

# UNOBTRUSIVE MEASURES

## LEARNING OBJECTIVES

1. Define *unobtrusive measures*, and discuss their use in research, providing examples.
2. Describe the process of content analysis, and give one example.
3. Define both historical research methods and comparative research methods, and give an example of each.
4. Explain the process of event-structure analysis.
5. Identify the strengths and limitations of oral history.
6. Discuss the major methodological challenges that arise in comparative and historical research.
7. List some of the cautions and ethical issues to keep in mind when using unobtrusive methods.



©Stockphoto.com/Brian Jackson

Perhaps the most commonly used methods of social science research today are surveys (including political and opinion polling of all kinds) and face-to-face interviews. These methods can elicit tremendous amounts of valuable information, precisely tailored to the researcher's purposes, at a relatively low cost and with very little "dross," or irrelevant information. They can also use sophisticated sampling and create a close-up, human view on what is happening in social life.

But surveys and interviews have a great disadvantage: They are **reactive methods** in which the people being studied know they are being studied, and so may modify their answers or even the behavior being studied. Adult Americans routinely, for instance, overstate how much they vote, how much they exercise, and how frequently they attend church, whereas they underreport how frequently they tell lies. In an effort to offset the weaknesses of reactive measures, Eugene Webb and his colleagues (Webb et al. 1966; revised edition, 2000) assembled a wide variety of examples of what they called **unobtrusive measures**—that is, research techniques that would gather data without alerting the people under study. As Webb and company said, “So long as one has only a single class of data collection, and that class is the questionnaire or interview, one has inadequate knowledge” (p. 175). They urged that researchers use multiple methods in an effort to validate findings in various ways, and they put together a fascinating compendium of creative (some called them “oddball”) ideas for studying social life: measuring interest in different museum exhibits by the frequency with which floor tiles need to be replaced, discovering the most popular radio stations in town by having car mechanics note the settings on car radio dials, or glancing at the hands of patrons in a neighborhood bar to judge the level of manual work done by the patrons (calluses!).

Actually, there are many kinds of nonreactive research methods available. Webb et al. described four categories of data that might provide unobtrusive measures: physical traces, archives, simple observation, and contrived observation. We begin this chapter with a variety of examples of these more “creative” methods, mainly to suggest how broad these possibilities are. For the remainder of the chapter, we outline three far more commonly used important kinds of research that are also typically nonreactive: content analysis, historical research, and comparative analysis.

## CREATIVE SOURCES

### Physical Traces

As criminal forensic scientists can attest, when human beings do almost anything, they tend to leave behind **physical traces** of themselves—hair, fingerprints, and sweat, but also wear and tear on the things they touch. Simply becoming aware of such traces (we might call it “seeing like a detective”) can provide social scientists with valuable research data. On your way to class you might notice that the carpeting or tile on certain stairways is more worn than on others (as Webb suggested), that the chairs in some classrooms are more likely to be damaged, and that paper towels in one particular restroom always seem to run out first. These all these point to heavier traffic in some areas than in others, so that even without watching human beings moving, you might be able to estimate where they go. Your professor in class might well notice that some students’ paperback books seem remarkably fresh, their backs uncracked and their pages unfilled with notes or underlinings; maybe those students aren’t doing the reading. Wear and tear on a book may only mean that it’s a used book, but lack of such wear almost certainly suggests that no one, now or earlier, has read it.

Patterns of physical wear may change over time, revealing changes in usage. For instance, the famous tennis tournament at Wimbledon, in England, is played each year on grass courts, which, of course, will show usage more readily than would, say, a concrete court. Paul Kedrosky, an entrepreneur who thinks creatively about “data

#### Reactive methods:

When the people being studied know they are being studied, and so may modify their answers or even the behavior being studied.

#### Unobtrusive measures:

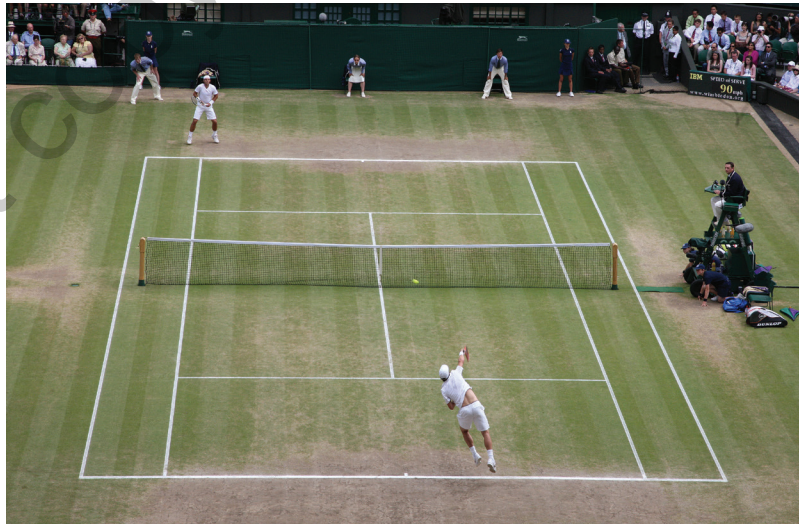
Measurements based on physical traces or other data that are collected without the knowledge or participation of the individuals or groups that generated the data.

#### Physical traces:

Either the erosion or the accumulation of physical substances that can be used as evidence of activity. For instance, footprints in snow indicate that someone has walked there.

exhaust,” or leftover sources of information, has suggested how, in looking at photographs of a match from 25 years ago, you can see that the grass is worn in a pattern that moves up the middle of the court to the net (Exhibit 11.1a). The pattern shows how players rush up to “volley” after their serves. But in the more recent photograph, the grass has been worn thin back at the rear of the court, reflecting a “power baseline” game that has come to predominate in tennis currently (Exhibit 11.1b) (Ryan, 2011). Refuse, trash, even excretions of all sorts can be fruitful sources of information (as physicians have long known). “In December 2011, a pair of data collectors came to

Exhibit 11.1 /// Patterns of Tennis Court Wear Showing Different Styles of Play



Sources: ©Hugo Philpott/UPI/Newscom; Tony Duffy/Alls/Getty Images.

Boston. . . . They made 29 stops . . . walking the neighborhood streets and picking up discarded cigarette packs. They collected 253 packs in all,” and by looking at the state excise tax stamp on each pack, determined that nearly 40% of cigarettes smoked in the Boston area were sold on the black market—they had been illegally imported, to avoid the high cigarette taxes in the state (Hartnett 2014). And in one of the more creative uses of simple wastewater, “since all drug users urinate, and since the urine eventually winds up in the sewers, [Oregon State University chemist Jennifer] Field and her fellow researchers figured that sewer water would contain traces of whatever drugs the citizens were using” (Thompson 2007: n.p.). Samples detected varying usage, by city, of cocaine, methamphetamine, and—most popular of all—caffeine. Cocaine use, interestingly, peaked on weekends, whereas methamphetamine use tended to hold steady across the week (Thompson 2007).

## Archives

By *archives*, we just mean records of all sorts that are already being kept, aside from any social science purpose. These may be quite formal, as in government records of births, deaths, marriages, tax records, building permits, crime statistics, and the like. Law enforcement and health statistics provide, for example, a variety of community-level indicators of substance abuse (Gruenewald et al. 1997). Statistics on arrests for the sale and possession of drugs, drunk driving arrests, and liquor law violations (such as sales to minors) can usually be obtained on an annual basis, and often quarterly. Health-related indicators include single-vehicle fatal crashes, the rate of mortality from alcohol or drug abuse, and the use of treatment centers. All sorts of media create archives that can be mined for data, including newspapers, magazine articles, TV or radio talk shows, legal opinions, historical documents, personal cards and letters, diaries, or e-mail messages. Or one could learn about different U.S. cities, for instance, by looking at the “yellow book” business telephone directories that are still used by many establishments. You would discover there that Sarasota, Florida, has many pages devoted to nursing homes and hospital appliances, but Chattanooga, Tennessee, with roughly the same number of people, has fewer facilities for older people but a huge number of family-friendly churches.

Less obvious, or even totally unintentional, archival sources can also be useful. An abandoned juvenile detention facility was the site, for John M. Klofas and Charles R. Cutshall (1985), of 2,765 instances of graffiti, in settings from the orientation corridors to inmates’ cells to the bathrooms. The authors classified the graffiti by a number of variables including location and content, and concluded that juveniles upon entry seemed more concerned with establishing their individual identity and status; later, their concerns shifted to group affiliations. Romance, politics, and criticisms of the criminal justice system also figured prominently in what inmates wrote about on the walls. Archives of various sorts can also serve as a check on respondents’ self-reports in surveys or in interviews. In Michael Pollan’s best-selling book *In Defense of Food* (2008), he first states that “validation studies of dietary trials like the Women’s Health Initiative or the Nurses’ [Health] Study [conducted on more than 100,000 women over several decades] . . . indicate that people on average eat between a fifth and a third more than they say they do on questionnaires.” He then adds, in a footnote, that “in fact, the magnitude of the error could be much greater, judging by the disparity between the total number

of food calories produced every day for each American (3,900) and the average number of those calories Americans own up to chomping each day: 2,000. Waste can account for some of the disparity, but not nearly all of it” (Pollan 2008: 74).

With the proliferation of smartphones and handheld video recorders, photographic data have become far more available, providing archives of all sorts of routine as well as extraordinary historic events. The Japanese tsunami of 2011 was exceptionally well documented, with real-time recordings of the wave as it came in, as water levels rose, and as the destruction ensued. As recently as the year 2000, almost no such evidence was easily available for study, but now even unpredicted tsunamis, tornadoes, flashfloods, and other catastrophes can and are being fully documented by people on the scene. YouTube and other video websites are wonderful sources for such recordings.

Photography has long provided valuable archival research material. Randall Collins, in research for his sweeping study *Violence: A Micro-Sociological Theory* (2008), assembled many hundreds of photos of people in violent situations from bank robberies to wartime combat to street riots. Collins's book is valuable methodologically for his detailed descriptions of how he selected photos, the sampling and interpretations involved, and the limitations of such data. Even given those issues, though, he was able to conclude (among many other important points) that in groups, violent activity tends to be confined to a few leaders—for instance in a riot in which a handful of protestors throw rocks while many more participants are just supportive or even passive (Exhibit 11.2).

**Archival data:**

Written or visual records, not produced by the researcher.

**Archival data** can be enormously useful, but as always you should be aware in using all sorts of archives that they may not accurately sample or represent reality. Even officially kept records, not to say personal documents, often have built-in biases.

Exhibit 11.2 /// Leaders, Supporters, and Onlookers in a Riot



Source: MUSA AL-SHAER/AFP/Getty Images.

For instance, the level of blood alcohol legally required to establish intoxication can vary among communities, creating an appearance of different rates of abuse even though drinking and driving patterns may in fact be similar. Enforcement practices can vary as well among police jurisdictions, so that conclusions based on these records may be unjustified (Gruenewald et al. 1997: 14).

## Observation

Of course, either moving or still photography is really just a recording of an observation—simply watching people. Fully developed, this is what we've called ethnography or field research (see Chapter 9), but even very brief observations can be revealing. Excellent work has been done, for instance, on the psychology of emotions, so that watching a person's face for even a fraction of a second can often tell you what that person is feeling. Paul Ekman, a psychologist who has helped police forces establish when a suspect is lying or telling the truth (by their facial expressions), is an expert at making detailed observations of the facial features associated with different emotions. Here, in a tragic situation, Ekman describes the look on the face of a woman just told that her missing child has been found murdered (Exhibit 11.3):

One very strong and reliable sign [of intense sadness] is the angling upward of the inner corners of her eyebrows. It is reliable because few people can make this movement voluntarily, so it could rarely be deliberately fabricated. . . . Even when people are attempting not to show how they are feeling, these obliquely positioned eyebrows will often leak their sadness. Look at the space between her eyebrows. In most people a vertical wrinkle between the brows will appear, as it does here. (Ekman 2003: 97)

A person well trained in Ekman's methods could do fascinating studies of different groups of people in public, following their emotional responses to various events, including sporting events, parties, or weddings.

Even simple and obvious sorts of observations, though, can be used to validate other sorts of measures. The tiny Scandinavian island nation of Iceland has very low official crime rates, according to standard police measures. But even casual observation suggests the same conclusion: It is common, for instance, to see babies in strollers lined up outside stores in Reykjavik, the capital, while mothers are inside shopping, a practice unthinkable (or even illegal, as parental negligence) in the United States. When Dan Chambliss lived in Iceland, at night he saw children as young as 6 years old walking alone in downtown Reykjavik, and young women,

### Exhibit 11.3 /// Intense Emotion, Apparent by Close Observation



Source: Ekman, Paul. 2003. *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. New York: Henry Holt.

obviously drunk, staggering home alone from dance clubs. What would be dangerous in an American city was a perfectly safe, if perhaps embarrassing, practice in this benign environment.

At a far more complex level of “observation” stands the massive surveillance programs unveiled by the Edward Snowden leaks, in which the U.S. government was discovered to have been monitoring literally millions of telephone records, as well as hacking the intelligence services of other countries. Our computer-based lives are essentially being observed all the time, of course—by online providers, eager to see what we watch and click on, as well as by employers, who frequently keep track of e-mail and websurfing.

## Contrived Observation

**Contrived observation:**  
Observation of situations in which the researcher has deliberately intervened.

Sometimes researchers with access with online usage data carry out what Webb et al. (1966/2000) called “**contrived observation**,” that is, observation in which the researchers deliberately intervene in the observed activity—for instance, by experimenting. In June 2014, Facebook “revealed that it had manipulated the news feeds of over half a million randomly selected users to change the number of positive and negative posts they saw” (Goel 2014). Investigating the concern that perhaps seeing positive content posted by friends will make viewers feel negative or left out, the researchers (including academics as well as Facebook employees) deliberately modified what was shown on users’ news feeds, to see how users would react. It turns out that people who see more positive content then produce more positive posts themselves. Facebook never asked explicit permission from the people who were studied (there were 689,003), although the company said that the 1.28 billion users give blanket permission when they begin using the service.

A more traditional form of contrived observation would be the groundbreaking linguistic field experiments conducted in the 1960s by William Labov, who hypothesized that people of different social classes pronounced their words differently (Labov 1972). (Specifically, Labov was curious about the way working-class residents of New York City sometimes drop their *rs* in casual conversation: “Hey, come over hee-ah!” instead of “Hey, come over here!” might be an example.) If he used scheduled research interviews, Labov realized, subjects would speak more formally, but he wanted to find out how people pronounce their words in daily life, when they have no idea that they’re being studied.

So Labov sent his research team members into three different New York City department stores (very popular in the 1960s), each representing a different social stratum of the city, as determined by various measures (prices, advertising budgets, etc.). Saks, on the upper East Side, was the expensive store, catering to an upper-class clientele; Macy’s, at Herald Square, was somewhat more middle class; and S. Klein, now closed, was more a budget-level store. Assuming that sales people would to some extent mirror the accents of their customers, researchers would approach employees in each store and ask for directions to items they knew were stocked on the fourth floor of the building. Notice: “fourth floor,” as a response, will provide two different uses of the letter *r*; when the researcher would ask for clarification, the responding sales person would then emphasize the words clearly—giving in total, then, four different examples of the *r* sound. Labov and his team asked 264 subjects for the directions, and found that indeed, the more “upper crust” the store was, the more likely the letter *r* was to be clearly sounded

out—thus confirming his hypothesis of what Labov called “stylistic stratification.” It was an excellent example of a contrived observation.

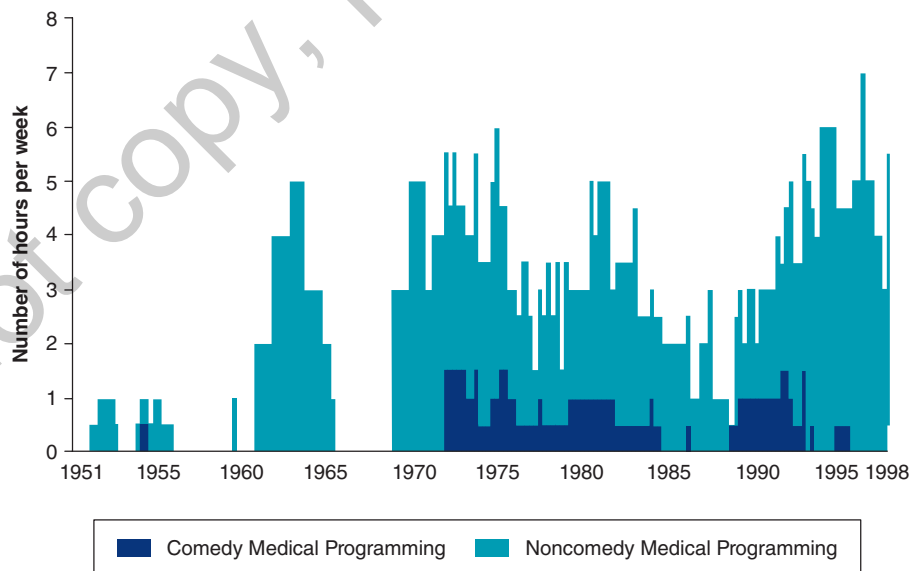
## CONTENT ANALYSIS

One kind of archival observation is content analysis (introduced in Chapter 4), which studies human communication in any of its forms, including books, articles, magazines, songs, films, and speeches (Weber 1990: 9). This method was first applied to the study of newspaper and film content and then developed systematically for the analysis of Nazi propaganda broadcasts in World War II. Since then, content analysis has been used to study historical documents, records of speeches, and other “voices from the past” as well as media of all sorts (Neuendorf 2002: 31–37). The same techniques can now be used to analyze blog sites, wikis, and other text posted on the Internet (Gaiser and Schreiner 2009: 81–90).

Content analysis can be used to study all sorts of topics appropriate for student research projects. How are medical doctors regarded in U.S. culture? Do newspapers use the term *schizophrenia* in a way that reflects what this serious mental illness actually involves? Does the portrayal of men and women in video games reinforce gender stereotypes? Are the body images of male and female college students related to their experiences with romantic love?

Content analysis is particularly well suited to the study of popular culture (Neuendorf 2002: 1). For instance, Kimberly Neuendorf’s (2002: 3) content analysis of medical prime time network television (Exhibit 11.4) shows how medical programming has been dominated by noncomedy shows, but there have been two

Exhibit 11.4 /// Medical Prime Time Network Television Programming, 1951 to 1998



Content analysis typically proceeds according to a regular series of steps.

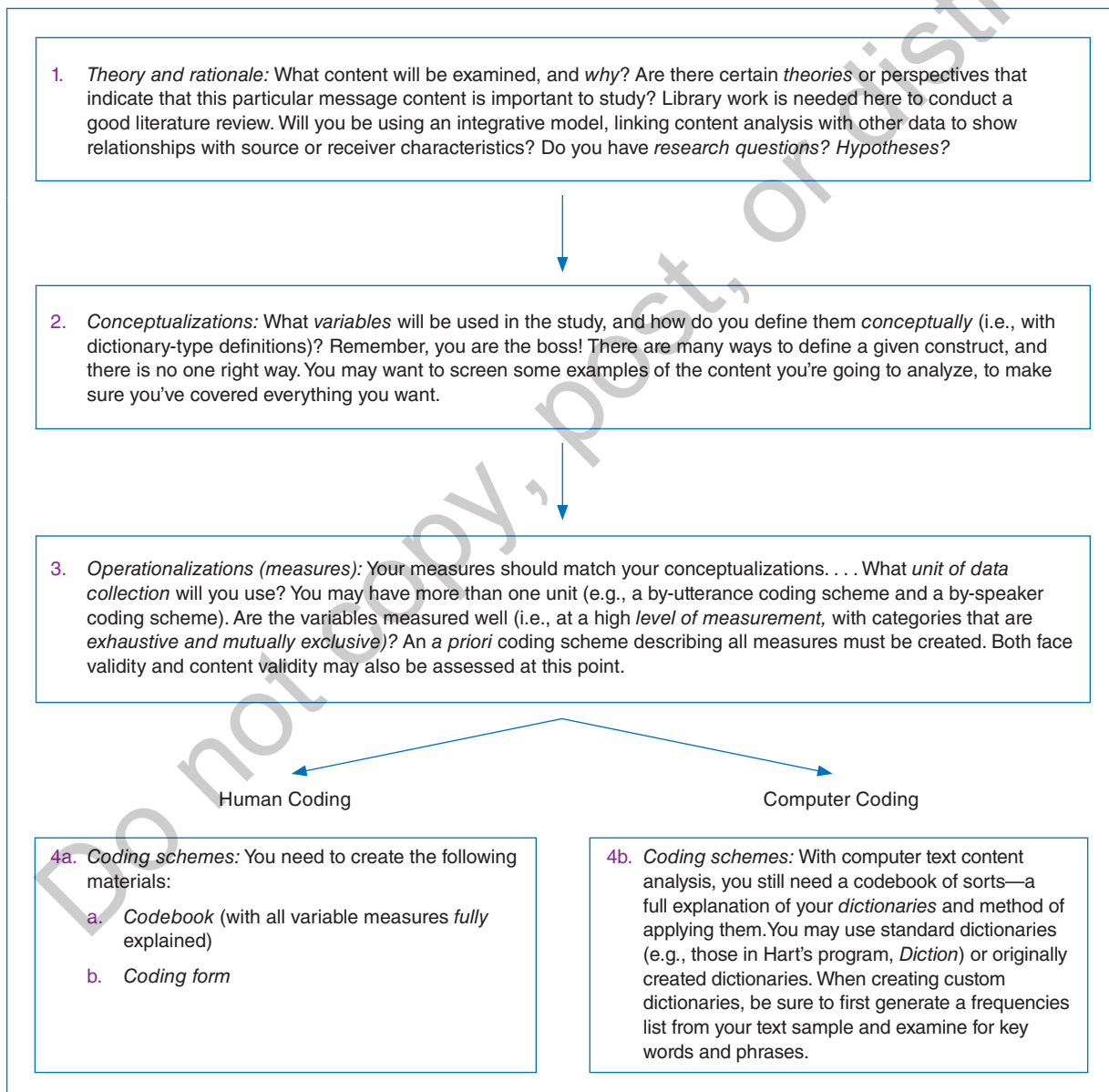
Source: Neuendorf, Kimberly A. 2002. *The content analysis guidebook*. Thousand Oaks, CA: Sage.

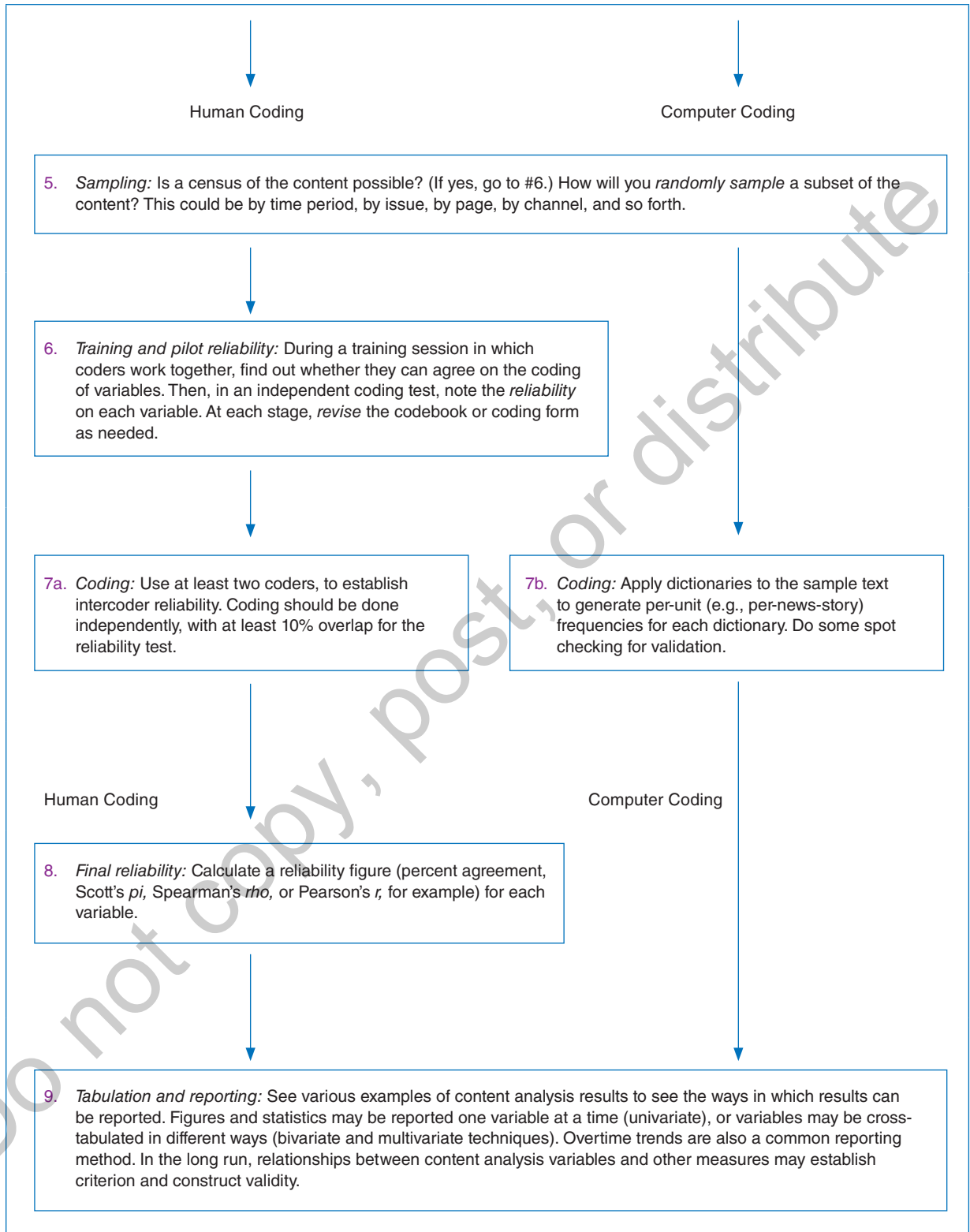


significant periods of comedy medical shows—during the 1970s and early 1980s and then again in the early 1990s. It took a quantitative analysis of medical show content to reveal that the 1960s shows represented a very distinct “physician-as-God” era, which shifted to a more human view of the medical profession in the 1970s and 1980s. This era has been followed, in turn, by a mixed period that has had no dominant theme.

The steps in a content analysis are represented in the flowchart in Exhibit 11.5. Note that the steps are comparable to the procedures in quantitative survey research. Use this flowchart as a checklist when you design or critique a professional content analysis project. We describe a simplified version in the following pages.

Exhibit 11.5 /// Flowchart for the Typical Process of Content Analysis Research





Source: Neuendorf, Kimberly A. 2002. *The content analysis guidebook*. Thousand Oaks, CA: Sage.

## Identify a Population of Documents or Other Textual Sources

Documents to be sampled could include, for instance, all newspapers published in the United States, college student newspapers, nomination speeches at political party conventions, or “state of the nation” speeches by national leaders. Books or films are also common sources for content analysis projects. For her analysis of prime time programming since 1951, Neuendorf (2002: 3–4) used a published catalog of all TV shows. For Russ Schutt’s analysis with Duckworth and others (Duckworth et al. 2003: 1402) of newspapers’ use of the terms *schizophrenia* and *cancer*, they requested a sample of articles from the LexisNexis national newspaper archive. Matthias Gerth and Gabriele Siegert (2012) focused on TV and newspaper stories during a 14-week Swiss political campaign, and Karen Dill and Kathryn Thill (2007: 855–856) turned to video game magazines for their analysis of the depiction of gender roles in video games. For their analysis of gender differences in body image and romantic love, Suman Ambwani and Jaine Strauss (2007: 15) surveyed students at a small midwestern liberal arts college and then analyzed the written responses.

## Determine the Units of Analysis

Units of analysis could be items such as newspaper articles, whole newspapers, speeches, or political conventions, or they could be more microscopic units such as words, interactions, time periods, or other bits of a communication (Neuendorf 2002: 71). The units of analysis for Neuendorf (2002: 2) were “the individual medically oriented TV program”; for Duckworth et al. (2003: 1403), they were newspaper articles; for Gerth and Siegert (2012: 288), they were arguments made in media stories; and for Dill and Thill (2007: 856), they were images appearing in magazine articles. The units of analysis for Ambwani and Strauss (2007: 15) were individual students.

## Design Coding Procedures for the Variables to Be Measured

Coding is probably the trickiest part of doing content analysis. It requires first deciding what variables to measure and what unit of text is to be coded. Do you code words, sentences, themes, or paragraphs? Then, the categories into which the text units are to be coded must be defined. These categories may be broad, such as *supports democracy*, or narrow, such as *supports universal suffrage*. Development of clear instructions and careful training of coders is essential.

As an example, Exhibit 11.6 is a segment of the coding form that Schutt developed for a content analysis of union literature that he collected during a mixed-method study of union political processes (Schutt 1986). His sample was of 362 documents: all union newspapers and a stratified sample of union leaflets given to members during the years of the investigation. The coding scheme included measures of the source and target for the communication, as well as measures of concepts that the theoretical framework indicated were important. (The analysis showed a decline in concern with client issues and an increase in focus on organizational structure, over the period studied.)

*I. Preliminary Codes*

1. Document # \_\_\_\_\_
2. Date \_\_\_\_\_  
                    mo                    yr
3. Length of text \_\_\_\_\_ pp. (round up to next 1/4 page; count legal size as 1.25)
4. Literature Type
  1. General leaflet for members/employees
  2. Newspaper/Newsletter article
  3. Rep Council motions
  4. Other material for Reps, Stewards, Delegates (e.g., budget, agenda)
  5. Activity reports of officers, President's Report
  6. Technical information-filing grievances, processing forms
  7. Buying plans/Travel packages
  8. Survey Forms, Limited Circulation material (correspondence)
  9. Non-Union
  10. Other \_\_\_\_\_ (specify)

4A. If newspaper article

Position

1. Headline story
2. Other front page
3. Editorial
4. Other
5. Other

4B. If Rep Council motion

Sponsor

1. Union leadership
2. Office
3. Leadership faction
4. Opposition faction

5. Literature content—Special issues

1. First strike (1966)
2. Second strike (1967)
3. Collective bargaining (1977)
4. Collective bargaining (1979)
5. Election/campaign literature
6. Affiliation with AFSCME/SEIU/other national union
7. Other

*II. Source and Target*

6. Primary source (code in terms of those who prepared this literature for distribution).
  1. Union-newspaper (Common Sense; IUPAE News)
  2. Union-newsletter (Info and IUPAE Bulletin)
  3. Union-unsigned
  4. Union officers
  5. Union committee
  6. Union faction (the Caucus; Rank-and-File; Contract Action, other election slate; PLP News; Black Facts)
  7. Union members in a specific work location/office
  8. Union members—other
  9. Dept. of Public Aid/Personnel
  10. DVR/DORS

(Continued)

11. Credit Union
12. Am. Buyers' Assoc.
13. Other non-union
7. Secondary source (use for lit. at least in part reprinted from another source, for distribution to members)
  1. Newspaper—general circulation
  2. Literature of other unions, organizations
  3. Correspondence of union leaders
  4. Correspondence from DPA/DVR-DORS/Personnel
  5. Correspondence from national union
  6. Press release
  7. Credit Union, Am. Buyers'
  8. Other \_\_\_\_\_ (specify)
  9. None
8. Primary target (the audience for which the literature is distributed)
  1. Employees—general (if mass-produced and unless otherwise stated)
  2. Employees—DVR/DORS
  3. Union members (if refers only to members or if about union elections)
  4. Union stewards, reps, delegates committee
  5. Non-unionized employees (recruitment lit, etc.)
  6. Other \_\_\_\_\_ (specify)
  7. Unclear

III. Issues

- A. Goal
- B. Employee conditions/benefits (Circle up to 5)
  1. Criteria for hiring
  2. Promotion
  3. Work out of Classification, Upgrading
  4. Step increases
  5. Cost-of-living, pay raise, overtime pay, "money"
  6. Layoffs (nondisciplinary); position cuts
  7. Workloads, Redeterminations, "30 for 40," GA Review
  8. Office physical conditions, safety
  9. Performance evaluations
  10. Length of workday
  11. Sick benefits/leave—holidays, insurance, illness, vacation, voting time
  12. Educational leave
  13. Grievances—change in procedures
  14. Discrimination (race, sex, age, religion, national origin)
  15. Discipline—political (union-related)
  16. Discipline—performance, other
  17. Procedures with clients, at work
  18. Quality of work, "worthwhile jobs"—other than relations with clients

\*Coding instructions available from author.

Source: Reprinted by permission from Schutt, Russell K. 1986. *Organization in a changing environment*. Albany: State University of New York Press. Reprinted by permission of The State University of New York Press. All rights reserved.

Developing reliable and valid coding procedures is not easy. The meaning of words and phrases is often ambiguous. Coding procedures cannot simply categorize and count words; text segments in which the words are embedded must also be inspected before codes are finalized. Because different coders may perceive different meanings in the same text segments, explicit coding rules are required (Weber 1990: 23–29).

Dill and Thill (2007) used two coders and a careful training procedure for their analysis of the magazine images about video games:

One male and one female rater, both undergraduate psychology majors, practiced on images from magazines similar to those used in the current investigation. Raters discussed these practice ratings with each other and with the first author until they showed evidence of properly applying the coding scheme for all variables. Progress was also checked part way through the coding process, as suggested by [Gloria] Cowan (2002). Cowan (2002) reports that this practice of reevaluating ratings criteria is of particular value when coding large amounts of violent and sexual material because, as with viewers, coders suffer from desensitization effects. (Dill and Thill 2007: 856)

After coding procedures are developed, their reliability should ideally be assessed by comparing different coders' codes for the same variables. Computer programs for content analysis can enhance reliability by facilitating the consistent application of text-coding rules (Weber 1990: 24–28). Validity can be assessed with a construct validation approach by determining the extent to which theoretically predicted relationships occur (see Chapter 4).

## Select a Sample of Units From the Population

The simplest strategy might be a simple random sample of the documents you are using. However, a stratified sample might be needed to ensure adequate representation of community newspapers in large and in small cities, or of weekday and Sunday papers, or of political speeches during election years and in off years (Weber 1990: 40–43). Nonrandom sampling methods have also been used in content analyses (Neuendorf 2002: 87–88).

The selected samples in our exemplar projects were diverse. For Schutt's content analysis with Duckworth (Duckworth et al. 2003), they had a student, Chris Gillespie, draw a stratified random sample of 1,802 articles published in the five U.S. newspapers with the highest daily circulation in 1996 to 1997 in each of the four regions identified in the LexisNexis database, as well as the two high-circulation national papers in the database, the *New York Times* and *USA Today* (pp. 1402–1403).

Because individual articles cannot be sampled directly in the LexisNexis database, a random sample of days was drawn first. All articles using the terms *schizophrenia* or *cancer* (or several variants of these terms) were then selected from the chosen newspapers on these days. Gerth and Siebert (2012: 285) selected 24 different newspapers and 5 TV news programs that targeted the population for the campaign, and then coded 3,570 arguments made in them about the campaign during its 14 weeks. Dill and Thill (2007: 855–856) used all images in the current

issues (as of January 2006) of the six most popular video game magazines sold on Amazon.com. Ambwani and Strauss (2007: 15) used an availability sampling strategy, with 220 students from introductory psychology and a variety of other sources. Neuendorf (2002: 2) included the entire population of medically oriented TV programs between 1951 and 1998.

## Code Documents and Apply Appropriate Statistical Analyses

In a content analysis, your data are created by counting occurrences of particular words, themes, or phrases, and then testing relations between the resulting variables. These analyses could use some of the statistics that were introduced in Chapter 8, including frequency distributions, measures of central tendency and variation, cross-tabulations, and correlation analysis (Weber 1990: 58–63). In many cases, computer-aided qualitative analysis programs can help in developing coding procedures and carrying out the content coding.

Final results may be presented in a number of ways. Exhibit 11.5 is the simple chart that Neuendorf (2002: 3) used to present the frequency of TV medical programming. Schutt's content analysis with Duckworth and others (Duckworth et al. 2003) was simply a comparison of percentages showing that 28% of the articles mentioning schizophrenia used it as a metaphor, compared with only 1% of the articles mentioning cancer; it also gave examples of the text that had been coded into different categories. For example, *the nation's schizophrenic perspective on drugs* was the type of phrase coded as a metaphorical use of the term *schizophrenia* (p. 1403). Dill and Thill (2007: 858) presented percentages and other statistics that showed that, among other differences, female characters were much more likely to be portrayed in sexualized ways in video game images than were male characters. Ambwani and Strauss (2007: 16) used other statistics showing that in survey responses, body esteem and romantic love experiences are related, particularly for women. They also examined the original written comments and found further evidence for this relationship. For example, one woman wrote, “[My current boyfriend] taught me to love my body. Now I see myself through his eyes, and I feel beautiful” (p. 17).

Content analysis, then, has the power to reveal broad patterns in how people understand even the most intimate sorts of experiences.

## HISTORICAL METHODS

The central insight behind both historical and comparative research, as we will see, is that we can improve our understanding of social process when we make comparisons with other times and places. Max Weber's comparative study of world religions (Bendix 1962) and Émile Durkheim's (1984) historical analysis of the division of labor are two examples of the central role of historical and comparative research during the period sociology emerged as a discipline. Although the popularity of this style of research ebbed with the growth of survey methods and statistical analysis in the 1930s, exemplary works such as Reinhard Bendix's (1956) *Work and Authority in Industry* and Barrington Moore Jr.'s (1966) *Social Origins of Democracy and Dictatorship* helped fuel a resurgence of historical

and comparative methods in the 1970s and 1980s that has continued into the 21st century (Lange 2013: 22–33). In recent years, the globalization of U.S. economic ties and the internationalization of scholarship have increased the use of unobtrusive methods for comparative research across many different countries (Kotkin 2002).

Historical methods are used increasingly by social scientists in sociology, anthropology, political science, and economics, as well as by many historians (Monkkonen 1994). The late 20th and early 21st centuries have seen so much change in so many countries that many scholars have felt a need to investigate the background of these changes and to refine their methods of investigation (Hallinan 1997; Robertson 1993).

Much historical research is qualitative. Like other qualitative methods, qualitative historical research is inductive: It develops an explanation for what happened from the details discovered about the past. In addition, qualitative historical research is **case-oriented**; it focuses on the nation or other unit as a whole, rather than only on different parts of the whole in isolation from each other (Ragin 2000: 68). The research question is “What was Britain like at the time?” rather than “What did Queen Elizabeth do?” Related to this case orientation, qualitative historical research is **holistic**—concerned with the context in which events occurred and the interrelations between different events and processes: “how different conditions or parts fit together” (Ragin 1987: 25–26). Charles Ragin (2000) uses the example of case-oriented research on the changing relationship between income and single parenthood in the United States after World War II:

In the end, the study is also about the United States in the second half of the twentieth century, not just the many individuals and families included in the analysis. More than likely, the explanation of the changing relation between income and single parenthood would focus on interrelated aspects of the United States over this period. For example, to explain the weakening link between low income and single parenthood the researcher might cite the changing status of women, the decline in the social significance of conventional family forms, the increase in divorce, the decrease in men’s job security, and other changes occurring in the United States over this period. (pp. 67–68)

Qualitative historical research is also likely to be *historically specific*—limited to the specific time(s) and place(s) studied. Qualitative historical research uses **narrative explanations**—in which the research tells a story involving specific actors and other events occurring at the same time (Abbott 1994: 102) or one that accounts for the position of actors and events in time and in a unique historical context (Griffin 1992). Larry Griffin’s (1993) research on lynching, in the next section, provides a good example.

The focus on the past presents particular methodological challenges:

- Documents and other evidence may have been lost or damaged.
- Available evidence may represent a sample biased toward more newsworthy figures.

**Case-oriented research:**

Research that focuses attention on the nation or other unit as a whole.

**Holistic research:**

Research concerned with the context in which events occurred and the interrelations between different events and processes.

**Narrative explanation:**

An explanation that involves developing a narrative of events and processes that indicate a chain of causes and effects.



- Written records will be biased toward those who were more prone to writing.
- Feelings of individuals involved in past events may be hard, if not impossible, to reconstruct.

Before you judge historical social science research as credible, you should look for convincing evidence that each of these challenges has been addressed.

## Event-Structure Analysis

### Event-structure analysis:

A systematic method of developing a causal diagram showing the structure of action underlying some chronology of events; the result is an idiographic causal explanation.

One technique useful in historical research is **event-structure analysis**. Event-structure analysis is a qualitative approach that relies on a systematic coding of key events or national characteristics to identify the underlying structure of action in a chronology of events. The codes are then used to construct event sequences, make comparisons between cases, and develop an idiographic causal explanation for a key event.

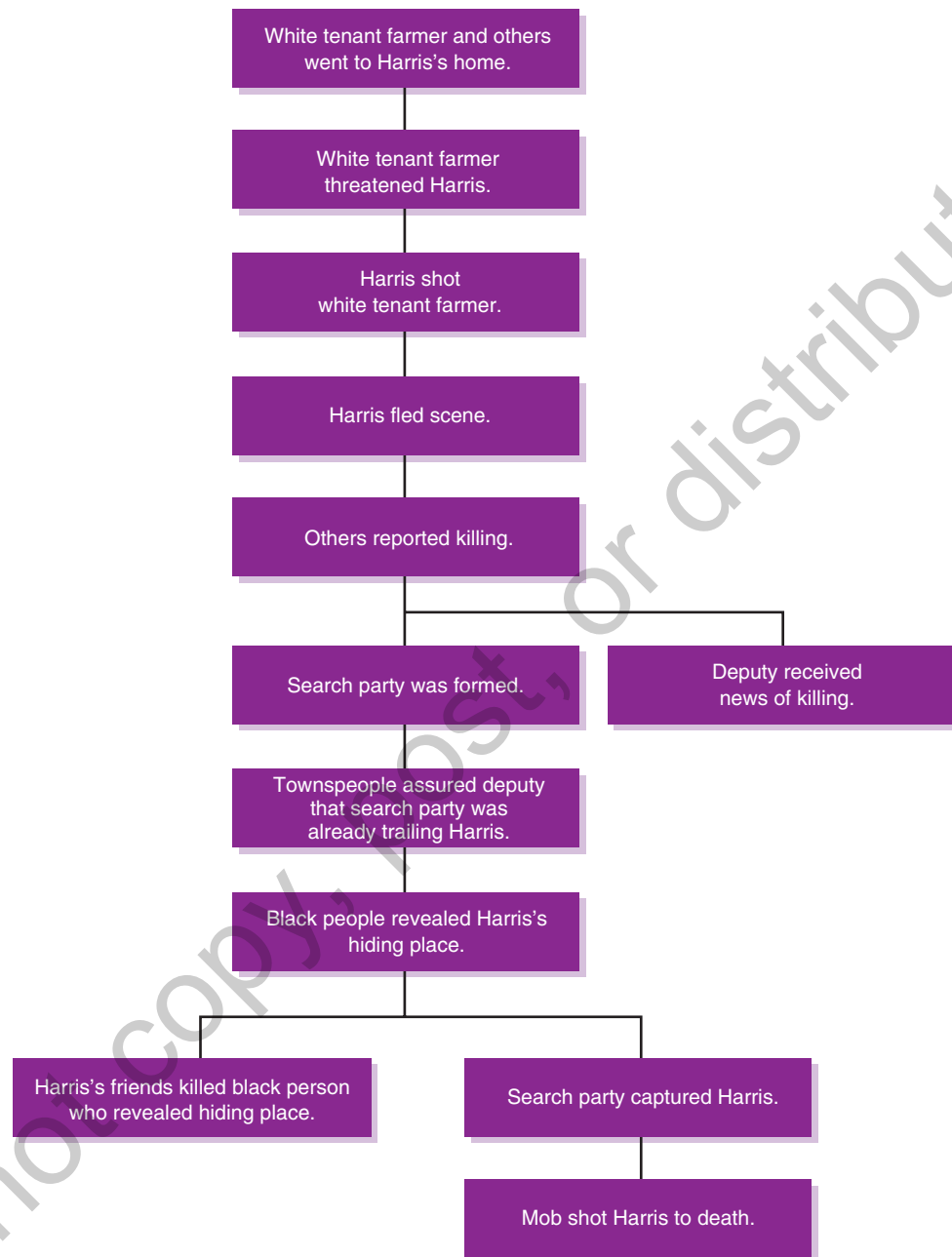
An event-structure analysis consists of the following steps:

1. Classifying historical information into discrete events
2. Ordering events into a temporal sequence
3. Identifying prior steps that are prerequisites for subsequent events
4. Representing connections between events in a diagram
5. Eliminating from the diagram connections that are not necessary to explain the focal event

Griffin (1993) used event-structure analysis to explain a unique historical event, a lynching in the 1930s in Mississippi. According to published accounts and legal records, the lynching occurred after David Harris, an African American who sold moonshine from his home, was accused of killing a white tenant farmer. After the killing was reported, the local deputy was called and a citizen search party was formed. The deputy did not intervene as the search party trailed Harris and then captured and killed him. Meanwhile, Harris's friends killed another African American who had revealed Harris's hiding place. This series of events is outlined in Exhibit 11.7.

Which among the numerous events occurring between the time that the tenant farmer confronted Harris and the time that the mob killed Harris had a causal influence on that outcome? To determine these idiographic causal links, Griffin identified plausible counterfactual possibilities—events that might have occurred but did not—and considered whether the outcome might have been changed if a counterfactual had occurred instead of a particular event.

If, contrary to what actually happened, the deputy had attempted to stop the mob, might the lynching have been averted? . . . Given what happened in comparable cases and the Bolivar County deputy's clear knowledge of the existence of the mob and of its early activities, his forceful intervention to prevent the lynching thus appears an objective possibility. (Griffin 1993: 1112)



Source: Adapted from Griffin, Larry J. 1993. "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98 (March 1993):1094–1133. Reprinted with permission from the University of Chicago Press.

So, Griffin concluded that nonintervention by the deputy had a causal influence on the lynching.

**Oral history:**

Data collected through intensive interviews with participants in past events.

## Oral History

History that is not written down is mostly lost to posterity (and social researchers). However, **oral histories** can be useful for understanding historical events that occurred within the lifetimes of living individuals. As the next example shows, sometimes oral histories even result in a written record that can be analyzed by researchers at a later point in time.

Thanks to a Depression-era writers' project, Deanna Pagnini and Philip Morgan (1996) found that they could use oral histories to study attitudes toward births out of wedlock among African American and white women in the South during the 1930s.

Almost 70% of African American babies are born to unmarried mothers, compared with 22% of white babies (Pagnini and Morgan 1996: 1696). This difference often is attributed to contemporary welfare policies or problems in the inner city, but Pagnini and Morgan thought it might be the result of more enduring racial differences in marriage and childbearing. To investigate these historical differences, they read 1,170 life histories recorded by almost 200 writers who worked for a New Deal program during the Depression of the 1930s, the Federal Writers' Project Life History Program for the Southeast. The interviewers had used a topic outline that included family issues, education, income, occupation, religion, medical needs, and diet.

In 1936, the divergence in rates of nonmarital births was substantial in North Carolina: 2.6% of white births were to unmarried women, compared with 28.3% of nonwhite births. The oral histories gave some qualitative insight into community norms that were associated with these patterns. A white seamstress who became pregnant at age 16 recalled, "I'm afraid he didn't want much to marry me, but my mother's threats brought him around" (Pagnini and Morgan 1996: 1705). There were some reports of suicides by unwed young white women who were pregnant. In comparison, African American women who became pregnant before they were married reported regrets but rarely shame or disgrace. There were no instances of young black women committing suicide or getting abortions in these circumstances.

We found that bearing a child outside a marital relationship was clearly not the stigmatizing event for African Americans that it was for whites. . . . When we examine contemporary family patterns, it is important to remember that neither current marriage nor current childbearing patterns are "new" for either race. Our explanations for why African Americans and whites organize their families in different manners must take into account past behaviors and values. (Pagnini and Morgan 1996: 1714–1715)

Whether oral histories are collected by the researcher or obtained from an earlier project, the stories they tell can be no more reliable than the memories that are recalled. Unfortunately, memories of past attitudes are "notoriously subject to modifications over time" (Banks 1972: 67), as are memories about past events, relationships, and actions. Use of corroborating data from documents or other sources should be used when possible to increase the credibility of descriptions based on oral histories.

One common measurement problem in historical research projects is the lack of data from some historical periods (Rueschemeyer, Stephens, and Stephens 1992: 4; Walters, James, and McCammon 1997). For example, the widely used U.S. Uniform Crime Reporting System did not begin until 1930 (Rosen 1995). Sometimes, alternative sources of documents or estimates for missing quantitative data can fill in gaps (Zaret 1996), but even when measures can be created for key concepts, multiple measures of the same concepts are likely to be out of the question; as a result, tests of reliability and validity may not be feasible (Bollen, Entwisle, and Alderson 1993; Paxton 2002).

The available measures are not always adequate. What is included in the historical archives may be an unrepresentative selection of materials that remain from the past. At various times, some documents could have been discarded, lost, or transferred elsewhere for a variety of reasons. *Original* documents may be transcriptions of spoken words or handwritten pages and could have been modified slightly in the process; they could also be outright distortions (Erikson 1966: 172, 209–210; Zaret 1996). When relevant data are obtained from previous publications, it is easy to overlook problems of data quality, but this simply makes it all the more important to evaluate the primary sources.

## COMPARATIVE METHODS

The limitations of single-case historical research have encouraged many social scientists to turn to comparisons between nations. These studies allow for a broader vision about social relations than is possible with cross-sectional research limited to one country or other unit.

### Cross-Sectional Comparative Research

Comparisons between countries during one time period can help social scientists identify the limitations of explanations based on single-nation research. Such comparisons can suggest the relative importance of universal factors in explaining social phenomena compared with unique factors rooted in specific times and places (de Vaus 2008: 251). These comparative studies may focus on a period in either the past or the present. Peter Houtzager and Arnab Acharya (2011) also point out that it can be more appropriate to compare cities or regions when the nations in which they are embedded vary internally in their social characteristics. For example, they compare the impact of engagement in associations on citizenship activity in São Paulo, Brazil, and Mexico City because the conditions exist for such an impact in these cities, rather than in the surrounding countries.

Researchers engaged in quantitative historical and comparative research may obtain data from national statistics or other sources of published data; if it is contemporary, such research may rely on cross-national surveys. Like other types of quantitative research, quantitative historical and comparative research can be termed **variable-oriented research**, with a focus on variables representing particular aspects of the units studied (Demos 1998).

Causal reasoning in quantitative comparative research is nomothetic, and the approach is usually deductive, testing explicit hypotheses about relations between these variables (Kiser and Hechter 1991). For example, Clem Brooks and Jeff

#### Variable-oriented research:

Research that focuses attention on variables representing particular aspects of the cases studied and then examines the relations between these variables across sets of cases.



## Careers and Research

### Kurt Taylor Gaubatz, PhD, Independent Scholar



Kurt Taylor Gaubatz

Kurt Taylor Gaubatz is the quintessential comparative researcher whose book *Elections and War* (1999) exemplifies the approach. But he started college at the University of California, Berkeley, majoring in music. He became fascinated by the challenge of understanding and modeling human

behavior only after he took a required economics class. He realized, “All of the biggest problems we face as a society, indeed as human beings, come down to research questions in the social sciences!”

Driven by his desire to influence public policy, Gaubatz went on to earn one master’s degree from the Fletcher School of Law and Diplomacy and another from Princeton Theological Seminary. He then earned his PhD in political science from Stanford

University and, several prestigious fellowships later, joined the faculty in the graduate program in international studies of the Department of Political Science & Geography at Old Dominion University. He is now an independent scholar. He describes his career in research as “a life of posing and answering questions, of trying to think about things in new and more interesting ways.”

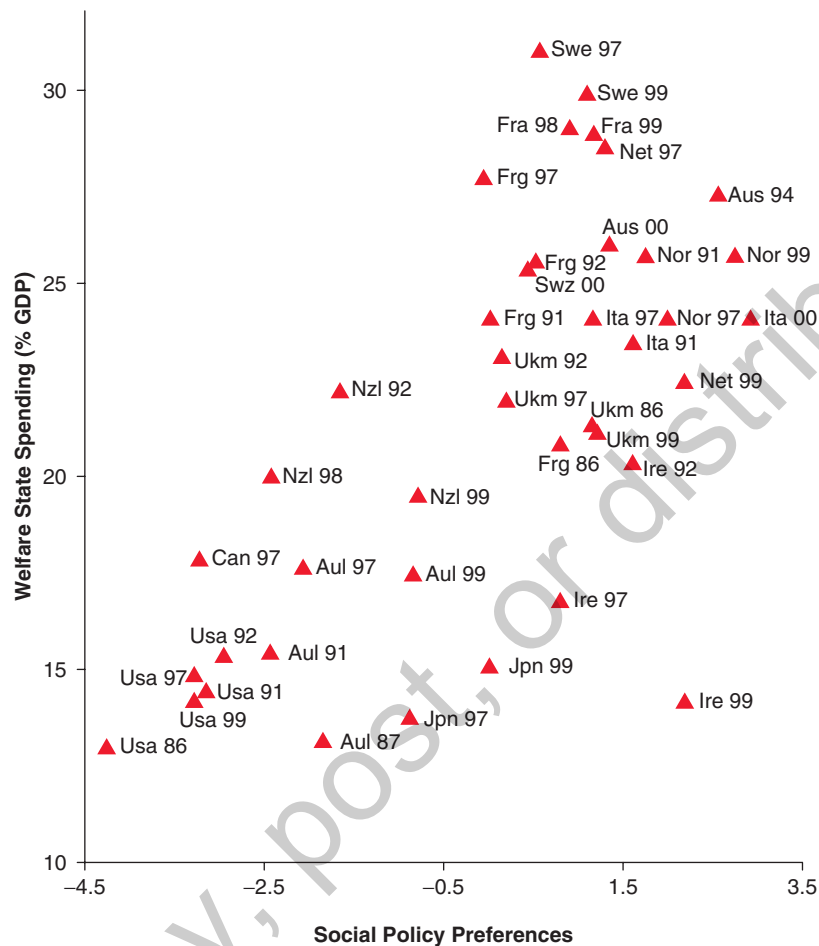
Gaubatz’s advice for students interested in research careers focuses on the ongoing revolution in information technology:

We are in the middle of a revolution in data creation and computing power. Just 25 years ago, people could make a career from knowing information. A huge amount of information is now increasingly available to everyone who carries a phone. The critical skill is knowing how to build new ideas from the organization and analysis of that information, and being able to communicate those ideas effectively. Students need to focus on filling their toolboxes with those analytic and communication skills.

Manza (2006: 476–479) deduce from three theories about welfare states—national values, power resources, and path dependency theory—the hypothesis that voters’ social policy preferences will influence welfare state expenditures. Using country-level survey data collected by the International Social Survey Program (ISSP) in 15 democracies in five different years and expenditure data from the Organisation for Economic Co-operation and Development (OECD), Brooks and Manza were able to identify a consistent relationship between popular preferences for social welfare spending and the actual national expenditures (Exhibit 11.8).

Popular preferences are important factors in political debates over immigration policy. Christopher Bail (2008) asked whether majority groups in different European countries differ in the way that they construct “symbolic boundaries”

Exhibit 11.8 /// Interrelationship of Policy Preferences and Welfare State Output



Note: Scattergram shows data for policy preferences and welfare state spending in 15 OECD democracies. Data are from the ISSP/OECD (International Social Survey Program/Organisation for Economic Co-operation and Development).

Source: Brooks, Clem, and Jeff Manza. 2006. Social policy responsiveness in developed democracies. *American Sociological Review* 71(3): 474–494. Reprinted with permission from the American Sociological Association.

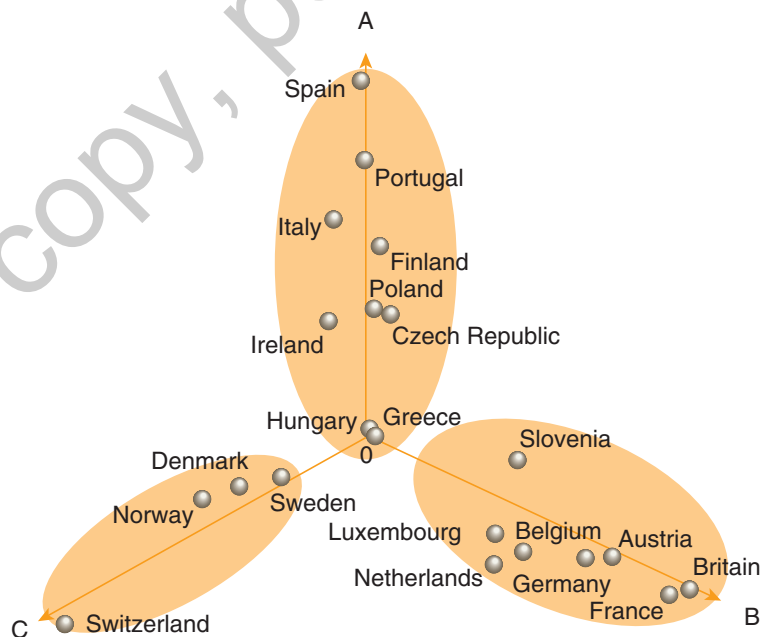
that define “us” versus an immigrant “them.” For his cross-sectional comparative investigation, he drew on 333,258 respondents in the 21-country European Social Survey (ESS). The key question about immigrants in the ESS was “Please tell me how important you think each of these things should be in deciding whether someone born, brought up and living outside [country] should be able to come and live here.” The “things” whose importance they were asked to rate were six individual characteristics: (1) being white, (2) being well educated, (3) having a Christian background, (4) speaking the official national language, (5) being committed to the country’s way of life, and (6) having work skills needed in the country. Bail then calculated the average importance rating in each country for each of these characteristics and used a statistical procedure to cluster the countries by the extent to which their ratings and other characteristics were similar.

Bail's (2008: 54–56) analysis identified the countries as falling into three clusters (Exhibit 11.9). Cluster A countries are on the periphery of Europe and have only recently experienced considerable immigration; their populations tend to draw boundaries by race and religion. Cluster B countries are in the core of Western Europe (except Slovenia), have a sizable and long-standing immigrant population, and their populations tend to base their orientations toward immigrants on linguistic and cultural differences. Countries in Cluster C are in Scandinavia, have a varied but relatively large immigrant population, and attach much less importance to any of the six symbolic boundaries than do those in the other countries. Bail (2008: 56) encourages longitudinal research to determine the extent to which these different symbolic boundaries are the product or the source of social inequality in these countries.

Cross-sectional comparative research has also helped explain variation in voter turnout. This research focuses on a critical issue in political science: Although free and competitive elections are a defining feature of democratic politics, elections cannot orient governments to popular sentiment if citizens do not vote (LeDuc, Niemi, and Norris 1996). As a result, the low levels of voter participation in U.S. elections have long been a source of practical concern and research interest.

International data give our first clue for explaining voter turnout: The historic rate of voter participation in the United States (48.3%, on average) is much lower than it is in many other countries that have free, competitive elections; for example, Italy has a voter turnout of 92.5%, on average, since 1945 (Exhibit 11.10).

Exhibit 11.9 /// Symbolic Boundaries Against Immigrants in 21 European Countries



Source: Bail, Christopher A. February 2008. "The Configuration of Symbolic Boundaries against Immigrants in Europe." *American Sociological Review* 73(1):37–59. Reprinted with permission from the American Sociological Association.

Exhibit 11.10 /// Average Percentage of Voters Who Participated in Presidential or Parliamentary Elections, 1945–1998\*

Country	Vote %	Country	Vote %
Italy	92.5	St. Kitts and Nevis	58.1
Cambodia	90.5	Morocco	57.6
Seychelles	96.1	Cameroon	56.3
Iceland	89.5	Paraguay	56.0
Indonesia	88.3	Bangladesh	56.0
New Zealand	86.2	Estonia	56.0
Uzbekistan	86.2	Gambia	55.8
Albania	85.3	Honduras	55.3
Austria	85.1	Russia	55.0
Belgium	84.9	Panama	53.4
Czech	84.8	Poland	52.3
Netherlands	84.8	Uganda	50.6
Australia	84.4	Antigua and Barbuda	50.2
Denmark	83.6	Burma/Myanmar	50.0
Sweden	83.5	Switzerland	49.3
Mauritius	82.8	USA	48.3
Portugal	82.4	Mexico	48.1
Mongolia	82.3	Peru	48.0
Tuvalu	81.9	Brazil	47.9
Western Samoa	81.9	Nigeria	47.6
Andorra	81.3	Thailand	47.4
Germany	80.9	Sierra Leone	46.8
Slovenia	80.6	Botswana	46.5
Aruba	80.4	Chile	45.9
Namibia	80.4	Senegal	45.6
Greece	80.3	Ecuador	44.7
Guyana	80.3	El Salvador	44.3
Israel	80.0	Haiti	42.9
Kuwait	79.6	Ghana	42.4
Norway	79.5	Pakistan	41.8
San Marino	79.1	Zambia	40.5
Finland	79.0	Burkina Faso	38.3
Suriname	77.7	Nauru	37.3
Malta	77.6	Yemen	36.8
Bulgaria	77.5	Colombia	36.2
Romania	77.2	Niger	35.6

\*Based on entire voting-age population in countries that held at least two elections during these years. Only countries with highest and lowest averages are shown.

Source: Reproduced by permission of International IDEA from Turnout in the world—country by country performance (1945–1998). From Voter Turnout: A Global Survey (<https://www.idea.int/data-tools/data/voter-turnout>) © International Institute for Democracy and Electoral Assistance.



Does this variation result from differences between voters in knowledge and wealth? Do media and political party get-out-the-vote efforts matter? Mark Franklin's (1996: 219–222) analysis of international voting data indicates that neither explanation accounts for much of the international variation in voter turnout. Instead, the structure of competition and the importance of issues are influential. Voter turnout is maximized where structural features maximize competition: compulsory voting (including, in Exhibit 11.10, Austria, Belgium, Australia, and Greece), mail and Sunday voting (including the Netherlands and Germany), and multiday voting. Voter turnout also tends to be higher where the issues being voted on are important and where results are decided by proportional representation (as in Italy and Israel, in Exhibit 11.10) rather than on a winner-take-all basis (as in U.S. presidential elections)—so individual votes are more important.

Franklin concludes that these characteristics explain the low level of voter turnout in the United States rather than the characteristics of individual voters. The United States lacks the structural features that make voting easier, the proportional representation that increases the impact of individuals' votes, and, often, the sharp differences between candidates that are found in countries with higher turnout. Because these structural factors generally do not vary within nations, we would never realize their importance if our analysis was limited to data from individuals in one nation.

Despite the unique value of comparative analyses like Franklin's (1996), such cross-national research also confronts unique challenges (de Vaus 2008: 255). The meaning of concepts and the operational definitions of variables may differ between nations or regions (Erikson 1966: xi), so the comparative researcher must consider how best to establish measurement equivalence (Markoff 2005: 402). For example, the concept of being a *good son or daughter* refers to a much broader



## In the News

### Research in the News

#### Britain Cracking Down on Gender Stereotypes in Ads

Britain's Advertising Standards Authority reported that gender stereotypes in ads could “restrict the choices, aspirations and opportunities” of girls and teenagers and others who view the ads. It is developing new standards for advertising that it will then enforce. Ads that fail “to demonstrate the mother's value to the family” or otherwise endorse gender equality could be banned.

*News Source:* Magra, Iliana. 2017. Britain cracking down on gender stereotypes in ads. *New York Times*, July 18.

Feminist groups, marketing groups, and journalists are debating the proposed standards.

#### For Further Thought

1. What are the expectations about gender equality in your country? Can you imagine rules like those under consideration in Britain being endorsed there?
2. What indicators of gender inequality would you propose for historical and comparative research?

range of behaviors in China than in most Western countries (Ho 1996). Rates of physical disability cannot be compared between nations because standard definitions are lacking (Martin and Kinsella 1995: 364–365). Individuals in different cultures may respond differently to the same questions (Martin and Kinsella 1995: 385). Alternatively, different measures may have been used for the same concepts in different nations, and the equivalence of these measures may be unknown (van de Vijver and Leung 1997: 9). The value of statistics for particular geographic units such as counties in the United States may vary over time simply because of changes in the boundaries of these units (Walters et al. 1997). Such possibilities should be considered, and any available opportunity should be taken to test for their effects.

## Longitudinal Comparative Research

Dietrich Rueschemeyer et al. (1992) used a comparative historical method, combining the approaches, to explain why some nations in Latin America (excluding Central America) developed democratic politics, whereas others became authoritarian or bureaucratic–authoritarian states. First, Rueschemeyer et al. developed a theoretical framework that gave key attention to the power of social classes, state (government) power, and the interaction between social classes and the government. The researchers then classified the political regimes in each nation over time (Exhibit 11.11). Next, they noted how each nation varied over time relative to the variables they had identified as potentially important for successful democratization.

Exhibit 11.11 /// Classification of Regimes Over Time

	Constitutional Oligarchic	Authoritarian; Traditional, Populist, Military, or Corporatist	Restricted Democrat	Fully Democratic	Bureaucratic–Authoritarian
Argentina	Before 1912	1930–46 1951–55 1955–58 1962–63	1958–62 1963–66	1912–30 1946–51 1973–76 1983–90	1966–73 1976–83
Bolivia	Before 1930	1930–52 1964–82	1982–90	1952–64	
Brazil	Before 1930	1930–45	1945–64 1985–90		1964–85
Chile	Before 1920	1924–32	1920–24 1932–70 1990	1970–73	1973–89
Colombia	Before 1936	1949–58	1936–49 1958–90		

(Continued)

	Constitutional Oligarchic	Authoritarian; Traditional, Populist, Military, or Corporatist	Restricted Democrat	Fully Democratic	Bureaucratic-Authoritarian
Ecuador	1916–25	Before 1916 1925–48 1961–78	1948–61 1978–90		
Mexico		Up to 1990			
Paraguay		Up to 1990			
Peru		Before 1930 1930–39 1948–56 1962–63 1968–80	1939–48 1956–62 1963–68	1980–90	
Uruguay		Before 1903 1933–42	1903–19	1919–33 1942–73 1984–90	1973–84
Venezuela		Before 1935 1935–45	1958–68	1945–48 1968–90	

Source: Rueschemeyer, Dietrich, Evelyne Huber Stephens, and John D. Stephens. 1992. *Capitalist Development and Democracy*. Reprinted with permission from the University of Chicago Press

Their analysis identified several conditions for initial democratization: consolidation of state power (ending overt challenges to state authority), expansion of the export economy (reducing conflicts over resources), industrialization (increasing the size and interaction of middle and working classes), and some agent of political articulation of the subordinate classes (which could be the state, political parties, or mass movements). Historical variation in these conditions was then examined in detail.

The great classical sociologists also used comparative methods, although their approach was less systematic. For example, Max Weber's (Bendix 1962: 268) comparative sociology of religions contrasted Protestantism in the West, Confucianism and Taoism in China, Hinduism and Buddhism in India, and Ancient Judaism. As Bendix (1962) explained,

His [Weber's] aim was to delineate religious orientations that contrasted sharply with those of the West, because only then could he specify the features that were peculiar to Occidental [Western] religiosity and hence called for an explanation . . . to bring out the distinctive features of each historical phenomenon. (p. 268)



## Research That Matters

Is an increase in democratic freedoms associated with greater representation of women in powerful political positions? Prior research indicates that this is not the case; in fact, case studies have shown a drop in women's representation in government in some countries that have adopted democratic governance. But, there are complicating factors, including whether gender quotas were implemented. Kathleen Fallon, Liam Swiss, and Jocelyn Viterna conducted a historical comparative research project to investigate why more democracy can be associated with fewer women in government. They collected data from 118 developing countries over a 34-year period. The dependent variable was

the percentage of seats held by women in the national legislature. The researchers distinguished countries transitioning from civil strife, authoritarian regimes, and communist regimes, and they accounted for the use of quotas for women, the extent of democratic practices, and differences in national culture.

The results indicate that women's legislative representation drops after democratizing changes begin, but then increases with additional elections; the *process* of democratic change is critical to understanding its outcome for women.

*Source:* Adapted from Fallon, Kathleen M., Liam Swiss, and Jocelyn Viterna. 2012. Resolving the democracy paradox: Democratization and women's legislative representation in developing nations, 1975 to 2009. *American Sociological Review* 77(3): 380–408.

So, for example, Weber concluded that the rise of Protestantism, with its individualistic approach to faith and salvation, was an important factor in the development of capitalism.

### Cautions for Comparative Analysis

Of course, ambitious methods that compare different countries face many complications. The features of the cases selected for comparison have a large impact on the researcher's ability to identify influences. Cases should be chosen for their difference of key factors hypothesized to influence the outcome of interest and their similarity on other, possibly confounding, factors (Skocpol 1984: 383). For example, to understand how industrialization influences democracy, you would need to select cases for comparison that differ in industrialization, so that you could then see if they differ in democratization (King et al. 1994: 148–152). Nonetheless, relying on just a small number of cases for comparisons introduces uncertainty into the conclusions (de Vaus 2008: 256). The focus on comparisons between nations may itself be a mistake for some analyses. National boundaries often do not correspond to key cultural differences, so comparing subregions within countries or larger cultural units that span multiple countries may make more sense for some analyses (de Vaus 2008: 258). Comparing countries that have fractured along cultural or religious divides simply by average characteristics would obscure many important social phenomena.

With cautions such as these in mind, historical and comparative methods allow for rich descriptions of social and political processes in different nations or regions as well as for causal inferences that reflect a systematic, defensible weighing of the

evidence. Data of increasingly good quality are available on a rapidly expanding number of nations, creating many opportunities for comparative research. We cannot expect one study comparing the histories of a few nations to control adequately for every plausible alternative causal influence, but repeated investigations can refine our understanding and lead to increasingly accurate causal conclusions (King et al. 1994: 33).

## ETHICAL ISSUES IN UNOBTUSIVE METHODS

Ethical concerns arise when using unobtrusive measures that involve observing people, analyzing pictures of them, or collecting evidence of their activities. Although the potential harm to research participants may be delayed, it can still occur unless care is used to avoid disclosing identities—including covering faces in photos that are published. Pictures of individuals engaging in activities in public settings do not create as many concerns, but even such pictures may reveal behaviors that the participants would not want to be disclosed.

Analysis of historical documents, documents from other countries, or content in media does not create the potential for harm to human subjects that can be a concern when collecting primary data. It is still important to be honest and responsible in working out arrangements for data access when data must be obtained from designated officials or data archivists, but many data are available easily in libraries or on the web. Researchers in the United States who conclude that they are being denied access to public records of the federal government may be able to obtain the data by filing a Freedom of Information Act (FOIA) request. The FOIA stipulates that all persons have a right to access all federal agency records unless the records are specifically exempted (Riedel 2000: 130–131). Researchers who review historical or government documents must also try to avoid embarrassing or otherwise harming named individuals or their descendants by disclosing sensitive information.

Ethical concerns are multiplied when surveys are conducted or other data are collected in other countries. If the outside researcher lacks much knowledge of local norms, values, and routine activities, the potential for inadvertently harming subjects is substantial. For this reason, cross-cultural researchers should spend time learning about each of the countries in which they plan to collect primary data and strike up collaborations with researchers in those countries (Hantrais and Mangen 1996). Local advisory groups may also be formed in each country so that a broader range of opinion is solicited when key decisions must be made. Such collaboration can also be invaluable when designing instruments, collecting data, and interpreting results.

Cross-cultural researchers who use data from other societies have a particular obligation to try to understand the culture and norms of those societies before they begin secondary data analyses. It is a mistake to assume that questions asked in other languages or cultural contexts will have the same meaning as when asked in the researcher's own language and culture, so a careful, culturally sensitive process of review by knowledgeable experts must precede measurement decisions in these projects. Researchers must become familiar with gender norms in the societies they seek to study because they may result in cross-country variation in responses to survey questions, willingness to participate in surveys, definitions of

terms used in government statistics (e.g., the term *labor participation*), and distortions in statistical data (Ayhan 2001).

## CONCLUSION

We've covered a huge range of research methods in this chapter, but all of them intervene relatively little in the lives of people they study, unlike participant observation, surveys, or interviews; in that sense, all are "unobtrusive." Some of them represent among the finest examples of classical and contemporary social science and are capable of addressing sweeping topics of international importance. Ideally, in your own research you can use and combine different methods, as a way of compensating for the weaknesses of each, to improve the validity of your findings. The creative methods we suggested at the beginning of this chapter should help with that—and perhaps be enjoyable to develop and use, as well.

## /// KEY TERMS

Archival data	288	Holistic research	299	Reactive methods	285
Case-oriented research	299	Narrative explanation	299	Unobtrusive measures	285
Contrived observation	290	Oral history	302	Variable-oriented research	303
Event-structure analysis	300	Physical traces	285		

## /// HIGHLIGHTS

- Many social science projects rely on methods such as surveys, interviews, or participant observations that are inherently reactive, in that they may change the behavior they are intended to study. Unobtrusive measures try to avoid this weakness in research.
- Unobtrusive measures can be based on physical traces, archives, or observations.
- Content analysis is a tool for systematic quantitative analysis of documents and other textual data. It requires careful testing and control of coding procedures to achieve reliable measures.
- The central insight behind historical and comparative methods is that we can improve our understanding of social processes when we make comparisons with other times and places.
- Event-structure analysis is a systematic qualitative approach to developing an idiographic causal explanation for a key event.
- Oral history provides a means of reconstructing past events. Data from other sources should be used whenever possible to evaluate the accuracy of memories.
- Comparative methods may be cross-sectional, such as when variation between country characteristics is compared, or longitudinal, in which developmental patterns are compared between countries.
- Analysis of historical documents, documents from other countries, or content in media usually creates less potential for harm to human subjects than analysis of primary data, but it is still important to be honest and responsible in working out arrangements for data access when data must be obtained from designated officials or data archivists. Unobtrusive measures obtained from physical traces or observations require attention to the ethical issues also relevant in qualitative research.

## /// STUDENT STUDY SITE

### SAGE edge™

The Student Study Site, available at [edge.sagepub.com/chamblissmssw6e](http://edge.sagepub.com/chamblissmssw6e), includes useful study materials including practice quizzes, eFlashcards, videos, audio resources, journal articles, and more.

## /// EXERCISES

### Discussing Research

1. The creative measures suggested by Webb et al. (1966/2000) as well as those described in the beginning of this chapter span a wide range of approaches. Can you think of other unobtrusive measures you might use?
2. Review the differences between case-oriented, historically specific, inductive explanations and those that are more variable oriented, theoretically general, and deductive. List several arguments for and against each approach. Which is more appealing to you and why?
3. What historical events have had a major influence on social patterns in the nation? The possible answers are too numerous to list, ranging from any of the wars to major internal political conflicts, economic booms and busts, scientific discoveries, and legal changes. Pick one such event in your own nation for this exercise. Find one historical book on this event, and list the sources of evidence used. What additional evidence would you suggest for a social science investigation of the event?
4. Susan Olzak, Suzanne Shanahan, and Elizabeth McEneaney (1996) developed a nomothetic causal explanation of variation in racial rioting in the United States over time, whereas Griffin's (1993) explanation of a lynching can be termed *idiographic*. Discuss the similarities and differences between these types of causal explanation. Use these two studies to illustrate the strengths and weaknesses of each.

### Finding Research

1. Paul Ekman, the psychologist cited who studies evidence of emotions in people's faces, has written extensively on this topic, and his work is widely used by police departments and even intelligence agencies. Find and read his findings on how to spot if someone is lying.
2. The journals *Social Science History* and *Journal of Social History* report many studies of historical processes. Select one article from a recent journal issue about a historical process used to explain some event or other outcome. Summarize the author's explanation. Identify any features of the explanation that are temporal, holistic, and conjunctural. Prepare a chronology of the important historical events in that process. Do you agree with the author's causal conclusions? What additional evidence would strengthen the author's argument?

### Critiquing Research

1. What would be the weaknesses of using graffiti, such as in the Klofas and Cutshall (1985) study, to determine what prison inmates are thinking about? Might there be other ways of gathering such information that could be more accurate? What would be their weaknesses?

### Doing Research

1. If you've read some of Ekman's work as suggested in "Finding Research," use his methods to watch people at some event, for example, a sporting competition or a reception. Keep track in detail of what they look like, and see if you can spot unexpected or socially awkward reactions. What might they mean?
2. Consider the media that you pay attention to in your social world. How could you design a content analysis of the messages conveyed by these media? What research questions could you help to answer by adding a comparison with another region or country to this content analysis?

Exhibit 11.12 /// Voting Procedures in 10 Countries

	Voting Age	Number of Days Polling Booth Open	Voting Day on Work Day or Rest Day	Postal Voting	Proxy Voting	Constituency Transfer	Advance Voting	Voter Turnout (in %)	Year (P = presidential, L = legislative election)
Switzerland	20	2	Rest day	Automatic for armed forces, otherwise by application 4 days before voting	Varies by canton	No	No	46	1991L
Taiwan	20	1	Rest day					72	1992L
Thailand	20	1	Rest day	No				62	1995L
Turkey	20	1	Rest day	No	No	Special polling stations at border posts for citizens residing abroad	No	80	1991L
Ukraine	18	1	Rest day					71.6	1994P
United Kingdom	18	1	Work day	On application	On application	No	No	77.8	1992L
United States	18	1	Work day	By application; rules vary across states	In some states for blind and disabled	No		51.5	1992P
Uruguay	18	1	Rest day	No	No	No		89.4	1994P
Venezuela	18	1	Rest day	No	Assisted voting for blind and disabled	No	No	60	1993P
Zambia		1	Work day		No			50	1991P

Source: LeDuc, Lawrence, Richard G. Niemi, and Pippa Norris. 1996. "Introduction: The Present and Future of Democratic Elections." Pp. 1–48 in Lawrence LeDuc, Richard G. Niemi, and Pippa Norris (eds.), *Comparing Democracies: Elections and Voting in Global Perspective*. Thousand Oaks, CA: SAGE Publications, p. 19, Table 1.3.



3. Select a major historical event or process, such as the Great Depression, World War II, the civil rights movement, or the war in Iraq. Why do you think this event happened? Now, select a historical or comparative method that you think could be used to test your explanation. Why did you choose this method? What type of evidence would support your proposed explanation? What problems might you face in using this method to test your explanation?
4. Using your library's government documents collection or the U.S. Census site on the web, select one report by the U.S. Census Bureau about the population of the United States or some segment of it. Outline the report

### Ethics Questions

1. Facebook and other popular social media sites routinely collect, use, and sell massive amounts of personal data. Do you think that's ethically right? When could it be right, and when wrong? What about experimentation on users, such as giving some users certain information and others not? Do you think a blanket waiver, such as what all users must sign when joining many sites, provides a sufficient level of consent?
2. Oral historians can uncover disturbing facts about the past. What if a researcher were conducting an oral history project such as the Depression Writer's Project and learned from an interviewee about his previously undisclosed involvement in a predatory sex crime many years ago? Should the researcher report what

### Video Interview Questions

Listen to the researcher interview for Chapter 11 at [edge.sagepub.com/chamblissmssw6e](http://edge.sagepub.com/chamblissmssw6e), found in the Video and Multimedia Section.

1. What caused Cinzia Solari's research question to change? What was the comparative element in her research?

and list all the tables included in it. Summarize the report in two paragraphs. Suggest a historical or comparative study for which this report would be useful.

5. Consider the comparative historical research by Rueschemeyer et al. (1992) on democratic politics in Latin America. What does comparison between nations add to the researcher's ability to develop causal explanations?
6. Exhibit 11.12 identifies voting procedures and the level of turnout in one election for 10 countries. Do voting procedures appear to influence turnout in these countries?

he learned to a government attorney who might decide to bring criminal charges? What about informing the victim or her surviving relatives? Would it matter if the statute of limitations had expired, so that the offender could not be prosecuted any longer? Would it matter if the researcher were subpoenaed to testify before a grand jury?

3. In this chapter's ethics section, we recommended that researchers who conduct research in other cultures form an advisory group of local residents to provide insight into local customs and beliefs. What are some other possible benefits of such a group for cross-cultural researchers? What disadvantages might arise from use of such a group?

2. How did Solari build rapport between her and the migrant workers she was trying to research? Why is this step important when doing qualitative research?