and noted that team collaboration helped to make coding processes more explicit. The ideal associated with grounded theory that a researcher begins with a completely open mind and without prior expectations is just that, an ideal. In reality, a person's academic training, awareness of concepts and theoretical assumptions, and expectations from the audience who will read the research report shape data coding. In Weston's study, the process began with one researcher on the team creating a coding system that had four codes based on a first reading of open-ended interview transcript data. The system had a definition for each coded idea and rules with examples for converting raw data into codes. Others on the research team then used the system to code selections of raw data. Based on experiences with this preliminary system, they revised the coding system and added subtypes of the original codes. The process was repeated several times with the team members individually coding raw data, meeting together to discuss coding, and revising the coding system. After months of coding and meetings, the initial four codes became three master concepts with two of the three containing two types and each type having four to seven more refined codes. This yielded thirty-four coding distinctions. Over the next 2 years, the research team applied the system to hundreds of pages of raw data. Team members continued the process of reflecting on codes, meeting to discuss coding, and refining the system. Eventually their coding system had four tiers—three master concepts, seven types under the master concepts, two subtypes within three of the seven types, and several refined codes within each of the subtypes. In total, they created fifty-eight codes.

Over the next 2 years, as they continued to examine the data and present findings to the scientific community, the team kept refining and adjusting the coding system. They were following a strategy of *successive approximation* (see later in this chapter). A few new codes emerged and the system's structure shifted a little, but 4 years into the project, after hundreds of hours of meetings and repeated passes through the raw data, the coding system stabilized. As you see, a coding system can be more than a way to code raw data. It offers a system of analysis that provides a structured interpretation. By the way, the research topic Weston et al. studied was improving university teaching, and the team's data were from detailed open-ended interviews with six professors gathered during one semester.

concepts. In this second pass, you focus on the initial coded themes more than on the data. Additional codes or new ideas may emerge during this pass, and you should note them, but your primary task is to review and examine initial codes. You move toward organizing ideas or themes and identify the axis of key concepts in analysis.

Miles and Huberman (1994:62) have warned:

> *Whether codes are created and revised early or late is basically less important than whether they have some conceptual and structural order. Codes should relate to one another in coherent, study-important ways; they should be part of a governing structure.*

While axial coding, you ask about causes and consequences, conditions and interactions, strategies and processes. You look for categories or concepts that cluster together. You should ask questions such as: Can I divide existing concepts into subdimensions or subcategories? Can I combine several closely related concepts into one more general construct? Can I organize categories into a sequence (i.e., A, then B, then C), or by their physical location

(i.e., where they occur), or their relationship to a major topic of interest?

For example, when studying working-class life, LeMasters could have divided the general issue of marriage into subparts (e.g., engagement, weddings). He could mark all notes involving parts of marriage and then relate marriage to themes of sexuality, division of labor in household tasks, views on children, and so on. When the theme reappeared in different places, he could have made comparisons to see new themes (e.g., men and women have different attitudes toward marriage).

In the example of historical research on the Knights of Labor, I looked for themes related to temperance. I looked for discussions of saloons, drinking or drunkenness, and relations between the movement and political parties that supported or opposed temperance. Themes that clustered around temperance included drinking as a form of recreation, drinking as part of ethnic culture, different religious views on drinking, and gender differences regarding drinking.

Graneheim and Lundman (2003) used a process of axial coding in their study of interview data on diabetes. They compared codes based on differences and similarities and sorted them into six subcategories and three categories. The two researchers discussed tentative categories and revised them. A process of reflection and discussion resulted in agreement about how to sort the codes. Finally, the researchers identified underlying meaning—that is, the latent content—of the categories that they formulated into a broader theme.

Axial coding not only stimulates thinking about linkages between concepts or themes, but also raises new questions. It can suggest dropping some themes or examining others in more depth. It also reinforces the connections between evidence and concepts. As you consolidate codes, you may find evidence in many places for core themes and build a dense web of support in the qualitative data for them. This is analogous to the idea of multiple indicators described with regard to reliability and measuring variables. The connection between a theme and data is strengthened by multiple instances of empirical evidence.[5]

When I studied the Knights of Labor, I made the movement's failure to form alliances with other political groups a major theme. I reviewed notes looking for compromise and conflict between the Knights and other political parties, including temperance groups and the Prohibition Party. The array of concepts and themes related to temperance in axial coding helped me to see how the temperance issue facilitated or inhibited alliances.

Selective Coding. By the time you are ready for this last pass through the data, you have identified the major themes. **Selective coding** involves scanning all the data and previous codes, looking selectively for cases that illustrate themes, and making comparisons after most or all data collection has been completed. Selective coding should begin after concepts have been well developed and several core generalizations or ideas have been identified.

For example, as LeMasters studied working-class life in a tavern, he decided to make gender relations a major theme. In selective coding, he could have gone through his field notes, looking for differences in how men and women talked about dating, engagements, weddings, divorce, extramarital affairs, or husband/wife relations. He could then compare male and female attitudes on each part of the theme of marriage.

Graneheim and Lundman (2003) may have used selective coding in their study of interview data on diabetes. They provided readers of their study examples of codes, subcategories, categories, and a theme taken from text narratives about hyperglycemia, offering very explicit examples of each.

During selective coding, major themes or concepts ultimately guide the search process. You reorganize specific themes identified in earlier coding and elaborate more than one major theme. For example, in the working-class tavern study, LeMasters could have examined opinions on marriage to understand both the theme of gender relations and the theme of different stages of the life cycle.

Selective coding The last stage in coding qualitative data that examines previous codes to identify and select data that will support the conceptual coding categories that were developed.

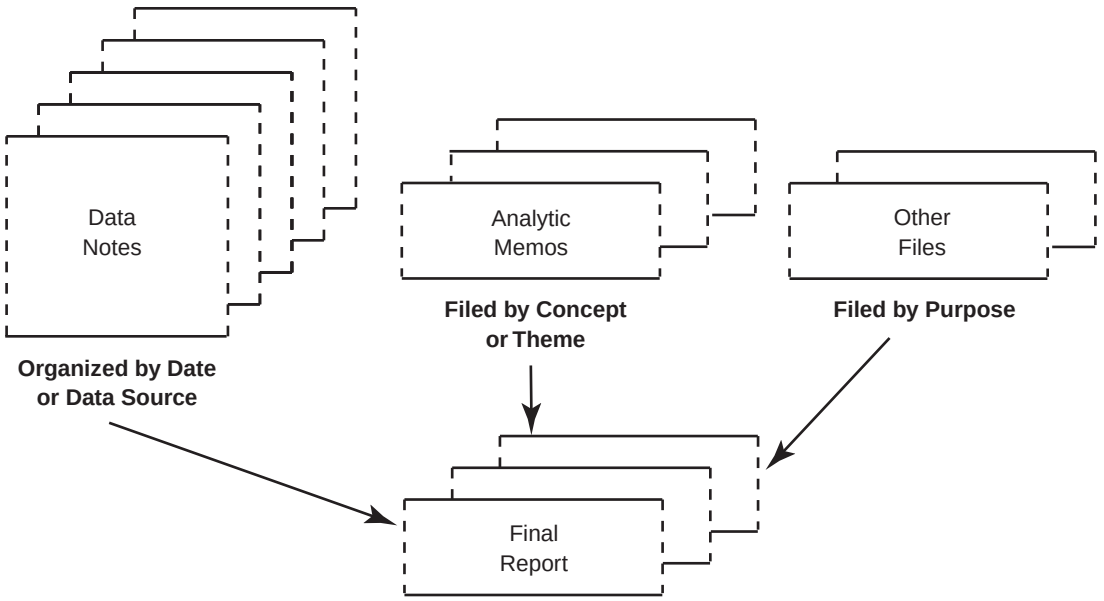


FIGURE 1 Analytic Memos and Other Files

Likewise, in the Knights of Labor study, I used temperance to understand the major theme of failed alliances and to understand another theme, sources of division within the movement that were based on ethnic or religious differences among members.

Analytic Memo Writing

In qualitative research, you are always writing notes. You record data in notes, write comments on method or research strategy in notes, and so on. You need to be a compulsive note taker, keep notes organized in files, and create many files with different subjects of the notes: a file on methodological issues (e.g., locations of sources or ethical issues), a file of maps or diagrams, a file on possible overall outlines of a final report or chapter, a file on specific people or events, and so on.

The analytic memo is a special type of note.[6] It is a memo or discussion of thoughts and ideas about the coding process that you write to yourself. Each coded theme or concept forms the basis of a separate memo. The memo is a discussion of the concept or theme. Rough theoretical ideas form the beginning of analytic memos.

The analytic memo links concrete data or raw evidence to abstract, theoretical thinking (see Figure 1). It contains your reflections on and thinking about the data and coding. Add to the memo and use it as you pass through the data with each type of coding. The memos form the basis for analyzing data in the research report. In fact, rewritten sections from good-quality analytic memos can become sections of the final report.

The tools involved in writing analytic memos are simple: pen and paper, a few notebooks, a stack of file folders, and photocopies of notes. Some researchers use computers, but it is not necessary. There are many ways to write analytic memos; you should develop your own style or method. See Expansion Box 3, Suggestions for Analytic Memo Writing, for concrete suggestions based on the experience of others. Some researchers make multiple copies of notes and then cut them and place various copies into an analytic memo file. This works well if the physical files are large and analytic memos are kept distinct within the file (e.g., on different-colored paper or placed at the beginning of the file). Other researchers list within the analytic memo file locations in the data notes where

EXPANSION BOX 3

Suggestions for Analytic Memo Writing

1. Start to write memos shortly after you begin data collection, and continue memo writing until just before the final research report is completed.
2. Put the date on memo entries so that you can see progress and the development of thinking. This will be helpful when rereading long, complicated memos because you will periodically modify memos as research progresses and add to them.
3. Interrupt coding or data recording to write a memo. Do not wait and let a creative spark or new insight fade away—write it down.
4. Periodically read memos and compare those on similar codes to see whether they can be combined, or whether differences between codes can be made clearer.
5. Keep a separate file for memos on each concept or theme. All memo writing on that theme or concept is kept together in one file, folder, or notebook. Label it with the name of the concept or theme so it can be located easily. It is important to be able to sort or reorganize memos physically as analysis progresses, so you should be able to sort the memos in some way.
6. Keep analytic memos and data notes separate because they have different purposes. The data are evidence. The analytic memos have a conceptual, theory-building intent. They do not report data but comment on how data are tied together or how a cluster of data is an instance of a general theme or concept.
7. Refer to other concepts within an analytic memo. When writing a memo, think of similarities to, differences between, or causal relationships with other concepts. Note these in the analytic memo to facilitate later integration, synthesis, and analysis.
8. If two ideas arise at once, put each in a separate memo. Try to keep each distinct theme or concept in a separate memo and file.
9. If nothing new can be added to a memo and you have reached a point of saturation in getting any further data on a theme or concept, indicate that in the memo.
10. Develop a list of codes or labels for the memos. This will let you look down the list and see all of the themes of memos. When you periodically sort and regroup memos, reorganize this list of memo labels to correspond to the sorting.

Sources: Adapted from Miles and Huberman (1994:72–76), Lofland and Lofland (1995:193–194), and Strauss (1987:127–129). Also see Lester and Hadden (1980).

a theme appears. Then it is easy to move between the analytic memo and the data. Because data notes contain highlighted or marked themes, it is easy to find specific sections in the data. An intermediate strategy is to keep a running list of locations where a major theme appears in the data and to include copies of a few key sections of the notes for easy reference.[7]

As you review and modify analytic memos, discuss ideas with colleagues and return to the literature with a focus on new issues. Analytic memos may help to generate potential hypotheses, which you can add and drop as needed. These notes also help you develop new themes or modify coding systems.

Outcropping An aspect of qualitative data analysis that recognizes some event or feature as representing deeper structural relations.

Outcroppings

The specific empirical evidence we gather is related to theoretical ideas and structures that are beneath observable reality. The relationship, modeled in Figure 2, shows that data are only samples of everything that happens on the visible, surface level. We use the data to generate and evaluate theories and generalizations and simultaneously assume that beneath the outer surface of reality lie deeper social structures or relationships.

The surface reality that we can easily see only partially reflects what goes on unseen, beneath the surface. To use a term from geology, events on the surface are **outcroppings**.[8] In geology, an outcropping is the part of bedrock that is exposed on the surface for people to see. It is the outward manifestation of central, solid features of the land. Geologists study outcroppings to get clues about what lies beneath the surface.

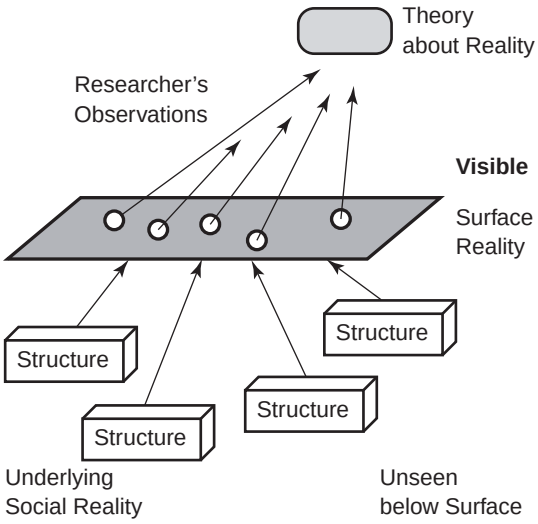


FIGURE 2 Theory, Surface Reality, and Underlying Structures

Often we cannot directly observe features of the social world. We cannot observe a deep loving relationship between two people. We can see its outward manifestation only in a kiss, specific deeds of affection, and acts of kindness. Likewise, we cannot directly observe a social structure such as social class. Nonetheless, we see its outward signs in differences in how people act, their career assumptions, their material possessions, and so forth. Sometimes we are misled by outward observation. We analyze data for both the surface level of reality and the deeper structures and forces that may lie unseen beneath the surface.

ANALYTIC STRATEGIES FOR QUALITATIVE DATA

Most qualitative researchers use techniques of coding, memo writing, and looking for outcroppings to some degree. This section introduces you to seven strategies you can use to analyze qualitative data: (1) ideal type, (2) successive approximation, (3) illustrative method, (4) domain analysis, (5) analytic comparison, (6) narrative analysis, and (7) negative case method.

As stated earlier in this chapter, strategies for qualitative data are more diverse, less standardized, and less explicit than in quantitative research. As Mahoney (1999:1192–1193) noted, "The absence of methodological explicitness has made it difficult for many readers to fully understand and appreciate the arguments of [qualitative data] researchers." Some researchers use only one strategy whereas others combine several.

In general, *data analysis* means a search for patterns in data—recurrent behaviors, objects, phases, or ideas. Once you identify a pattern, you need to interpret it in terms of a social theory or the setting in which it occurred. This allows you to move from the particular description of a historical event or social setting to a more general interpretation.

Data take many forms in qualitative research. For example, field research data include raw sense data that a researcher experiences, recorded data in field notes, and selected or processed data that appear in a final report (see Figure 3). Data analysis involves examining, sorting, categorizing, evaluating, comparing, synthesizing, and contemplating the coded data as well as reviewing the raw and recorded data.

Ideal Types

One of the most common strategies of qualitative data analysis is Max Weber's *ideal type*. It is a model or mental abstraction of social relations or processes. Ideal types are pure standards against which the data or "reality" can be compared. An ideal type is an artificial device used for comparison because no reality ever fits an ideal type. For example, I develop a mental model of the ideal democracy or an ideal college beer party. These abstractions with lists of characteristics do not describe any specific democracy or beer party; nevertheless, they are useful when applied to many specific cases to see how well each case measures up to the ideal.

Weber's method of ideal types also complements Mills' method of agreement (see analytic comparison). The method of agreement focuses attention on what is common across cases and looks for common causes in cases with a common outcome. By itself, the method of agreement implies a comparison against actual cases. This comparison of cases could also be made against an idealized model. You could develop an ideal type of

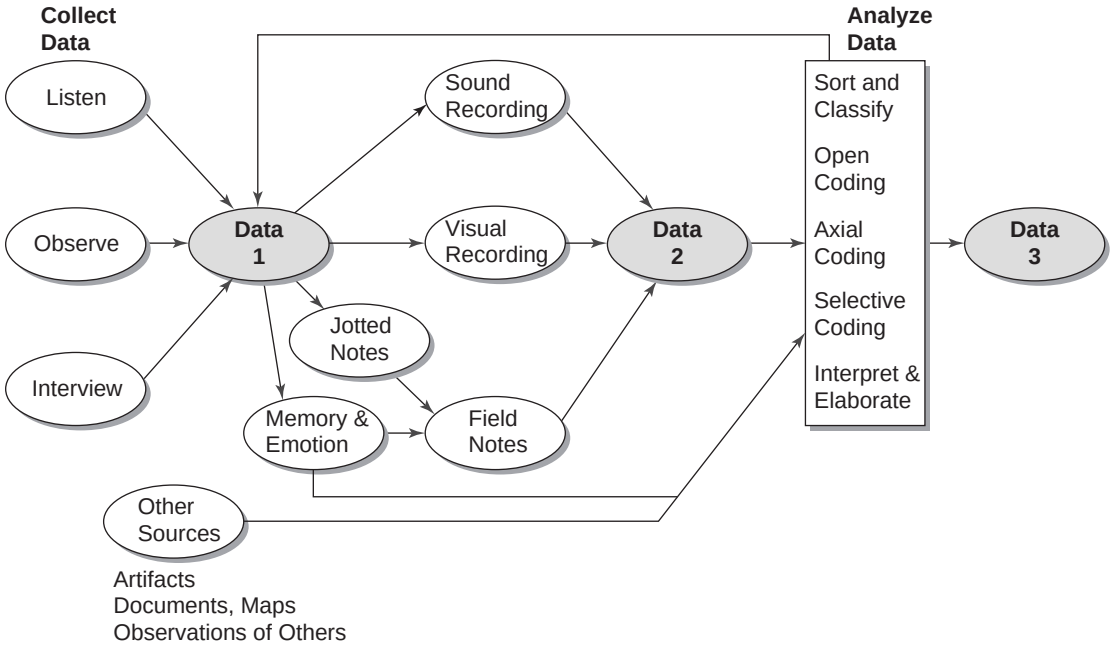


FIGURE 3 Data in Field Research (Data 1 = Raw sense data, experiences of researcher; Data 2 = Recorded data, physical record of experiences; Data 3 = Selected, processed data in a final report)
Source: Adapted from Ellen (1984a:214).

a social process or relationship and then compare specific cases to it.

In qualitative research, we can use ideal types in two ways: contrast contexts and analogy.

1. *Contrast contexts.* Researchers who adopt a strongly interpretive approach may use ideal types to interpret data in a way that is sensitive to the context and cultural meanings of members. Rather than develop hypotheses or create a generalizable theory, they use the ideal type to bring out the specifics of each case and to emphasize the impact of the unique context.[9] As they contrast between contexts, they may choose cases with dramatic contrasts or distinctive features. For example, in *Work and Authority in Industry*, Bendix (1956) compared management relations in very different contexts, Czarist Russia and industrialized England. When comparing contexts, some researchers do not use the ideal type to illustrate a theory in different cases or to discover regularities. Instead, they accentuate the specific and the unique. In contrast, others use ideal types to show how unique features shape the operation of general processes. As Skocpol and Somers (1980:178) explained:

> *"Above all, contrasts are drawn between or among individual cases. Usually such contrasts are developed with the aid of references to broad themes or orienting questions or ideal type concepts. Themes and questions may serve as frameworks for pointing out differences among cases. Ideal types may be used as sensitized devices—benchmarks against which to establish the particular features of each case."*

You might use the ideal type to show how specific circumstances, cultural meanings, and the perspectives of specific individuals are central for understanding a social setting or process. The ideal type becomes a foil against which you can highlight unique contextual features.

2. Analogies. We can also use ideal types as analogies to organize qualitative data. An *analogy* is a statement that two objects, processes, or events are similar to each other. We use it to communicate ideas and to facilitate logical comparisons. Analogies transmit information about patterns in data by referring to something that is already known or an experience familiar to the researcher. Analogies can describe relationships buried deep within many details. They are a shorthand method for seeing patterns in a maze of specific events. Making comparison of social processes across different cases or settings are easier.[10] For example, you might say a room went silent after person X spoke and "a chill like a cold gust of air" spread through it. This does not mean that the room temperature dropped or that a breeze was felt, but it succinctly expresses a rapid change in emotional tone. Likewise, you could report that gender relations in society Y were such that women were "viewed like property and treated like slaves." This does not mean that the legal and social relations between genders were identical to those of slave owner and slave. It implies that an ideal type of a slave-and-master relationship would show major similarities to the evidence on relations between men and women if applied to society Y. Ideal type analogies operate as heuristic devices (i.e., a device that helps one learn or see). Analogies are especially valuable when you try to make sense of or explain data by referring to a deep structure or an underlying mechanism.[11] Ideal types do not provide a definitive test of an explanation. Rather, they guide the conceptual reconstruction of the mass of details into a systematic format.

Successive Approximation

Successive approximation is a process that involves making repeated iterations. You cycle through steps, moving toward a final analysis. Over time, or after several iterations, you move from vague ideas and concrete details in the data toward a comprehensive analysis with generalizations. This is similar to coding discussed earlier. You begin with research questions and a framework of assumptions and concepts. You then probe into the data, asking questions of the evidence to see how well the concepts fit the evidence and reveal features of the data. You also create new concepts by abstracting from the evidence and adjusting concepts to fit the evidence better. You then collect additional evidence to address unresolved issues that appeared in the first stage and then repeat the process. At each stage, the evidence and the theory shape each other. The process is called successive approximation because the modified concepts and the model approximate the full evidence and are modified repeatedly to become successively more accurate.

Each pass through the evidence is provisional or incomplete. The concepts are abstract, but they are rooted in the concrete evidence and reflect the context. As the analysis moves toward generalizations that are subject to conditions and contingencies, you can refine generalizations and linkages to reflect the evidence better.[12]

The Illustrative Method

Another method of analysis anchors or illustrates theoretical concepts with empirical evidence. The **illustrative method** applies theory to a concrete historical situation or social setting and organizes data based on theory. Preexisting theory can provide conceptual **empty boxes** that you fill with the empirical evidence.[13] Evidence in the boxes confirms, modifies, or rejects the theory, which can be in the form of a general model, an analogy, or a sequence of steps (see Expansion Box 4, Three Variations of the Illustrative Method).[14]

A single case study with the illustrative method is not a strong test or verification of an explanation because data from one case can illustrate empty boxes from several competing explanations. In addition,

Successive approximation A method of qualitative data analysis that repeatedly moves back and forth between the empirical data and the abstract concepts, theories, or models, adjusting theory and refining data collection each time.

Illustrative method A method of qualitative data analysis that takes theoretical concepts and treats them as empty boxes to be filled with specific empirical examples and descriptions.

Empty boxes The conceptual categories in an explanation used as part of the illustrative method.

EXPANSION BOX 4

Three Variations of the Illustrative Method

1. *Case clarification.* A theoretical model used to illuminate or clarify a specific case or single situation, making the case more understandable by applying theory to it.
2. *Parallel demonstration.* Juxtapositioning of multiple cases (i.e., units or periods) to show that the same theory holds across multiple cases. Paige (1975) used parallel demonstration in a study of rural class conflict. He first developed an elaborate model of conditions that cause class conflict and then provided evidence to illustrate it from Peru, Angola, and Vietnam.
3. *Pattern matching.* This method matches the observations from one case with the pattern or concepts derived from theory or other studies. It allows for partial theory falsification; it narrows the range of possible explanations by eliminating some ideas, variables, or patterns from consideration.

finding evidence to illustrate an empty box using one case does not build a generalized explanation, which requires evidence from numerous cases.

Domain and Scheme Analysis

Cognitive anthropology, which studies relations between human culture and thought, has contributed greatly to qualitative data analysis. It treats cultures as mental creations or the cognitive organization of the physical, material world. Cognitive anthropologists study how people understand and organize material objects, events, and experiences. They note that we make sense of reality based on cognitive categories and that we order events, material life, and ideas based on cultural categories. Cognitive anthropology seeks to discover and document the rules of behavior or logical systems of thought that we use. To do this, it outlines what people see as culturally expected or appropriate in various situations more than what people actually do. Cognitive anthropology is part of a broader type of data analysis and theorizing that identifies how people or institutions classify and categorize the world, often implicitly. Such classifications then "take on a life of their own" to organize human experience (see Bowker and Leigh-Star, 1999).

Domain analysis A method of qualitative data analysis that describes and reveals the structure of a cultural domain.

Cultural domain A cultural setting or site in which people regularly interact and develop a set of shared understandings or "miniculture" that can be analyzed.

Early cognitive anthropologists asked people to arrange colors and plants into categories or to organize relatives into kinship systems. This helped the anthropologists to discover the organizing principles that underlie human social behavior. Cognitive anthropology evolved from studies in the 1950s–1960s (called "ethnoscience") to the study of "folk models" or "domains" in the 1970s, and it later evolved to "scheme analysis." *Schemas* are abstract entities and unconscious models of the world that we use to organize experience. In scheme analysis, we do not view the parts of a culture as either material or symbolic; rather we see culture as being composed of many parts. These parts are not static or integrated into a single whole; instead, we apply schemes to organize the parts. Schemes are cognitive units, such as prototypes, propositions, and cognitive categories. We can analyze the parts of culture to see whether they are shared, examine how they are distributed across people, and look for how the schemas relate to behaviors.

The anthropologist Spradley (1979a, 1979b) developed **domain analysis**. We will examine this system for qualitative data analysis in this section. For Spradley, the basic unit in a cultural setting is a **cultural domain**, an organizing idea or concept. The data analysis system focuses on analyzing domains. Later we can combine domains into taxonomies and broader themes that provide us an interpretation of a cultural scene or social setting.

Cultural domains have three parts: a cover term, included terms, and a semantic relationship. The *cover term* is simply the domain's name. *Included terms* are the subtypes or parts of the domain. A *semantic relationship* tells how the included terms fit logically within the domain. For

CHART 1 Forms of Relationships in Cultural Domains

SEMANTIC RELATIONSHIP	EXAMPLE OF USE
Is a type of	A bus *is a type of* motor vehicle [types of vehicles].
Is a part of/is a place in	A tire *is a part of* a car [parts of cars].
Is a way to	Cheating *is a way to* get high grades in school [ways students get high grades].
Is used for	A train *is used for* transporting goods [ways to transport goods].
Is a reason for	High unemployment *is a reason for* public unrest [reasons for public unrest].
Is a stage of	The charge *is a stage of* a battle [stages of battle].
Is a result of/ is a cause of	A coal power plant *is a cause of* acid rain [causes of acid rain].
Is a place for	A town square *is a place for* a mob to gather [places where mobs gather].
Is a characteristic of	Wearing spiked, colored hair *is a characteristic of* punks [characteristics of punks].

example, in the domain of a witness in a judicial setting, the cover term is *witness*. Two subtypes or included terms are *defense witness* and *expert witness*. The semantic relationship is "is a kind of." Thus, an expert witness and a defense witness are both types of witnesses. Other semantic relationships are listed in Chart 1.

Spradley developed domain analysis by analyzing the argot of members in ethnographic field research, although we can extend it to other qualitative data. For example, Zelizer (1985) studied the changing social value of children by examining documents on attitudes and behaviors toward a child's death in the late nineteenth century. She could have used a domain analysis in which "attitude toward child's death" was a domain, and the statements of various attitudes she discovered in documents were included terms. The attitudes could be organized by the semantic relationship "is a type of." Spradley identified three types of domains: folk domains, mixed domains, and analytic domains.

1. A **folk domain** contains terms from the argot of the members in a social setting. To use it, you pay close attention to language and usage. The domain uses the relationship among terms from a subculture's argot or in the language of historical actors to identify cultural meaning.
2. A **mixed domain** contains folk terms, but you add your own concepts. For example, types of runners are named by the terminology of runners (e.g., long-distance runner, track people), but you observe other types of people for whom no term exists in the argot and assign them labels (e.g., infrequent visitors, newcomers, amateurs).
3. An **analytic domain** contains terms from the researcher and social theory. They are most helpful when the meanings in a setting are tacit, implicit, or unrecognized by participants. You infer meaningful categories and identify patterns from observations and artifacts and then assign terms to them.

You can construct domains from data notes by proceeding as follows: read your notes and look for common semantic relationships (e.g., is a type of place, is a type of person, is a type of feeling) in order to find the organization of social relationships. Next, identify a list of cover terms. For example, a witness in a judicial setting could be a cover term. Once you have a list of cover terms, you next organize the information from the notes as included

Folk domain A cultural area based on the argot and categories used by the people being studied in a field site.

Mixed domain A cultural area that combines the argot and categories of members under study with categories developed by a researcher for analysis.

Analytic domain A cultural area developed by a researcher using categories or terms that he or she developed to understand a social setting.

EXAMPLE BOX 1

Example of Domain Analysis Worksheet

1. Semantic relationship: Strict inclusion
2. Form: *X* (is a type of) *Y*
3. Example: An oak (is a type of) tree

INCLUDED TERMS	SEMANTIC RELATIONSHIP	COVER TERM
laundromat, hotel lobby motor box, orchard	is a type of →	flop
flophouse, under bridge box car, alley public toilet, steam grate		
Structural question: Would you call an alley a flop?		

INCLUDED TERMS	SEMANTIC RELATIONSHIP	COVER TERM
trusty, ranger bull cook, mopper	is a type of →	jail inmate
head trusty, lockup bullet man, sweeper lawn man, inmate's barber		
Structural question: Would you call a trusty a type of jail inmate?		

terms. Prepare a worksheet for each domain relationship. The worksheet contains the cover term, the list of included terms, and the semantic relationship (see Example Box 1, Example of Domain Analysis Worksheet).

Next you locate your examples of the domain relationship from your notes. The analysis proceeds until all relevant domains have been identified. You then organize the domains by comparing their differences and similarities. Finally, reorganize domains into typologies or taxonomies and reexamine the domains to create new, broader ones that include other domains as included terms (see Expansion Box 5, Summary of Steps in Domain Analysis). The process builds up from specifics in the notes to an overall set of logical relationships.[15]

Analytic comparison Qualitative data analysis technique that uses the method of agreement and the method of difference to discover causal factors that affect an outcome among a set of cases.

Analytic Comparison

The British philosopher and theorist John Stuart Mill (1806–1873) developed a logic of comparison that is still widely used today. His method of agreement and method of difference form the basis for **analytic comparison**.[16] We can use the ideal type, successive approximation, the illustrative method, and domain analysis to examine qualitative data from a single case or from multiple cases; however, analytic comparison requires multiple cases. Analytic comparison uses a quasi-experimental approach that combines deductive with inductive theorizing. Basically, you identify many factors for a set of cases, sort through logical combinations of

EXPANSION BOX 5

Summary of Steps in Domain Analysis

Domain analysis formalizes six steps found in many types of qualitative data analysis:

1. Read and reread qualitative data notes that are full of details.
2. Mentally repackage the details into a few dozen organizing ideas.
3. Develop new ideas from the notes relying on subjective meanings or organizing ideas.
4. Look for relationships among the ideas and group them based on logical similarity.
5. Organize larger groups by comparing and contrasting the sets of ideas.
6. Reorganize and link the groups together into broader integrating themes.

factors, and compare them across cases. In certain ways, analytic comparison shares features with statistical reasoning more than with quantitative data analysis. It is even used with rational decision-making models, such that particular combinations of factors may make certain choices appear to be rational for people whereas other combinations do not.

Analytic comparison sometimes is called *nominal comparison* because the factors in the qualitative data are often at a nominal level of measurement but they can also be ordinal.[17] You organize data for a set of cases (often three to ten) into many mutually exclusive and exhaustive factors. When analytic comparison is formalized via a computer program (QCA for Qualitative Comparative Analysis, to be discussed later in this chapter), you construct what logicians and mathematicians call a *truth table.* A truth table contains all of the logically possible combinations of factors and outcomes among cases. This information is frequently organized as a chart (see Example Box 2, Example of Method of Agreement and Difference: Theda Skocpol's Theory of Revolution) that looks similar to a Guttman scale. Analytic comparison helps you identify the combination of factors, often measured at the nominal level, that are associated with outcomes among a small number of cases.

Ragin (1994b) contrasted case-oriented, analytic comparison with traditional variable-oriented statistical analysis. He noted that case-oriented comparison "sees cases as meaningful but complex configurations of events and structures, and treats cases as singular, whole entities purposefully selected" (p. 300). Analytic comparison involves qualitative data from a small number of cases and adopts an intensive (i.e., a great many in-depth details about a few cases) rather than an extensive (i.e., a few details about a great many cases) data analysis strategy. Moreover, explanation in analytic comparison tends to be interpretative or structural rather than nomothetic. Analytic comparison emphasizes the effect of particular configurations of conditions in cases or context. It allows different causal factors to produce an outcome and considers highly complex outcomes that have qualitative differences.[18]

Method of Agreement. The **method of agreement** focuses attention on what is common across cases. You establish that cases have a common outcome and then try to locate a common cause, although other features of the cases may differ. The method proceeds by a process of elimination. You eliminate features as possible causes if they are not shared across cases that have a common outcome. For example, you look at four cases. All four share two common features, but they also differ in many respects. You look for one or more common causes to explain the common outcome in all cases. At the same time, you eliminate alternative possibilities and identify a few primary causal factors so that you can argue that, despite the differences, the critical similarities exist.

Method of Difference. You can use the **method of difference** alone or in conjunction with the

Method of agreement A method of qualitative data analysis that compares characteristics that are similar across cases that share a significant outcome.

Method of difference A method of qualitative data analysis that compares characteristics among cases in which some share a significant outcome but others do not; focuses on the differences among cases.

EXAMPLE BOX 2

Example of Method of Agreement and Difference: Theda Skocpol's Theory of Revolution

	CAUSAL FACTOR		OUTCOME
CASE	*State Breakdown*	*Peasant Revolt*	*Revolution?*
France	Yes	Yes	Yes
Russia 1917	Yes	Yes	Yes
China	Yes	Yes	Yes
England	Yes	No	No
Russia 1905	No	Yes	No
Germany	No	No	No
Prussia	No	No	No
Japan	No	No	No

Source: Adapted from Mahoney 1999, Table 1.

method of agreement. The method of difference is usually stronger and is a "double application" of the method of agreement. First, locate cases that are similar in many respects but differ in a few crucial ways. Next pinpoint features in which a set of cases is similar with regard to an outcome and causal features and another set in which the cases differ on outcomes and causal features. The method of difference reinforces information from positive cases (e.g., cases that have common causal features and outcomes) by contrasting it with the negative cases (e.g., cases lacking the outcome and causal features). Thus, you look for cases that have many of the causal features of positive cases but lack a few key features and have a different outcome (see an example of analytic comparison in Example Box 3, Analytic Comparison to Study the Success and Failure of Homeless Organizations).

Narrative Analysis

Narrative, as well as the related idea of analyzing a sequence of events, has multiple meanings and is used in anthropology, archaeology, history, linguistics, literary criticism, political science, psychology, and sociology.[19] We encountered narrative regarding historical-comparative research in referring to a form of historical writing. In addition, *narrative* refers to a type of qualitative data, a form of inquiry and data gathering, a way to discuss and present data, a set of qualitative data analysis techniques, and a kind of theoretical explanation. As Griffin (1992a:419) observed, "Narrative is both a rhetorical form and a generic, logical form of explanation that merges theorized description of an event with its explanation."

Narratives as a way to examine the world have several features: a connected relationship among parts, a causal sequence of episodes to form a "plot," a selection that emphasizes important versus less important parts, and a specific mix of time and place. We use narratives for several purposes. They can address the issue of "who are we" as individual people, or they can be public narratives that link us to larger groups, communities, or nations. Some narratives describe social forces that act on us. Finally, *metanarratives* are overall frameworks with master ideas. They organize the thinking of entire populations for generations (e.g., the ideas of progress, industrialization, or globalization (see Somers and Gibson, 1994), Despite its many uses, a narrative shares core elements (see Expansion Box 6, Six Features of a Narrative).[20]

Next we briefly consider several types of narrative, and then turn to examine narrative analysis, a type of qualitative data analysis.

EXAMPLE BOX 3

Analytic Comparison to Study the Success and Failure of Homeless Organizations

Cress and Snow (1996) used analytic comparison to analyze field research data (1,500 pages of field notes) that they had gathered on fifteen social movement organizations to help homeless people in eight U.S. cities. They identified four general types of resources—moral, material, information, and human—that the movements could have. They measured a movement organization's resources by whether it had fourteen specific resources, at least two for each of the four types. For example, a specific moral resource was a public statement of support by an external organization, material support included supplies such as paper or telephone service, information support included people who were experienced at running meetings, and human support included individuals who volunteered time on a regular basis and followed orders.

The researchers classified whether the movement organizations were *viable* (seven were and eight were not), meaning that the organization had survived for one year or more during which meetings were held at least twice a month. They found that nine specific resources were necessary or the organization would fail, as well as combinations of the five other resources. The development of the fifteen organizations followed one of three "paths" based on the combination of the nine necessary and the five "other" resources.

As raw data, a narrative refers to text and practice in social life. Narratives are how people organize their everyday practices and subjective understandings. Narratives appear in oral or written texts to express understandings and the quality of lived experience. They are a form by which people construct identities and locate themselves in what is happening around them at the micro and macro levels.[21]

Narrative text refers to data in a storylike format that people apply to organize and express meaning and understandings in social life. "Schooling, clinics, counseling centers, correctional facilities, hospitals, support groups, and self-help organizations, among many other sites for storing experience, provide narrative frameworks for conveying personal experience through time" (Gubrium and Holstein, 1998:164). We find narratives in stories in novels, poems, myths, epic tales, dramatic performances, film, newspaper or media reports, sermons, oral histories, interviews, and the telling of events of a person's life. More than a form of expression, narrative is also a practice.

Narrative practice is the storylike form through which people subjectively experience and give meaning to their daily lives and their actions. The narrative organizes information, events, and experiences that flow across time. It offers a story line or plot from a particular point of view. The point of view is that of a motivated actor who expresses intentions. Because a narrative plot is embedded in a constellation of particular details, using it to make universal generalizations is difficult.

In a study of Caracas, Venezuela, Smilde (2003) emphasized the narrative he discovered in the beliefs of local Pentecostal churches. A local group of men used stories from the Pentecostal narrative to reinterpret their life experiences and it shaped their daily lives. The men adapted and used the narrative to reorganize their understandings of ongoing life events, and it gave a new coherence to

EXPANSION BOX 6

Six Features of a Narrative

1. It tells a story or tale (i.e., presenting unfolding events from a point of view).
2. It has a sense of movement or process (i.e., a before and after condition).
3. It contains interrelations or connections within a complex, detailed context.
4. It involves individuals or groups that engage in action and make choices.
5. It has coherence, that is, the whole holds together.
6. It has a temporal sequencing of a chain of events.

events. Thus, the narrative offered by the Pentecostal churches blended a religious conversion with a new self-understanding. The local men used it to reinterpret their past actions and guide their current activities. More than the telling of a story, the church narrative helped them to construct identity and find meaning in life.

Narrative inquiry is a method of investigation and data collection that retains a narrativelike quality from social life (Chase, 2005). Using it as inquiry, we try to capture people's ordinary lived experience without disrupting, destroying, or reducing its narrative character. The inquiry is self-reflective; that is, you place yourself in a flow of events and self-consciously become part of the "plot." The inquiry itself—engaging participant-observers in a field setting or examining historical-comparative documents—appears in narrative terms; that is, as a tale with a sense of movement and a coherent sequence of events about an engaged social actor in a specific context.

Narrative presentation grows out of the interpretative social science approach. Often called *storytelling* (Berger and Quinney, 2004), this mode of presentation blends description, empathetic understanding, and interpretation. It dissolves the space between a researcher and the people being studied. This makes the researcher an integral aspect of description, discussion, and interpretation in a study. Together, researcher and the researched coparticipate in creating/gathering data, and both reflect on the data. Such a process interweaves a researcher's life with the lives of the people being studied. As an individual social actor, the researcher becomes inseparable from the research process and from data presentation. For this reason, a researcher's personal biography and life situation are often included in the story format and in data presentation, discussion, and interpretation. Besides "giving voice" to the people who are studied, the researcher's voice, presence, and subjectivity appear. The storyteller-researcher is not a disembodied voice or detached observer; rather, he or she is a storyteller whose emotions, personal experiences, and life events become a part of the story that is told.

Narrative analysis, a method for analyzing data and providing an explanation, takes several forms. It is called *analytic narrative*, *narrative explanation*, *narrative structural analysis*, or *sequence analysis*.[22] Besides recognizing the core elements of a narrative (listed earlier), you may use narrative analysis techniques to map the narrative and give it a formalized grammar/structure. You can not only recognize the narrative character of social life but also analyze data in a manner that retains and unveils that character. The narrative is an outline or model for organizing data, but it also serves as a type of explanation.

Some researchers apply a few analytic concepts to qualitative data whereas others employ complex logical systems that outline the structure of a narrative, often with the aid of computer software. As you examine and analyze qualitative data for its narrative form and elements—whether it is an individual's life history, a particular historical event, the evolution of an organization over the years, or a macro-level historical process—you focus on events (rather than variables, individuals, or cases) and connections among them. You find that temporal features (e.g., order, pace, duration, frequency) are essential organizing concepts. You soon start to treat the sequence of events itself as an object of inquiry.

Franzosi (1998) argued that once we recognize narrative within data, we try to extract and preserve it without destroying its meaning-making ability or structure. We also look for what Abell (2004:293) called "action linkages"—that is, how a social actor engages in actions to transform one condition or situation into another or, simply put, makes things happen. As we map the structure of a narrative's sequence, the process operates as both a mode of data analysis and a type of explanation. It is an answer to this question: Why do events occur as they do? Some researchers believe that narrative explanations are not causal, but others believe

Narrative analysis Both a type of historical writing that tells a story and a type of qualitative data analysis that presents a chronologically linked chain of events in which individual or collective social actors have an important role.

narrative analysis is a causal explanation although perhaps involving a different type of causality, from that common in a traditional positivist science approach.[23]

Tools of Narrative Analysis. We next examine three analytic tools: path dependency, periodization, and historical contingency.

1. *Path dependency.* The way that a unique beginning can trigger a sequence of events and create a deterministic path is called **path dependency**. The path is apparent in a chain of subsequent events, constraining or limiting the direction of the ongoing events that follow. The outcome explained using path dependency is sensitive to events that occurred very early in the process. Path dependency explanations emphasize how the choices of one period can limit future options, shape later choices, and even accelerate events toward future crises in which options may be restricted.[24]

When building a path dependency explanation, start with an outcome. You then show how the outcome follows from a sequence of prior events. As you trace and demonstrate each event's effect on another, you go backward in the process to initial events or conditions. The initial conditions you identify are a "historical fork in the road" (Haydu, 1998:352).

Explanations that use path dependency assume that the processes that generated initial events (a social relationship) or institution may differ from the processes that keep it going. There may be one explanation for the "starting event" and another for the path of subsequent events. Researchers often explain the starting event as the result of a contingent process (i.e., a specific and unique combination of factors in a particular time and place that may never repeat). In addition, causal processes in one historical period may not operate in another. "There is no good reason to assume that findings from one period support causal claims for another period" (Haydu, 1998:345).

Path dependency comes in two forms: self-reinforcing and reactive sequence.[25] If you use a *self-reinforcing* path dependency explanation, you examine how, once set into motion, events continue to operate on their own or propel later events in a direction that resists external factors. An initial "trigger event" constrains, or places limits on, the direction of a process. Once a process begins, "inertia" comes into play to continue the process along the same path or track.

A classic example of inertia is the QWERTY pattern of letters on a keyboard. The pattern is inefficient. It takes longer for the fingers to hit keys than alternative patterns do, and it is difficult to learn. Engineers created QWERTY more than a century ago to work with early crude, slow, mechanical typewriters. They designed a keyboard pattern that would slow human typists to prevent the primitive machines from jamming. Later, mechanical typewriters improved and were replaced by electric typewriters and then by electronic keyboards. The old keyboard pattern was unnecessary and obsolete, but it continues to this day. The inertia to use an obsolete, inefficient system is strong. It overwhelms efforts to change existing machinery and people to a more rational, faster keyboard. Social institutions are similar. Once social relations and institutions are created in specific form (e.g., decentralized with many local offices), it is difficult to change them even if they are no longer efficient under current conditions.

The *reactive sequence* path dependency emphasizes a different process. It focuses on how each event responds to an immediately preceding one. Thus, instead of tracing a process back to its origins, it studies each step in the process to see how one influences the immediate next step. The interest is in whether the moving sequence of events transforms or reverses the flow of direction from the initial event. The path does not have to be unidirectional or linear; it can "bend" or even reverse course to negate its previous direction.

We can think of reactive sequence path dependency as a sequence of events that is like a pendulum; it swings back and forth. A single event may

Path dependency An analytic idea used in narrative analysis to explain a process or chain of events as having a beginning that triggers a structured sequence so that the chain of events follows an identifiable trajectory over time.

set into motion a reaction that changes or reverses the direction of the events that preceded it. For example, as part of the long process of the U.S. civil rights movement, the assassination of Martin Luther King Jr. triggered more vigorous civil rights law enforcement and an expansion of welfare programs. Events had been moving in the direction of increased social equality, reduced discrimination, and expanded legal rights, yet vigorous civil rights enforcement and welfare expansion disrupted existing status and power relations. This created tensions and triggered a backlash by resentful Whites. The White backlash sought to restrict or reverse civil rights law enforcement and cut back social welfare programs. Thus, a reaction to events in the sequence reversed the direction of its path.

2. *Periodization.* In historical-comparative research, we know that historical reality flows as discontinuous stages. To recognize this, researchers may use **periodization** to divide the flow of time in social reality into segments or periods. For example, we may divide 100 years of history into several periods. We break continuous time into several discrete periods that we define theoretically through periodization. Theory helps us to identify what is significant and what is common within periods or between different periods. As Carr (1961:76) remarked, "The division of history into periods is not a fact, but a necessary hypothesis." The breaks between periods are artificial; they are not natural in history, but they are not arbitrary.

You cannot determine the number and size of periods and the breaks between them until you have examined the evidence. You may begin with a general idea of how many periods are necessary to create and what distinguishes them, but you should adjust the number and size of the periods and the location of breaks after you examine the evidence. You may then reexamine the evidence with added data, adjust the periodization, and so forth. After several cycles of doing this, you get an approximate set of periods across 100 years based on successively theorizing and looking at evidence.

Periodization Dividing the flow of time in social reality into segments or periods; a field researcher might discover parts or periods in an ongoing process (e.g., typical day, yearly cycle).

Historical contingency An analytic idea in narrative analysis that explains a process, event, or situation by referring to the specific combination of factors that came together in a particular time and place.

3. *Historical contingency.* **Historical contingency** refers to a unique combination of particular factors or specific circumstances that may not be repeated. The combination is idiosyncratic and unexpected from the flow of prior conditions. As Mahoney (2000a:513) explained, "Contingency refers to the inability of theory to predict or explain, either deterministically or probabilistically, the occurrence of a specific outcome. A contingent event is therefore an occurrence that was not expected to take place." A contingent situation may be unexpected, but once it occurs, it can profoundly influence subsequent events. Because many possible idiosyncratic combinations of events occur, we use theory to identify important contingent events for an explanation.

A *critical juncture* is often a part of historical contingency (see Example Box 4, Path Dependency, Critical Junctures, and Historical Contingency). We use it to explain how several viable options may exist at a specific point in time. After one option is selected, many idiosyncratic events converge, which often has a powerful continuing influence. We can combine historical contingency and path dependency.

Roy (1997) combined historical contingency and path dependency to explain the rise of the large corporation in the United States. He argued the preexisting power relations among investors and government officials in the mid-nineteenth century did not cause the large private corporation to rise to prominence. Instead, a unique set of factors at a particular time and place favored its appearance (i.e., historical contingency). Once the institution of the large modern corporation appeared, it encouraged the ascendance of certain groups and fostered new power arrangements. These groups and arrangements then operated to maintain the corporate form of organization. An elite of financiers, wealthy investors, and executives rose in power and benefited from the private corporation form of business organization. They actively supported it through new laws, government regulations, financial relations, and other conditions. The corporate form sustained the growing power and

EXAMPLE BOX 4

Path Dependency, Critical Junctures, and Historical Contingency

Researchers combine the concepts of path dependency and conjunction in narrative analysis to discover how a specific short-term combination of circumstances can set subsequent events off along a new trajectory, and they try to identify these "critical junctures" or historical turning points. Kiser and Linton (2002) used this idea in their study of France from 1515 to 1789, and noted, "Particular historical turning points change the relationships between variables" (p. 905). They focused on rebellions against taxation in France. Tax revolts occurred in about 20 percent of the years 1515 to 1789. The taxes were primarily gathered to pay for ongoing wars (wars took place in 65 percent of the time period). The Fronde was a set of large-scale revolts (1648 to 1653) that the king's army successfully suppressed. Prior to the Fronde, tax increases and offensive wars regularly generated local revolts, but after it they very rarely did. The theoretical implication is that researchers may find that one set of causal relations are stable and operate for a time period but find little evidence for them in another period. Moreover, researchers might identify a specific event or short-term period that operates as a critical juncture or tipping point after which important relations dramatically shift and then begin to operate differently. It is a pattern of continuity along a path that is interrupted at a juncture and then is redirected to a new trajectory.

privilege of the elites. Thus, the "chance" convergence of particular events at one time selected one form of business organization among alternatives; it was not inevitable. However, once established, this business form set into motion new dynamics that perpetuated it into the future and altered surrounding conditions. It made alternative business forms less viable because it reinforced sociopolitical arrangements and realigned economic power in ways that undermined the alternatives. Thus, the corporate form of organization created a path along which the events that followed in time depended.

The path dependency may be self-reinforcing to continue with inertia along one direction, or particular events might set off a reaction that alters its direction. Along the flowing sequence of events across time, periodic critical junctures may occur. The process or conditions that were initially set into motion may resist change, or the contingent conditions may be powerful enough to trigger a major change in direction and initiate a new path of events.

Negative Case Method

We usually focus on what is evident in the data, yet we can also study what is *not* explicit in the data, or what did *not* happen. At first, studying what is not there may appear counterintuitive, but an alert observer who is aware of all clues notices what is missing as well as what is there. In the story "Silver Blaze," Sherlock Holmes solved a mystery when he noticed that a guard dog did not bark during the theft of an expensive racehorse, suggesting that the watchdog knew the thief. When what was expected did not occur, it was important information.

Negative evidence takes many forms (see Expansion Box 7, Types of Negative Evidence). It includes silences, absences, and omissions. For example, a field researcher notices that no one of a certain age, race, or gender is present in a social setting. This absence can be very revealing about the nature of the setting. Likewise, you notice some money lying on the floor, yet no one picks it up. The failure to pick it up can be an important clue. Perhaps in a historical-comparative study, you notice that there are no reports of a type of crime (e.g., hate crime, child abuse) in certain locations or times. You may find that the absence of reports or incidences can be equally important as their presence.

The **negative case method** is a way to systematically examine the absence of what is expected.[26] It combines the method of difference from analytic comparison with deviant case analysis. Deviant case analysis focuses attention on a few cases among a great many (including quantitative data sets) that do not conform to the general

Negative case method A qualitative data analysis that focuses on a case that does not conform to theoretical expectations and uses details from that case to refine theory.

EXPANSION BOX 7

Types of Negative Evidence

1. *Events that do not occur.* Some events are expected to occur on the basis of past experience, but do not. For example, research on the Progressive Era of U.S. history found that large corporations did not veto moderate labor reform legislation. Such a veto was expected after corporations had showed hostility toward labor for years. Instead, they actually encouraged the reform because it would quiet growing labor unrest.

   Likewise, nondecisions may occur when powerful groups do not participate directly in events because their powerful positions shape which issues arise. For example, a city has terrible air pollution, but there is no public action on the problem because "everyone" implicitly recognizes the power of polluting industry over jobs, tax revenue, and the community's economy. The polluting industry does not have to oppose local regulations over pollution because no such regulations are ever proposed.
2. *Events of which the population is unaware.* Some activities or events are not noticed by people in a setting or by researchers. For example, at one time the fact that employers considered a highly educated woman only for clerical jobs was not noticed as an issue. Until societal awareness of sexism and gender equality grew, few saw this practice as limiting the opportunities of women. Another example is that country-western song writers deny writing with a formula. Despite their lack of awareness, a formula is apparent through a content analysis of lyrics. The fact that members or participants in a setting are unaware of an issue does not mean that a researcher should ignore it or fail to look for its influence.
3. *Events the population wants to hide.* People may misrepresent events to protect themselves or others. For example, elites often refuse to discuss unethical behavior and may have documents destroyed or held from public access for a long period. Likewise, for many years, cases of incest went unreported in part because they violated such a serious taboo that incest was simply hushed up.
4. *Overlooked commonplace events.* Everyday, routine events set expectations and create a taken-for-granted attitude. For example, television programs appear so often in conversations that they are rarely noticed. Because most people have a television set and watch TV regularly, only someone who rarely watches television or who is a careful analyst may notice the topic. Or a researcher observes a historical period in which cigarette smoking is common. He or she may become aware only if he or she is a nonsmoker or lives in a period when smoking has become a public health issue.
5. *Effects of a researcher's preconceived notions.* Researchers must take care not to let their prior theoretical framework or preconceived notions blind them to contrary events in a social setting. Strong prior notions of where to look and what data are relevant may inhibit a researcher from noticing other relevant or disconfirming evidence. For example, a researcher expects violent conflict between drug addicts and their children and notices it immediately but fails to see that they also attempt to form a loving relationship.
6. *Unconscious nonreporting.* Some events appear to be insignificant and not worthy of being reported in the mind of a researcher, yet if detailed observations are recorded, a critical rereading of notes looking for negative cases may reveal overlooked events. For example, at first a researcher does not consider company picnics to be important. However, after rereading data notes and careful consideration, he or she realizes that they play an important symbolic role in building a sense of community.
7. *Conscious nonreporting.* Researchers may omit aspects of the setting or events to protect individuals or relations in the setting. For example, a researcher discovers an extramarital affair involving a prominent person but wishes to protect the person's good name and image. A more serious problem is a breach of ethics. This occurs when a researcher fails to present evidence that does not support his or her argument or interpretation of data. Researchers should present evidence that both supports and fails to confirm an interpretation. Readers can then weigh both types of evidence and judge the support for the researcher's interpretation.

Source: Lewis and Lewis (1980).

pattern. We use unusual cases to understand processes or generate new ideas.

Negative case methodology uses detailed knowledge of one particular case that does not conform to what would be expected based on a theory that has supporting evidence from many other cases. You use the single negative case to reexamine the theory, noticing lapses or problems in it. You can then apply insights from the negative case to revise the theory.

For example, Emigh (2003) observed that fifteenth century Tuscany, at the peak of the highly developed northern Italian Renaissance culture, had all preconditions predicted by major theories for producing a rapid "take off" to industrial capitalism: efficient agriculture, well-developed commercial manufacturing, no feudal nobility, a large urban economy, and a stable political organization. Yet it did not happen. Emigh asked why this was a negative case and gained an in-depth knowledge of the one such case. She then uncovered previously unknown factors (about local rural investment) that the major theories had failed to take into account. The types of analytic strategies used in qualitative analysis are summarized graphically in Figure 4.

OTHER TECHNIQUES

Qualitative research involves using many analysis techniques. Here we briefly consider other techniques to illustrate the variety.

Network Analysis

In qualitative research, we often "map" the connections among a set of people, organizations, events, or places. Using sociograms and similar mapping techniques, we can discover, analyze, and display sets of relations. For example, in a company, Harry gives Sue orders; Sue and Sam consult and help one another. Sam gets materials from Sandra. Sandra socializes with Mary. We find that networks help us see and understand the structure of complex social relations.[27]

Time Allocation Analysis

Time is an important resource in research. We examine the way people or organizations spend or invest time to reveal implicit rules of conduct or priorities. We document the duration or amount of time devoted to various activities. Qualitative research examines the duration or amount of time devoted to activities. An analysis of how people, groups, or organizations allocate the valuable resources they control (such as time, space, money, prestige) can reveal much about their real, as contrasted with officially professed, priorities. Often people are unaware of or do not explicitly acknowledge the importance of an activity on which they spend time. For example, you notice that certain people are required to wait before seeing a manager, but others do not wait. You may analyze the amount of time, who waits, what they do while waiting, and whether they feel waiting is just. Or you document that people say that a certain celebration in a corporation is not important. Yet everyone attends and spends 2 hours at the event. The collective allocation of 2 hours for the celebration during a busy week signals its latent or implicit importance in the culture of the corporation.[28]

Flowchart and Time Sequence

In addition to the amount of time devoted to various activities, we analyze the order of events or decisions. Historical researchers have focused on documenting the sequence of events, but comparative and field researchers also look at their flow or sequence. In addition to when events occur, we can use a decision tree or flowchart to outline the order of decisions to understand how one event or decision is related to others. For example, we can outline an activity as simple as making a cake (see Figure 5). Researchers applied the idea of mapping out steps, decisions, or events and investigating their interrelationship to many settings. For example, Brown and Canter (1985) developed a detailed flowchart for house-buying behavior. They divided it into fifty steps with a time line and many actors (e.g., involved

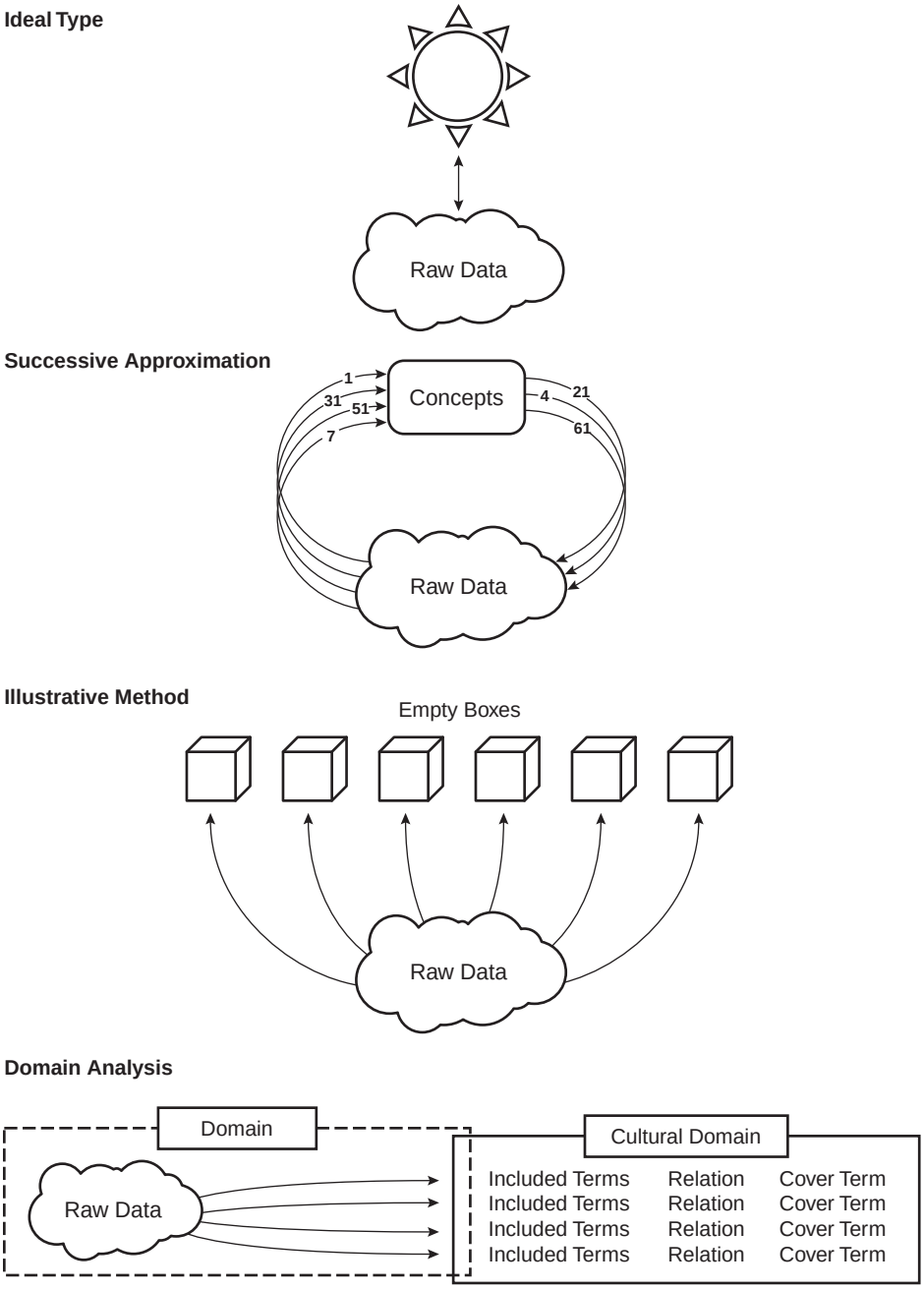
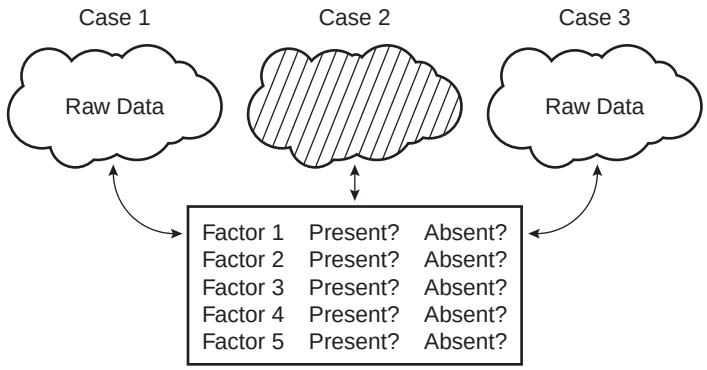
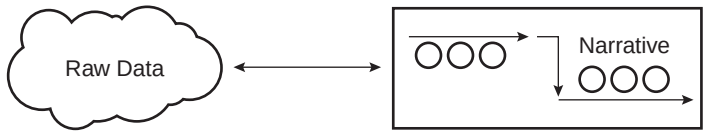


FIGURE 4 Summary of Analytic Strategies Used in Qualitative Data Analysis

Analytic Comparison



Narrative Analysis



Negative Case Method

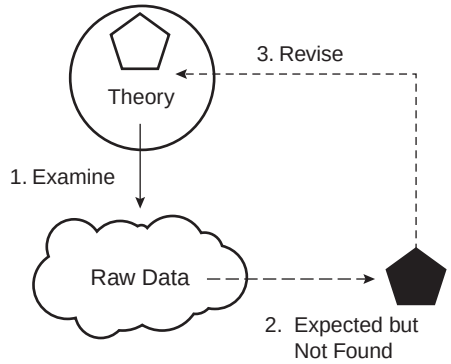


FIGURE 4 (Continued)

buyer, financial official, surveyor, buyer's attorney, advertising firm/realtor, seller, seller's attorney).[29]

Multiple Sorting Procedure

Multiple sorting is a technique similar to domain analysis found in field research or oral history. Its purpose is to discover how people categorize their experiences or classify items into what is similar or different. Cognitive anthropologists and psychologists often use a multiple sorting procedure. You can use multiple sorting to collect, verify, or analyze data. Here is how it works. You give the people you are studying a list of terms, photos, places, names of people, and so on, and ask them to organize the lists into categories or piles. They use categories of their

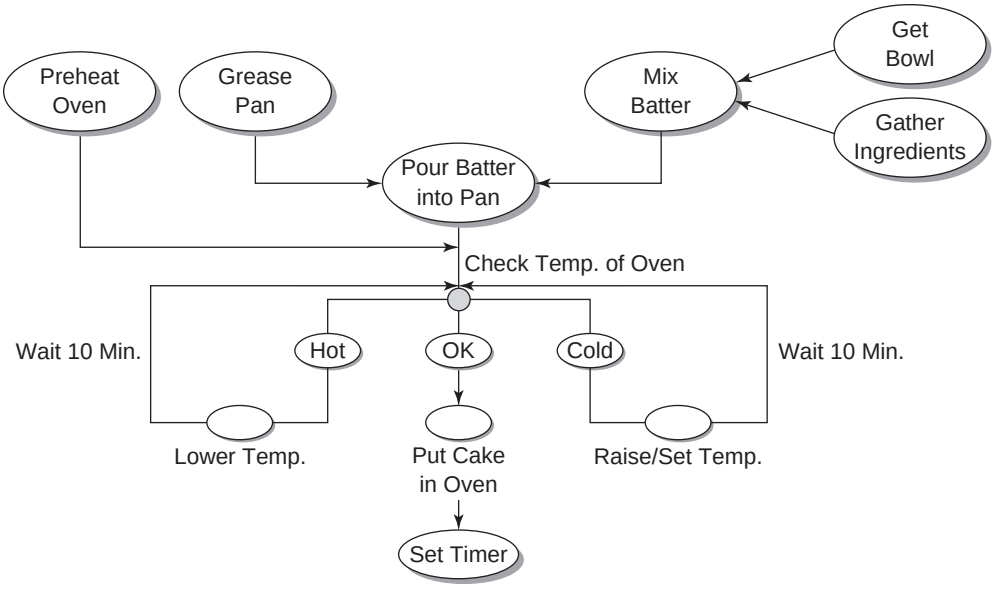


FIGURE 5 Partial Flowchart of Cake Making

own devising. Once sorted, you ask about the criteria used. You next give the people the items again and ask them to sort them in other ways that they may think of them. There is a similarity to Thurstone scaling in that people sort items, but here, the number of piles and types of items differ. The purpose of the sorting is not to create a uniform scale; rather, it is to discover how people understand the world. Canter et al. (1985:90) provide the example of a gambler who sorted a list of eight gambling establishments five times. Each sort had three to four categories. One of the sorts was organized based on "class of casino" (high to low). Other sorts were based on "frills," "size of stake," "make me money," and "personal preference." By examining the various sorts, you see how people organize their social reality.[30]

Diagrams

Qualitative research often presents data analysis as visual representations, such as diagrams and charts. Diagrams and charts help organize ideas and assist in systematically investigating data. They also communicate results to readers. We can use spatial or temporal maps, typologies, and sociograms. Thus, in a study of Little League baseball, Fine (1987) used sociograms to show the social relations among players. In addition to taxonomies, maps, and lists, we use flowcharts, organizational charts, causal diagrams, and various lists and grids to advance analysis and illustrate findings (see Figure 6).

Maps

Both quantitative and qualitative researchers place data on maps to help them see spatial relations and to supplement or reinforce results from other data analyses. For example, Ballen and Richardson (2002) used maps of France and United States to examine data on geographic patterns in suicide rates and to support theories of social integration and imitation from Émile Durkheim. Kiser and Linton (2002) presented a map of France with sites of rebellions marked in their study (discussed in Example Box 4). Villarreal (2002, 2004) used a map of Mexico in his study of violence and social-political change. In their study of differences in local hate crime law enforcement, McVeigh and colleagues (2003) offered a map of

EXAMPLE 1

Person	*Worked before College*	*Part-Time Job in College*	*Pregnant Now*	*Had Own Car*
John	Yes	Yes	N/A	No
Mary	Yes	DK	No	Yes
Martin	No	Yes	N/A	Yes
Yoshi	Yes	No	Yes	Yes

DK = don't know, N/A = not applicable

EXAMPLE 2

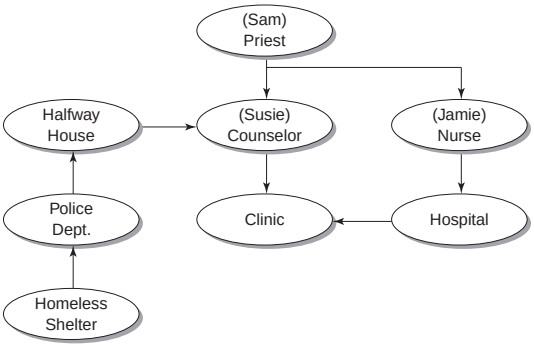


FIGURE 6 Examples of the Use of Diagrams in Qualitative Analysis

counties across the United States. Myers and Caniglia (2004) used a map of regions in the United States in a study of whether newspapers reported protest events. In their study of the endurance of distinct regional cultures, Griswold and Wright (2004) labeled areas of a U.S. map. Maps can be helpful in analyzing and presenting data to bolster an explanation; however, as a visual representation of information, they can also be misleading, so we should use them with care (see Monmonier, 1996).

Software for Qualitative Data

Since the mid-1960s, researchers have used computer technology to generate tables, graphs, statistical tests, and charts to analyze numerical data. By contrast, qualitative research has used computer technology only since the mid-1980s.[31] If you enter notes into a word processing program, you can quickly search for words and phrases. It is a small step to adapt such searching to data coding or linking codes to analytic memos. Word processing can also help you revise and move codes and parts of field notes.

Software has been specifically created for qualitative data analysis and new computer programs are continuously being developed or modified. Most come with highly detailed and program-specific user manuals, so the review here does not go into detail about specific software. It covers only the major approaches to qualitative data analysis at this time.

Text Retrieval. Some programs perform searches of text documents similar to the search function in word processing software. The specialized text retrieval programs are faster and have the capability of finding close matches, slight misspellings, similar sounding words, and synonyms. For example, if you look for the keyword *boat,* the program might also tell you whether any of the following appeared: *ship, battleship, frigate, rowboat, schooner, vessel, yacht, steamer, ocean liner, tug, canoe, skiff, cutter, aircraft carrier, dinghy, scow, galley, ark, cruiser, destroyer, flagship,* and *submarine.* In addition, some programs identify the combination of words or phases using logical terms (*and, or, not*) in what are called *Boolean searches* (named after George Boole, 1815–1864). For example, you may search long documents to identify where the keywords *college student, drinking,* and *smoking* occur within four sentences of one another and only when the word *fraternity* is not present in the block of text. This Boolean search uses *and* to seek the intersection of *college student* with either of the other two behaviors that are connected by the logical term *or,* whereas the logical search word *not* excludes situations in which the term *fraternity* appears.

Most programs show a keyword or phrase and the surrounding text. The programs may also permit you to write separate memos or add short notes to the text. Some programs count the keywords found and give their location. Most programs create a very specific index for the text based only on the terms of interest.

Textbase Managers. Textbase managers are similar to text retrieval programs. The key difference is their ability to organize or sort information about search results. Many programs create subsets of text data that help you compare and sort notes by a key

idea or to add factual information. For example, to detailed notes on interviews you can add the date and length of the interview, the gender of interviewee, the location of the interview, and so on. You can then sort and organize each interview or part of the interview notes using a combination of keywords and added information.

In addition, some programs have *hypertext* capability linking terms to other information, so when you click on one term it opens a new screen that has related information. You can identify keywords or topics and then link them to text. For example, in a field research study, you want to examine the person Susan and the topic of hair (including haircuts, hairstyles, hair coloring, and hats or hair coverings). You can use hypertext to connect all places that Susan's name appears to all discussions of hair. By clicking on Susan's name, one block of text quickly jumps to another in the notes, allowing you to see where Susan and the hair topic appear together.

Some text-based manager software creates cross-tabulation or scatterplot cross-classifications from information in text documents. For example, students keep journals on a course. They write their feelings about each day using one of four categories (boring, stimulating, challenging, or creative). The students also describe the major activities of each day (e.g., group work, discussion, videotape viewing, lecture, or demonstration). You can cross-classify student feelings by activity. By adding other information (e.g., male or female, academic major), you can learn whether students with different characteristics felt differently about the activities and see whether the feelings changed with the topic being presented in class.

Code-and-Retrieve Programs. We often assign codes or abstract terms to qualitative data (i.e., text field notes, interview records, and video or audiotape transcripts). Code-and-retrieve programs let us attach codes to lines, sentences, paragraphs, and blocks of text. The programs permit the use of multiple codes for the same data. In addition to attaching codes, most programs also help to organize the codes. For example, a program can help create outlines or "trees" of connections (e.g., trunks, branches, and twigs) among the codes, and among the data to which the codes refer. The program rearranges the text data based on the codes used and the relations among the indicated codes.

Qualitative comparative analysis (QCA) Qualitative data analysis and computer software based on Boolean logic that examines combinations of explanatory factors and various outcome measures to help a researcher identify complex, contingent causal relations.

Code-Based Theory Builders. Researchers using qualitative research are often interested in the evaluation and generation of theory. To do this, code-based theory builders require first assigning codes to the data. The programs provide ways for manipulating or drawing contrasts and comparisons among the codes. The relationships among the codes then become the basis for testing or generating a theory. The types of relations created among the codes may vary by program. A program may permit "if-then" logical relations. For example, Corsaro and Heise (1990) described how they coded field research data on young children into separate events. They then examined the logical sequence and relations among the events to search for principles or a "grammar" of implicit rules. They looked for rules that guided the sequencing, combination, or disconnection among events. The computer software ETHNO asks for logical connections among the events (e.g., time order, necessary precondition, co-occurrence) and then shows the pattern among events.

In contrast to other qualitative programs, code-based theory builders have a powerful ability to manipulate codes to reveal patterns or show relations in data that are not immediately evident. It becomes easier for researchers to compare and classify categories of data.

Qualitative comparative analysis (QCA) is an analytic strategy and type of software that uses Boolean logic or algebra.[32] Charles Ragin created QCA in 1987. An entire system of logical, mathematic-like relations has become the basis for computer software and digital electronics. It includes set theory, binary relations, logic gates, Venn diagrams, and truth tables. The logic's principle lets you organize concepts into sets. For example, when you search a computer database, you often include a keyword and Boolean operators *or, and,* and *not*.

QCA's strength is its ability to analyze multiple conjunctural causation. This implies that a combination of conditions produces the outcome, and different combinations of conditions may produce the same outcome. Depending on the context (conjuncture), a particular condition can have different impacts on the outcome. Together, QCA recognizes that different causal paths may yield the same outcome. QCA as a method of analysis and software works as an iterative process. It requires active engagement by the researcher. As Rihoux (2003: 354) describes it:

> *In a nutshell, the researcher must first produce a raw data table in which each case displays a specific combination of conditions (with 0 or 1 values) and an outcome (with 0 or 1 values). The software then produces a truth table that displays the data as a list of configurations. A configuration is a given combination of some conditions (each one receiving a 1 or 0 value) and an outcome (receiving a 1 or 0 value). A specific configuration may correspond to several observed cases, and different cases may display the same configuration. Then the key step of the analysis is Boolean minimization—that is, reducing the long Boolean expression (the long description that is expressed by the truth table) to the shortest possible expression (the minimal equation) that unveils the causal regularities in the data. It is then up to the researcher to interpret this minimal equation.*

QCA can help to analyze the characteristics of several cases and apply the method of difference and method of agreement. It performs the logical computations to identify common and unique characteristics among a set of cases. The algebra is not difficult, but it can be time consuming and subject to human error without the program (see Example Box 5, Example of QCA).

Conceptual Network Builders. This category of programs helps to build and test theory by presenting graphic displays or networks. The displays do more than diagram data; they help organize a researcher's concepts or thinking about the data. The programs use nodes, or key concepts, that the researcher identifies in data. They then show links or relationships among the nodes. Most programs give graphic presentations with boxes or circles that are connected by lines with arrows. The output looks similar to a flowchart diagram with a web or network of connections among concepts. For example, the data might be a family tree in which the relationships among several generations of family members are presented. Relations among family members (X is a sibling of Y, Z is married to Y, G is an offspring of X) can be used to discuss and analyze features of the network.

Event-Structure Analysis

Many qualitative researchers organize data chronologically in a narrative analysis. **Event-structure analysis (ESA)** is used to organize the sequence of events in ways that facilitate seeing causal relations. Researchers first used the method and ETHNO, a computer program used with it, to analyze field research data, but it can also be used for historical data. ESA first organizes the data into events and then places them into a temporal sequence.[33]

ESA facilitates narrative analysis. It helps to outline a set of links between events that happened. You separate what *had* to happen before other events from what *could* have happened. The computer program makes you answer questions about the logical relationships among events. For example, a situation has events A, B, C, X, and Y. You are asked: Must event A occur prior to X causing Y (i.e., is A a necessary precondition for the X:Y causal relationship?) or would X affect Y without A? If it is required, A must recur before X will again affect Y. This process forces you to explain whether the causal relation between two events is a unique and one-time relationship or a recurring relationship that can be repeated indefinitely or for a limited number of cycles.

Event-structure analysis has limitations. It does not provide the theory or causal logic; you must supply that. ESA creates only maps or diagrams (with the computer program) that make it easier for you to see relationships. When you decide about logically

Event-structure analysis (ESA) Qualitative data analysis often conducted with computer software that forces a researcher to specify the links among a sequence of many events; it clarifies causal relationships by asking whether one event logically had to follow another or just happened to follow it.

EXAMPLE BOX 5

Example of QCA

Roscigno and Hodson (2004) used QCA to analyze qualitative data from workplaces. They were interested in worker resistance including the collective response in the form of union activity and strikes as well as individualized forms such as sabotage, theft, and work avoidance. They asked whether grievances and resistance unfold as a function of workplace organization or are caused by interpersonal mistreatment on the shop floor. The data (eighty-two workplace ethnographies) represented the population of available ethnographic evidence on organizations. As the authors noted (p. 18), "There is good reason to expect that organizational structure and social relations condition one another and, thus, have contingent effects on worker grievances and resistance strategies."

QCA forced Roscigno and Hodson to specify and focus on variables deemed theoretically important. QCA had theoretical rigor, a case-oriented logic, and specification of potentially complex, conditional configurations. It helped to identify typologies that denote unique combinations of attributes in the data. By coding the qualitative data, the authors identified six workplace conditions as explanatory factors: bureaucracy, good organization, conflict, abuse, union presence, and a history of strikes. They created a "truth table" with each possible confirmation of the factors, each coded 1 = present, or 0 = absent. The configurations denoted the minimum number of factors needed to cover all positive, negative, and contradictory cases in the data. The researchers also identified six possible outcomes or forms of worker resistance: strikes, social sabotage, work avoidance, play dumb, absenteeism, and theft. Based on the various combinations revealed in QCA, the authors identified three types of workplaces each with a type of worker resistance: contentious workplaces, cohesive workplaces, and unorganized workplaces. Contentious workplaces had high levels of all forms of resistance, cohesive workplaces showed low levels of resistance, and unorganized workplaces largely had individual acts of resistance. See the following excerpt of the researchers' truth table. (Note that only the first three of thirty-six possible combinations are shown).

EXPLANATORY MEASURES

B = Bureaucracy. Workplace is bureaucratically organized with operational control of daily procedures in written rules.
G = Good Organization. There is coherence and integration of production practices.
C = Conflict. Ongoing conflict between workers and supervisors is common.
A = Abuse. Verbal, emotional, or physical abuse by supervisor of individual employees occurs.
U = Union Presence. Union representation exists in the particular workplace.
H = History of Strikes. Workplace has experienced strikes in the past.

RESISTANCE MEASURES

S = Strikes. There was a strike during the period of observation.
SS = Social Sabotage. There is undermining of superiors through mocking and ridicule.
AV= Work Avoidance. Avoiding work and/or work tasks occurs.
PD = Play Dumb. Workers pretend not to understand particular job tasks or organizational procedures.
AB = Absenteeism. Absenteeism is a response to workplace problems.
T = Theft. Stealing by workers while on the job takes place.

EXPLANATORY MEASURES (ALL POSSIBLE CONFIGURATIONS)						NUMBER OF CASES	RESISTANCE MEASURES WITHIN EACH CONFIGURATION					
B	G	C	A	U	H		S	SS	AV	PD	AB	T
0	0	1	1	0	0	2	0	2	2	1	2	0
1	0	0	1	1	0	4	1	2	3	1	3	1
1	1	1	1	1	0	1	0	0	0	0	0	0

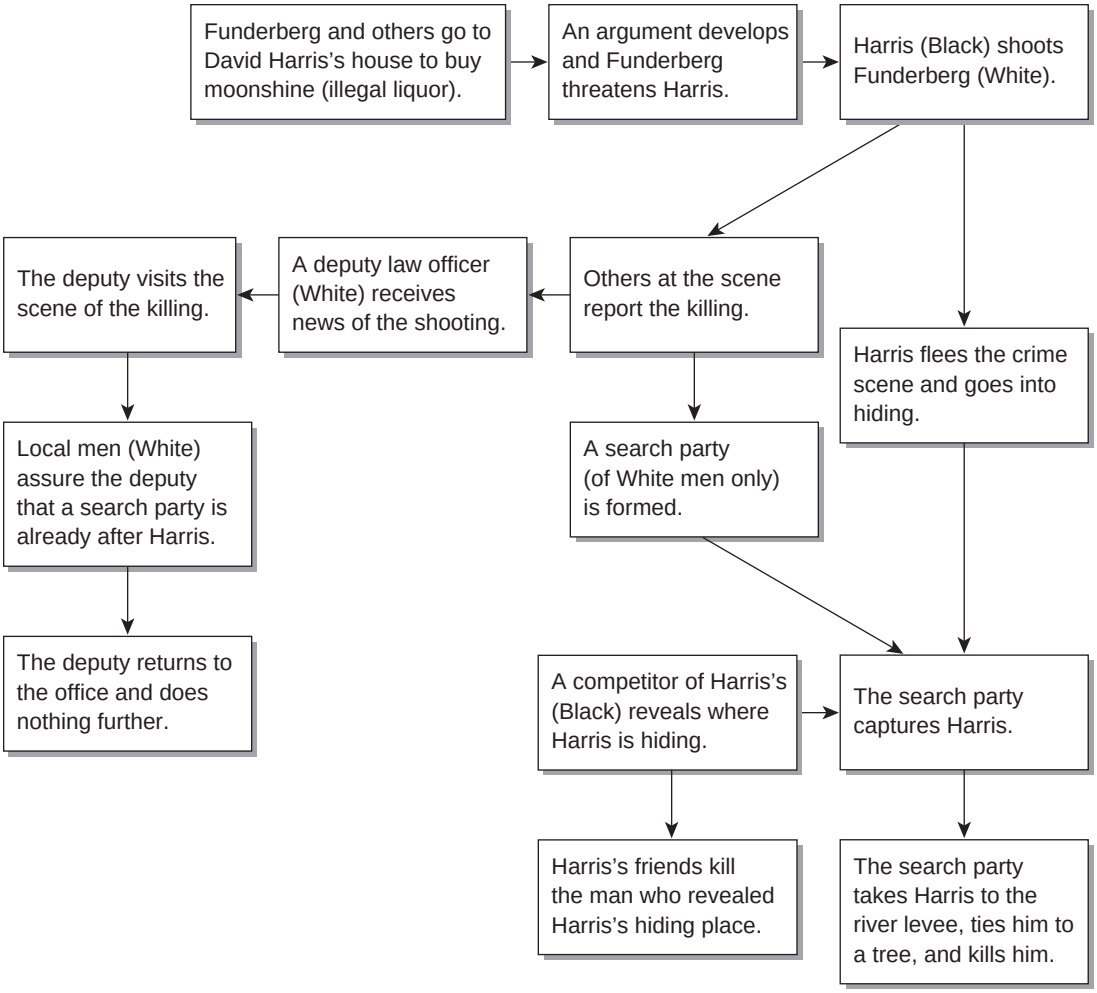


FIGURE 7 Example of Event-Structure Analysis of the Lynching of David Harris
Source: Adapted from Griffin (1993).

possible relations, ESA clarifies a chain of events and highlights those that might have been different. ESA does not have a place for enduring social structures that frame the action of event sequences.

Griffin's (1993) analysis of a lynching illustrates ESA. Based on many oral histories, a book, and newspaper reports, he reconstructed the sequence of events surrounding the lynching of David Harris in Bolivar County, Mississippi, in April 1930. After answering many yes/no questions about possible linkages among a series of events and analyzing the linkages, Griffin was able to conclude that the critical factor was the inaction of the local deputy who could have stopped the process. An abbreviated summary of the ESA diagram is presented in Figure 7.

CONCLUSION

This chapter discussed how we analyze qualitative data. In many respects, qualitative data are more difficult to deal with than data in the form of numbers. Numbers have mathematical properties that allow

us to use statistical procedures. Qualitative analysis requires more effort to read and reread data notes, reflect on what is read, and make comparisons based on logic and judgment.

Most forms of qualitative data analysis involve coding and writing analytic memos. Both are labor-intensive and time-intensive efforts. They require reading data carefully and thinking about them seriously. In addition, the chapter presented methods we used for the analysis of qualitative data. The techniques presented in this chapter are only a sample of the full range of qualitative data analysis techniques. The chapter discussed also the importance of thinking about negative evidence and events that are not present in the data.

KEY TERMS

analytic comparison
analytic domain
axial coding
cultural domain
domain analysis
empty boxes
event structure analysis (ESA)
folk domain
historical contingency
illustrative method
method of agreement
method of difference
mixed domain
narrative analysis
negative case method
open coding
outcropping
path dependency
periodization
qualitative comparative analysis (QCA)
selective coding
successive approximation

REVIEW QUESTIONS

1. Identify four differences between quantitative and qualitative data analysis.
2. How does the process of conceptualization differ for qualitative and quantitative research?
3. How does data coding differ in quantitative and qualitative research, and what are the three types of coding used by a qualitative researcher?
4. What is the purpose of analytic memo writing in qualitative data analysis?
5. Describe *successive approximation.*
6. What are the *empty boxes* in the illustrative method, and how are they used?
7. What is the difference between the method of agreement and the method of difference? Can a researcher use both together? Explain why or why not.
8. What are the parts of a domain, and how are they used in domain analysis?
9. What are the major features of a narrative?
10. Why is it important to look for negative evidence, or things that do not appear in the data, for a full analysis?

NOTES

1. See Miles and Huberman (1994) and Ragin (1987). These should not be confused with statistical techniques for "qualitative" data (see Haberman, 1978). These are sophisticated statistical techniques (e.g., logit and log linear) for quantitative variables in which the data are at nominal or ordinal levels. They are better labeled as techniques for categorical data.

2. Sprague and Zimmerman (1989) discuss the importance of an explicit theory.
3. See Hammersley and Atkinson (1983:174–206) for a discussion of questions.
4. See Boyatzis (1998), Lofland and Lofland (1995: 192–193), Miles and Huberman (1994:57–71), Sanjek (1990:388–392), and Wolcott (1994) for additional discussions of coding.
5. See also Horan (1987) and Strauss (1987:25) for multiple indicator measurement models with qualitative data.
6. For more on memoing, see Lester and Hadden (1980), Lofland and Lofland (1995:193–197), Miles and Huberman (1994:72–77), and Strauss (1987:107–129).
7. Also see Barzun and Graff (1970:255–274), Bogdan and Taylor (1975), Lofland and Lofland (1984:131–140), Shafer (1980:171–200), Spradley (1979a, 1979b), and Schatzman and Strauss (1973:104–120) on notes and codes.
8. See Fetterman, 1989:68.
9. See Skocpol (1984) and Skocpol and Somers (1980).
10. For a discussion of analogies and models, see Barry (1975), Glucksmann (1974), Harré (1972), Hesse (1970), and Kaplan (1964).
11. For a discussion of the importance of analogies in social theory, see Lloyd (1986:127–132) and Stinchcombe (1978).
12. For more on successive approximation and a debate over it, see Applebaum (1978a), McQuaire (1978, 1979), P. Thompson (1978), Wardell (1979), and Young (1980).
13. For a discussion of empty boxes, see Bonnell (1980) and Smelser (1976).
14. For a discussion of the illustrative method, see Bonnell (1980) and Skocpol (1984). Bogdan and Taylor (1975:79) describe a similar method.
15. See Coffrey et al. (2002) for an example of domain analysis.
16. For a discussion of methods of difference and agreement, see Ragin (1987:36–42), Skocpol (1984), Skocpol and Somers (1980), and Stinchcombe (1978:25–29).
17. See Mahoney (1999) on a nominal comparison.
18. See Griffin (1993) and Mahoney (1999).
19. On various uses see Abbott (1995) and Franzosi (1998)
20. The six core elements are derived from the following: Abell (2001, 2004), Abbott (1995, 2001), Büthe (2002), Franzosi (1998), Griffin (1992, 1993), Gubrium and Holstein (1998), Haydu (1998), Mahoney (2000), Pedriana (2005), Sewell (1992, 1996), and Stryker (1996).
21. On narrative as a condition of social life, see Abbott (2001) and Somers (1994).
22. Abell (2004:288) remarked, "Although the term narrative and cognate concepts . . . are widely used . . . no settled definition is yet established." Some of the terms used include: analytic narrative (Pedriana, 2005), causal narrative (Sewell, 1996), comparative narrative (Abell, 2001), event structural analysis (Griffin, 1993), historical narrative (Mahoney, 2000b), narrative explanation (Abell, 2004), sequence analysis (Abbott, 1995), and structural analysis of narrative (Franzosi, 1998).
23. On debates about causality in narrative analysis and narrative as explanation, see Abbott (2001:290), Abell (2004), Büthe (2002), Griffin (1993), and Mahoney (2000b). For debate about the narrative, see Haydu (1998), Mahoney (1999), Sewell (1996), and Stryker (1996). Researchers such as Goldthrope (1991, 1997) and Lieberson (1991) question the narrative approach whereas Goldstone (1997) and Rueschemeyer and Stephens (1997) defend its utility.
24. See Haydu (1998:353).
25. Mahoney (2000a) gives a detailed description of the path dependency method and provides many examples of its use. Altman (2000) provides a discussion from the economics literature. Also see Blute (1997) and Pedriana (2005).
26. See Becker and Geer (1982) and Emigh (1997) on the negative case method. Blee and Billings (1986) discuss analyzing "silences" in ethnographic or historical text.
27. See Sanjek (1978) and Werner and Schoepfle (1987a).
28. See Gross (1984) and Miles and Huberman (1994:85, 119–126).
29. See Lofland and Lofland (1995:199–200) and Werner and Schoepfle (1987a:130–146).
30. See Canter et al. (1985) and Werner and Schoepfle (1987a:180–181).
31. See Dohan and Sanchez-Jankowski (1998) and Weitzman and Miles (1995) for a comprehensive review of software programs for qualitative data analysis. Also see Fielding and Lee (1991) and Richards and Richards (1994).
32. See http://www.u.arizona.edu/~cragin/fsQCA/software.shtml.
33. For a more in-depth discussion of event-structure analysis, see Abbott (1992), Griffin (1993), Griffin and Ragin (1994), Heise (1991), and Issac et al. (1994).