Content Analysis

Chapter Overview

Content analysis involves the examination of various documents and **texts**, which may be printed, visual, aural, or virtual. It can be **quantitative**, focused on **coding** data into predetermined categories in a systematic and easily **replicable** manner, or **qualitative**, seeking to uncover deeper meanings in the materials. Although both approaches are discussed, this chapter deals primarily with the qualitative variety. It considers:

- the kinds of research question that content analysis can answer;
- the features of documents or texts that are commonly analyzed;
- how to code (a key part of content analysis);
- **semiotics** and **hermeneutics**;
- active and passive audiences;
- **conversation analysis** and its roots in **ethnomethodology**;
- the assumptions and analytic strategies of **discourse analysis**, including **critical discourse analysis**; and
- the advantages and disadvantages of content analysis.

Rebecca and her boyfriend, Daniel, were sitting at home watching TV on a Saturday night. Neither was in a particularly good mood; both thought there must be a more exciting way to spend the weekend. "Have you ever noticed that there are way more food commercials—burgers, pizza, chicken—at night than in the daytime?" Rebecca asked. "Are there?" Daniel replied. "I know hockey games always seem to come with a lot of beer ads."

"You ever notice how the women look in those ads?" Rebecca fumed. "They're always in bathing suits or shorts, even if they're standing on a ski hill in winter. And by some miracle they all have perfect bodies." "Yeah," said Daniel. "It's a little different with the guys. A lot of them look like dorks, doing stupid things. But some guys look cool, and they're the ones who get the girls. They're better-looking and taller than the dorks." "We could probably do a study comparing men and women in

beer commercials," Rebecca said, her mood beginning to lighten. "We could even throw in a comparison with detergent ads. Ever see a man selling detergent?" "You could do that," Daniel said. "You'd just have to find a way to categorize the portrayals—flattering or unflattering, active or passive, the age of the people, how attractive they are, overtly sexual or non-sexual behaviour—that kind of thing." "That wouldn't be too hard to do," Rebecca answered. "Sounds like a lot more fun than watching this stupid stuff."

Introduction

There can be similarities between quantitative and qualitative approaches to social research, and even where there are differences the two orientations can sometimes be combined or at least used in a complementary fashion. Content analysis in some ways is a blend of the two. It examines forms of communication to see what they reveal about a society, a culture, or even the relationships between individuals. Meanings and interpretations are important aspects of content analysis, but so is formal categorization of the communications, which often includes some quantification. For example, suppose you are interested in how newspapers cover crime. You come up with the following questions:

- Do certain newspapers report more crimes than others?
- How much crime is reported? Where in the paper do crime stories appear—on the front page, or inside?
- Do columnists as well as reporters write about crime?
- Are some crimes given more attention than others?
- Do more crime stories appear during the week or on weekends?
- What sorts of crime predominate in newspaper articles (crimes against the person, property crime, crimes in which society is the victim)?

Most content analysis of the media is likely to entail several research questions, generally revolving around the same five W's that are the basis of any news report: *who* (does the reporting); *what* (gets reported); *where* (does the issue get reported); *why* (does the issue get reported); and *when* (does it get reported). But researchers are also interested in what media coverage omits.

For example, interviews with the family of the accused are rare; such inattention is itself notable, revealing what is and is not important to writers and publishers.

Another common theme in content analysis is change in the coverage of an issue over time. For example, Buchanan (2014) found that in the period 1894–2005 there was a significant decrease in the amount of local news featured in the *Toronto Star* and the *Ottawa Citizen*, and that both papers had increased their national news coverage over this period.

Content analyses of this kind yield a predominantly quantitative description of the characteristics of a communication. Those who perform content analysis in this way claim to be objective and systematic, clearly specifying the rules for the categorization of the material in advance. Researchers try to create transparency in the coding procedures so that personal biases intrude as little as possible. The coding rules in question may, of course, reflect the researcher's interests and concerns and therefore be a product of subjective bias to some degree, but once the rules have been formulated, it should be possible to apply them without bias.

Content analysis can be performed on unstructured information, such as transcripts of semi- and unstructured interviews, and even qualitative case studies of organizations (for example, Hodson, 1996). Until recently it was used mainly to examine printed texts and documents, but other sorts of materials may also be analyzed. If you were interested in gender roles or representations, for example, you might look at:

- differences in the degree to which men and women are sexualized on the covers of popular magazines (Hatton & Trautner, 2013);
- websites showing pictures of men and women, such as dating sites;

Social Research Methods by Alan Bryman

- animated cartoons (Thompson & Zerbinos, 1995); and
- lyrics of popular songs (e.g., to look for changes in the representation of women) (Marcic, 2002).

Here we will call all such materials "documents," and define them as any data source that:

- can be "read" (including visual materials such as photographs); and
- was *not* produced specifically for the purpose of social research.

Documents are important because they are unobtrusive and non-**reactive**. The fact that the subjects don't know they are being studied removes a common threat to the **validity** of the data.

In discussing the different kinds of documents used in social science, Scott (1990) distinguished between personal and official documents and, within the latter group, between private and state documents. These distinctions are used in much of the discussion that follows. Scott also enumerated four criteria for assessing the quality of documents (1990, p. 6):

- *Authenticity.* Is the evidence genuine and of unquestionable origin?
- *Credibility.* Is the evidence free from error and distortion?
- *Representativeness.* Is the evidence typical of what it is supposed to represent (for example, social life at a particular time and place)? If not, is the extent of its uniqueness known?
- *Meaning.* Is the evidence clear and comprehensible?

Note that this use of "authenticity" is different from how the term was used in Chapter 9 to assess qualitative research.

Personal documents

Diaries, letters, and autobiographies

Diaries and letters are often used by historians, but they have been given less attention by other social researchers (see Box 12.1). Whereas letters are written to communicate with other people, diarists presumably write for themselves. When written for wider consumption, diaries are difficult to distinguish from another kind of personal document: the autobiography. Used with a **life history** or **biographical method**, diaries, letters, and autobiographies (whether solicited or unsolicited) can be either primary sources of data or adjuncts to other sources such as life story interviews. However, this chapter is primarily concerned with unsolicited documents.

The distinction between biographies and autobiographies can sometimes break down. Walt Disney provides a case in point. The first biography of Disney, written by his daughter, Diane Disney Miller (1956), almost certainly included information from Mr Disney himself. Subsequent biographies were very similar, because of the tight control exercised by the Walt Disney Corporation over the primary materials in the Disney archive (letters, notes of meetings, and so on) out of which the biographies were fashioned. As a consequence, even though Walt Disney never wrote an autobiography in the conventional meaning of the term, he (and, later, his company) clearly controlled what was written about him.

In evaluating personal documents, the *authenticity* criterion is obviously important. Is the purported author of the letter or diary the real author? In the case of autobiographies, the increasing use of "ghost" writers has made this a standard question. But the same kind of problem can arise with other documents. For example, in the case of Augustus Lamb (Box 12.1), Dickinson (1993, pp. 126–127) noted that there are "only three letters existing from Augustus himself (which one cannot be certain were written in Augustus's own hand, since the use of amanuenses was not uncommon)." This raises the question of whether Augustus was in fact the author of the letters, especially in the light of his apparent learning difficulties.

Turning to the issue of *credibility*, Scott (1990) observed that there are at least two major concerns with respect to personal documents: their factual accuracy and whether they express the true thoughts and feelings of the writer. The case of Augustus Lamb suggests that a definitive, factually accurate account is at the very least problematic. Scott recommended a healthy skepticism regarding the sincerity

BOX 12.1 Using historical personal documents: The case of Augustus Lamb

West7megan/Dreamstime.com/GetStock

What sorts of dilemmas do you face if you use historical personal documents as primary sources for research?

Dickinson (1993) provided an interesting account of the use of historical personal documents in the case of Augustus Lamb (1807–36), the only child of the novelist Lady Caroline Lamb and her husband, the second Viscount Melbourne, who was prime minister of the UK from 1834 to 1841. It is possible that Augustus suffered from epilepsy, though he seems to have had other medical problems as well throughout his short life. Dickinson was drawn to him because of her interest in nineteenth-century reactions to non-institutionalized people with mental disabilities. In fact, Dickinson doubted that Augustus was mentally disabled, and suggested the somewhat milder diagnosis of "learning difficulties." At the same time, she found that the people around him had difficulty coming to terms with his conditions, in large part because of the challenge of finding words to describe him that were consistent with his high social status.

The chief sources of data were "letters from family and friends; letters to, about and (rarely) from Augustus" (1993, p. 122). Other sources included the record of the post-mortem examination conducted on his remains and extracts from the diary of his resident tutor and physician for the years 1817–21. Despite the many sources she consulted, Dickinson still could not conclude with certainty that she understood definitively what Augustus was like.

with which the writer reports his or her feelings. Famous people who are aware that their letters or diaries will be of interest to others may be very careful about how much of themselves they reveal in their writings. Similarly, adolescent diary-keepers are often aware that their parents could "accidentally" read what they write. In short, what is *not* said can be of great importance. A particularly poignant illustration is Sugiman's (2004) suggestion that the reason many Japanese-Canadian women interned during the Second World War did not write down their experiences was to shield their children from that painful episode in their family histories.

Representativeness is clearly a major concern in assessing these materials. Since literacy rates were very low in the past, letters, diaries, and autobiographies

were largely the preserve of a small class of wealthy, literate people. Moreover, because boys were more likely to receive an education than girls, the voices of women tended to be under-represented in documents of that kind. Some researchers also argue that in the past it was less socially acceptable for women than men to write diaries and autobiographies. Therefore such historical documents are likely to be biased in terms of their applicability to the society as a whole.

A further problem is the selective survival of documents like letters. What proportion are damaged, lost, or thrown away? How representative are suicide notes of the thoughts and feelings of all suicide victims? It seems that only a relatively small percentage of suicide victims leave notes, and some of the notes that are written may be destroyed by the victims' relatives, especially if they contain accusations against family members.

Finally, understanding the meaning of the documents we do have is often problematic. Some pages in a letter or diary may be missing or damaged, or the writer may have used abbreviations or codes that are difficult to decipher. Also, as Scott (1990) observed, letter writers tend to leave much unsaid when they take it for granted that the people they are writing to have the same values, assumptions, and information as they do.

Researchers often have to search long and hard for new or surprising materials in diaries. However, this was not the case with the diaries of William Lyon Mackenzie King. When the King diaries were first released to the public in the 1970s, they caused a sensation (see Box 12.2).

Visual objects

There is a growing interest in the use of visual materials, especially photographs, in social research (see Box 12.3). One of the main reasons why photographs are of interest to social researchers is their potential to reveal important information about families.

As Scott (1990) observed, many family photos are taken as records of ceremonial occasions such as weddings and graduations, or of recurring events such as reunions and holidays. Scott distinguished three

---

BOX 12.2 The Mackenzie King diaries

William Lyon Mackenzie King was prime minister of Canada for 21 of the 27 years between 1921 and 1948. He was an avid diarist for most of his life, recording his thoughts and feelings about both personal issues and national and international events. The contents of the diaries suggest that he did not expect them to be made public. They reveal his political views in considerable detail, but also a side of the man that shocked many people when the transcripts were released. King believed strongly in psychic phenomena and spiritualism—he attended séances, consulted mediums, and believed that he was in communication with various deceased persons, including Franklin Delano Roosevelt and his own mother. He also believed that he could commune with his deceased dog, Pat. (To place these revelations in their proper historical context, the practice of spiritualism was quite common when King was a young man in the late nineteenth and early twentieth centuries.) Sensationalism aside, the King diaries are a valuable primary source of information for a significant portion of Canadian and world history.

Nevertheless, the publication of the King diaries does raise ethical questions. Is it a violation of privacy rights to publish or quote from someone's diary, especially if the person is deceased and cannot raise objections? King's diaries are now available online at Library and Archives Canada: https://www.bac-lac.gc.ca/eng/discover/politics-government/prime-ministers/william-lyon-mackenzie-king/Pages/diaries-william-lyon-mackenzie-king.aspx (accessed 29 October 2018).

BOX 12.3 Photographs in social research

JD Moodie / Library and Archives Canada / C-001823

This photo of Niviaqsarjuk (left), Suzie (centre), and Jennie (right) was taken by Geraldine Moody at Fullerton (in what is now Nunavut) in 1904, although the names of the women depicted were unknown to Library and Archives Canada until 2002. How might photos like this one be used as data? What issues around representation or ethics might arise when Indigenous cultures are represented by people who are not community members? How might community consultation help mitigate these issues?

Photographs can play a variety of roles in social research. They can be used in qualitative research, as well as in surveys and experiments. Three prominent roles photographs can play are the following:

- *As illustrations.* Photographs may be used to illustrate points and thus to enliven what might otherwise be rather dry discussions of findings. This was certainly the case with some classic reports by anthropologists.
- *As data.* Photographs may be data in their own right. When produced for research purposes, they essentially become part of the researcher's field notes. When based on extant photographs, they may become the main source of data about the field in which the researcher is interested, as in the work of Blaikie (2001) and Sutton (1992, see Box 12.4).
- *As prompts.* Photographs may be used as prompts to entice people to talk about what they see in them. Both researcher-created photographs and previously existing ones may be used in this way. Sometimes research participants may volunteer their own photographs. For example, Riches and Dawson (1998) found in their interviews with bereaved parents that many parents showed them pictures of their deceased children, likely the same pictures they had shown to neighbours and friends when they were first grieving their loss.

types of home photograph: *idealization* (a formal portrait of a wedding party or the family in its finery); *natural portrayal* (an informal snapshot capturing action as it happens, though there may be a contrived component); and *demystification* (depicting the subject in an atypical—often embarrassing—situation). Scott suggested that researchers need to be aware of these different types in order to avoid being deceived by the superficial appearance of images. One must probe beneath the surface:

> There is a great deal that photographs do not tell us about their world. Hirsch [1981, p. 42] argued, for example, that "The prim poses and solemn faces which we associate with Victorian photography conceal the reality of child labour, women factory workers, whose long hours often brought about the neglect of their infants, nannies sedating their charges with rum, and mistresses diverting middle class fathers." (Scott, 1990, p. 195)

As Scott argued, this means not only that the photograph must not be taken at face value when used as a research source, but that the viewer must have considerable knowledge of the social context in order to get its full meaning. Sutton (1992) makes a similar point in Box 12.4. In fact, we may wonder whether photographs in such situations can be of any use to a researcher. At the same time, a researcher's interpretations should not be accepted uncritically. For example, the "prim poses and solemn faces" of Victorian portraits might well "conceal" a bleak and miserable reality; but is that all? Could they not also conceal moments of human warmth or compassion? Or fun?

A particular problem for the analyst of photographs, according to Scott, is judging *representativeness*. Photos that survive the passage of time—in archives, for example—are very unlikely to be representative for the simple reason that somebody at some time decided they should be preserved. The discussion in Box 12.4 suggests that decisions about which photos to keep or post online may reflect the needs and biases of the people who make those decisions. Another problem relates to the

issue of what is *not* photographed, as suggested by Sutton's idea that unhappy events at Disney theme parks may not be photographed at all. Awareness of what is not photographed can reveal the "mentality" of the person(s) behind the camera. It's clear that the question of representativeness is much more fundamental than the issue of what survives, because it points to how the selective survival of photographs may be part of a reality that family members (or others) deliberately seek to fashion. As in Sutton's example, that manufactured reality itself may then become a focus of interest for the social researcher.

Government documents

The state is a source of much information of potential significance for social researchers. It produces a great deal of quantitative statistical information. For example, Bell et al. (2007) used government voting records as well as census information to analyze the issue of one-party dominance in Alberta. The state is also the source of a great deal of textual material of potential interest, such as official reports. For instance, in his study of the issues surrounding synthetic bovine growth hormone, Jones (2000) used transcripts of a Canadian senate inquiry on the topic. Briefly, Monsanto, the manufacturer, lost the battle to have the hormone accepted for use; health groups claimed that it was unnecessary and, worse, that it posed a health risk both to cows and to the humans who drank their milk. Abraham (1994) used similar materials in his research on the role of self-interest and values in scientists' evaluations of the safety of medicines, specifically the drug Opren. The author described his sources as "publicly available transcripts of the testimonies of scientists, including many employed in the manufacture of *Opren*, Parliamentary debates, questions and answers in *Hansard*, and leaflets, letters, consultation papers and other documentation disposed by the [drug regulatory authority]" (Abraham, 1994, p. 720). His research showed inconsistencies in the scientists' testimonies, suggesting that self-interest can play an important role in such situations. He also used his findings to

BOX 12.4 Photographs of the Magic Kingdom

Stellajune3700/iStock photo

This family seems to be enjoying their vacation, but should we conclude that their holiday was fun in every respect? How can the knowledge that a photo may not provide a complete record of an event or experience affect how we approach the use of photographs as data?

Sutton (1992) noted a paradox about visitors to Disney theme parks. The Magic Kingdom is supposed to be "the happiest place on Earth," and employees ("cast members") are trained to enhance visitors' experiences. Yet it is clear that some people do not enjoy themselves while visiting. Time spent waiting in lines is a particularly common gripe for "guests" (Bryman, 1995). Nonetheless, because people expect their visit to be momentous, they take photographs that support their assumption that Disney theme parks are happy places. When they return home, they "discard photographs that remind them of unpleasant experiences and keep those that remind them of pleasant experiences" (Sutton, 1992, p. 283). In other words, positive feelings about an experience may be a post hoc reconstruction, substantially aided by the photos that are not thrown away or deleted. Thus photographs may provide an incomplete, somewhat distorted record.

Similar observations can be made with regard to pictures posted on social media. How common is it to see Facebook photographs of people having unpleasant experiences on their holidays? With social media, the purpose of photo-assisted reconstruction seems to be expanding: now photos may be used not only to create happy memories, but also to present a public image of oneself as a fun-loving, interesting person. Clearly, holiday photos represent more than an objective account of a vacation.

argue that the scientific ethos of "objectivity" has limited applicability in areas where self-interest may be a factor.

In terms of Scott's (1990) four criteria, such materials can certainly be seen as authentic and as having meaning (in that they are clear and comprehensible to the researcher). However, the credibility criterion requires us to consider whether the documentary source is biased. This is exactly the point of Abraham's (1994) research: such documents can be interesting precisely because of the bias they reveal. Equally, this point suggests that caution is necessary in attempting to treat them as depictions of reality.

Official documents from private sources

Official documents available from "private sources" vary widely, but a common category is company documents. Companies (and organizations generally) produce many documents, some of which are in the public domain; the latter include annual reports, press releases, advertisements, and public relations material both in printed form and on the Web. Other documents, such as company newsletters, organizational charts, minutes of meetings, memos, correspondence (internal and external), and manuals for new recruits, may not be accessible to the public. This kind of material is often used by organizational ethnographers in their investigations, but the difficulty of gaining access to it means that many other researchers have to rely on public domain documents. Even if the researcher is an insider with the organization, certain private documents may remain inaccessible.

Private documents also need to be evaluated using Scott's four criteria. As with the materials considered in the previous section, documents deriving from private sources such as companies are likely to be authentic and meaningful (in the sense of being clear and comprehensible to the researcher). However, issues of credibility and representativeness are still likely to require scrutiny.

People who write documents generally want to convey a particular point of view. An interesting illustration is provided by a study of career development issues in a major retail company (Forster, 1994). Forster analyzed company documentation as well as interviews and a questionnaire. Because he was able to interview many of the authors of the documents, "both the accuracy of the documents and their authorship could be validated by the individuals who had produced them" (1994, p. 155). In other words, the authenticity of the documents was confirmed and apparently the credibility as well. However, Forster also said that the documents revealed divergent interpretations of key events and processes:

One of the clearest themes to emerge was the apparently incompatible interpretations of the same events and processes among the three sub-groups within the company—senior executives, HQ personnel staff, and regional personnel managers. . . . These documents were not produced deliberately to distort or obscure events or processes being described, but their effect was to do precisely this. (1994, p. 160)

The perspectives that members of the different groupings expressed in the documents reflected their positions in the organization. Consequently, although the authors of the documents were able to confirm their authenticity, their contents could not be regarded as "free from error and distortion," as Scott put it. In other words, the accounts that documents provide may not be completely objective. They have to be analyzed critically and compared with other sources of data. As Forster's study suggests, the different stances taken by the authors of documents can be used to develop insights into the processes and factors that lie behind their creation.

Issues of representativeness are also important. Did Forster have access to a complete set of documents? It could be that some had been destroyed, or that he was not allowed access to certain sensitive ones. This is not to say that such documents necessarily exist, but a healthy skepticism is often warranted.

Mass media outputs

Newspapers, magazines, television programs, films, and other mass media are potential sources for social scientific analysis. For example, Parnaby's (2003) study of how Toronto tried to deal with its squeegee kids examined 200 newspaper articles that appeared between 1995 and 2000. Hallgrimsdottir et al. (2006) looked at how newspapers in Victoria, BC, characterized sex-trade workers in the province, while Harding (2006; 2010) analyzed press coverage of Indigenous people and their institutions. Bhuyan et al. (2017) used documents published by Citizenship and Immigration Canada as well as articles from the *Toronto Star*, the *Globe and*

Syrian refugees in Manitoba

From 1 January 2015 to 30 April 2017, 77,090 refugees arrived in Canada, 6305 of whom made their homes in Manitoba. Of the latter group, about 60 per cent were of Syrian origin (Olukoju, 2017).

University of Manitoba sociologist Lori Wilkinson led a research team examining how well Syrian refugees are adapting to their new surroundings in Manitoba. The researchers report that over 80 per cent of the newcomers are satisfied with their current housing. "Housing Syrian refugees in Manitoba is a good news story," Wilkinson told the *Winnipeg Free Press*. She added that housing refugees in permanent, safe accommodations rather than temporary housing is particularly important to their well-being. Permanent housing makes it easier for adults to get into language classes, and it expedites children's enrollment in school. Acquisition of the host language is crucial for their integration into the community, and it is something that virtually all the new arrivals had to contend with as over 97 per cent could not speak English or French when they arrived.

Mail, and the *National Post* to investigate how the use of language in those sources is related to the development of social ideologies and the implementation of government policies on immigration. All of these studies are examples of **critical discourse analysis** (discussed later in this chapter) in that they examined the social and political implications of the materials they examined, an approach that can be taken with virtually any medium of communication.

Films, television shows, and magazines have similar potential for research. For example, when Coté and Allahar (1994) examined magazines aimed at adolescents, they concluded that these "teenzines" turned adolescents into uncritical consumers and diverted them from protesting against their lack of adult privileges.

Authenticity is sometimes difficult to ascertain in the case of mass media outputs. The authors are not always identified, as in the case of a TV news report, so it is sometimes difficult to know whether a given account was prepared by someone in a position to know all the facts. Credibility is frequently an issue, and as the examples used in this section show, it is often the uncovering of error or distortion that is the objective of the analysis. Representativeness may not be an issue with newspaper or magazine articles, since many publications make a point of maintaining a consistent tone or ideological bent. Finally, although the literal meaning of mass media outputs is often clear, it usually takes some reflection and theoretical analysis to appreciate the broader societal impact these forms of communication can have.

Virtual outputs and the Internet as objects of analysis

"Text" is a word that has up to this point been used sparingly in this chapter. It originally referred to a written document, but in recent years it has been applied to an increasingly wide range of phenomena including theme parks, technologies, paintings, and buildings. Contemporary writers and researchers regard such "texts" as materials that can be interpreted to produce "readings."

Websites and webpages are further examples of "texts" that can be "read." The vastness of the Internet and its growing accessibility make it a valuable source of documents for both quantitative and qualitative data analysis. Hier (2000), for example, examined a Toronto-based racial supremacy website and found that it allowed people access to racist ideas in a relatively anonymous way (e.g., without having to subscribe to a hard-copy newsletter). For that reason he feared that it might be more successful in reaching ordinary citizens than non-Internet sources would be.

The use of images on websites can also be quite revealing. Crook and Light (2002) analyzed the photographs in 10 university prospectuses and found that although there were many photos of students apparently studying, the setting was rarely a library carrel or a dorm room. Instead, students were usually shown in "social" learning situations that allowed them to appear active and engaged, frequently out of doors. The authors argued that those less typical learning contexts were chosen because they are more seductive. In a study of Web-based family photography, Pauwels (2008) concluded that digital technology has made it easier for families to "construct fictions and fantasies" and to project their values and norms than was the case when paper-based family albums were more common.

Other forms of Internet-based communication, such as email lists, discussion groups, and chat rooms, have also been used as objects of analysis. For their study of online social support groups, Nettleton et al. (2002) examined the interactions between people who use those sorts of virtual communications. Similarly, social media such as Facebook, Twitter, and YouTube have been analyzed by social researchers. Clayton et al. (2013), for instance, found evidence that Facebook use may lead to conflict between romantic partners that can contribute to negative relationship outcomes such as cheating. Even sexting (sending sexual text messages) has been the subject of analysis (Drouin et al., 2014).

Websites clearly have huge potential, but it's important to keep Scott's criteria in mind. First, authenticity: anyone can set up a website, so you have no guarantee that a person offering information (such as financial advice) is an authority. Second, credibility: is the information on the site credible, or might it be distorted for some reason? For example, a site that encourages you to buy stock held by its author might exaggerate its value. Third, given the constant flux of the Internet, it's doubtful that one can ever know how representative websites on a certain topic are. New websites are continually appearing, others disappearing, still others being modified. Online searching is like trying to hit a target that is not only moving but also in a constant state of metamorphosis. In a related vein, any one search engine can provide

Practical Tip | Referring to Websites

It is becoming the common practice in academic work, when referring to websites, to include the date they were consulted. This convention is closely associated with the fact that websites often disappear and frequently change, so that if subsequent researchers want to check or follow up a Web reference, they may find that the site has changed or even disappeared. Citing the date(s) the site was visited can help to relieve anxieties about citing Web sources.

access to only a portion of the Web; there is evidence that even the combined use of several search engines gives access to just under half of the sites that exist, and there is no way of knowing if they are a biased sample. Finally, websites are notorious for a kind of "webspeak" that makes it difficult to comprehend what is being said without some insider knowledge.

Researchers basing their investigations on websites need to recognize these limitations as well as the opportunities available. Scott's suggestions invite consideration of why a website was constructed in the first place: for commercial reasons? Political reasons? In other words, you should be no less skeptical about websites than about any other kind of document. Employing both traditional printed documents and website materials will allow you to cross-validate sources.

Beyond Scott's criteria

So far, most of the biases discussed in the context of Scott's criteria have been of the personal type—someone may write a flattering if somewhat unrealistic autobiography, pictures for public display may be chosen in order to create a favourable impression, and so on. Keep in mind that there may be larger cultural and political biases involved in the creation and use of any type of text or document. Hales (2006), for example, points out that a colonialist mentality often pervades research conducted on subject or subordinated peoples. Where there is a power differential

> ### ✓ Checklist
>
> ### *Checklist for evaluating documents*
>
> Have the following questions been answered?
>
> ☐ Who produced the document?
>
> ☐ Why was the document produced?
>
> ☐ Was the person or group who produced the document in a position to write authoritatively about the subject?
>
> ☐ Is the material genuine?
>
> ☐ Did the author have an axe to grind or a particular slant to promote?
>
> ☐ Is the document typical of its kind? If not, is it possible to establish how atypical it is and in what ways?
>
> ☐ Is the meaning of the document clear?
>
> ☐ Can the events or accounts presented in the document be corroborated?
>
> ☐ Are there different interpretations of the document from the one you offer? If so, what are they? Have you discounted them? If so, why?

between the researcher and those who are being studied, one would do well to ask, "Why is the research done? How is it done? Who defines, initiates and conducts the research? On/with/for whom is the research carried out? What topics are addressed? Who benefits and how? Who interprets for whom and who represents whom?" (Hales, 2006, p. 243). Many sources of information, including much of the research in the social sciences, serves some larger if unstated political or cultural purpose. Kempf (2006), for example, in an analysis of history textbooks and Ontario Ministry of Education guidelines from 1860 to 2006, concludes that a lot of this material showed a disrespect for Indigenous perspectives and ways of knowing, and in effect provided a rationalization for the mistreatment of Indigenous Canadians.

What things need to be analyzed?

Obviously, what is to be analyzed depends on the specific research questions. In mainly quantitative studies, these elements are usually specified in advance in order to guide both the selection of the media to be assessed and the construction of the coding schedule. They usually include words, subjects and themes, and value positions, as discussed next.

Words

Determining the frequency with which certain words are used is often the first step in content analysis. Jagger (1998), for example, searched dating advertisements and counted words such as "slim" and "non-smoker" to compare women and men with respect to what each deemed desirable in a date (see Box 12.5). Simple counting of particular words can reveal emphasis, style of writing or presentation, and even the overplaying of certain events. For example, Dunning et al. (1988) noted a tendency for the British press to sensationalize disturbances at soccer matches by using emotive words such as "hooligans" and "louts" to refer to audience members, and terms such as "battle" to refer to a game; less dramatic terms such as "fans" or "contest" would have encouraged more neutral responses among readers.

A variation on the search for individual keywords is the search for pairings of keywords. The growing availability of the written news media online greatly facilitates this kind of search. Hier (2002) found "rave" and "drug use" to be frequently linked, which may have encouraged readers to believe that raves must be controlled. Parnaby (2003), on the other hand, noted that "squeegee kids" and "homelessness" were not paired, even though many of the "kids" were in fact homeless; as a result of this omission,

BOX 12.5 Finding love: Then and now

Jagger (1998) reported a content analysis of 1094 dating advertisements in four newspapers with a general readership. The sample was chosen over two four-week periods in 1996. Three research questions drove the study:

- What is "the relative significance of monetary resources and lifestyle choices as identity markers and desirable attributes for men and women"?
- How do men and women vary in the ways they choose to market themselves and describe their preferred (or ideal) partners in terms of the body?
- To what extent are "traditional stereotypes of masculinity and femininity . . . still in operation" (1998, p. 799)?

 Jagger noted the tendency for considerable percentages of both men and women to market themselves in terms of their lifestyle choices (film preferences, clubs frequented, sports activities, etc.). She also found that women were far more likely than men to stress the importance of economic and other resources in a preferred partner. There was also a somewhat greater propensity for women to market themselves in terms of physical appearance. As an aside, men were just as likely to market themselves as "slim," suggesting that certain preferences in body shape were not exclusive to one gender. More generally, her results pointed to the significance of the body in identity construction for both men and women. In a later publication, Jagger (2001) reported the findings from a qualitative content analysis of a sub-sample of the 1094 advertisements.

 Much has changed in the years since Jagger did her research. The Internet dating site has eclipsed the classified newspaper advertisement as the medium of choice for people whose in-person activities do not provide adequate opportunities to meet a romantic partner. Since space is not at a premium, such sites allow participants to give far more information than was the case with newspaper ads, and of course the provision of photographs means that verbal descriptions of physical appearance have become far less important than they once were. However, many sites still feature "dating profiles" in which people have a few paragraphs to describe themselves and to outline what they would like to find in a partner, and the content of these profiles is somewhat similar to the older-style ad descriptions. There are even sites that provide tips on writing a good dating profile.

 Not surprisingly, a major concern with people who engage in online dating is the misrepresentation of personal characteristics by potential mates (Hall et al., 2010). Most users want an in-person meeting before getting involved in a romantic relationship (Finkel et al., 2012), and not without reason. Hitsch et al. (2010) compared the information provided in online dating sites with American national averages, and found that both men and women on the sites claimed higher than average heights, and that women but not men indicated lower than average weights. Using self-reports of probable misrepresentation on dating sites, Hall et al. (2010) found that men were more likely than women to be deceptive about their personal assets (e.g., income), but that there was no gender difference in misrepresentation of relationship goals (e.g., interest in pursuing a serious relationship). Women were found to be more likely to misrepresent their weight, while men, contrary to expectation, were more likely to mislead about their age.

he argued, the "kids" were constructed "as a social problem requiring a law and order resolution" (2003, p. 281). The search for pairings of keywords can be a starting point for a more in-depth analysis.

Subjects and themes

A more interpretative approach is required to code text in terms of subjects and themes. At this point, the analyst is searching not just for the obvious or *manifest* content, but also for some of the underlying or *latent* content as well. This occurs if the researcher wants to probe beneath the surface to ask deeper questions about what is happening. In reports on crime, for example, is the victim blamed along with the accused? Is the occupation of the accused or the victim stated? If not, is it implied in an address or a picture? Why is it that news stories about men who have been mugged sometimes mention their marital status? In seeking to answer these sorts of questions, the quantitative aspects of content analysis can serve as a starting point for a study in which qualitative research becomes the central focus of the endeavour.

Value positions

A further level of interpretation is likely when the researcher is seeking to demonstrate that the writer of a particular text has taken a certain value position. For example, is a journalist who writes about crime sympathetic or hostile to the accused? Is all the blame put on the accused, with the implication that punishment is appropriate? Or is the focus on social conditions, with the assumption that less blame should be placed on the criminal? If there is no manifest indication of the writer's value positions, can inferences be made from the latent content?

Another way in which content analysis can reveal value positions is through the coding of ideologies, beliefs, and principles. For instance, Sturgeon and Morrissette (2010) examined suicide ideation among Manitoba farmers who called a rural crisis line. The data were gathered from forms that were filled out by the crisis line counsellors (which are similar in design to the coding manual depicted in Figure 12.2). One of the themes that emerged as codes were developed was "family salvation," which involved the belief that life is worth living if one has

meaningful relationships with family members. "I just keep going because I think someday my son will need me," said one farmer (Sturgeon & Morrissette, 2010, p. 199). Surprisingly, only a small proportion of farmers mentioned their spouses or life partners as the people who made their lives bearable; others such as siblings, children, and extended family members were more likely to be cited.

Coding

As the foregoing has implied, coding is a crucial part of content analysis. There are two main elements to a content analysis coding scheme: designing a coding schedule and creating a **coding manual**. To illustrate, imagine that you are interested in crime reports in a local newspaper. To simplify, assume that your study is limited to crimes where the victim is a person rather than an organization, and that it considers these variables:

1. nature of the offence(s)
2. gender of suspect
3. occupation of suspect
4. age of suspect
5. gender of victim
6. occupation of victim
7. age of victim
8. depiction of victim
9. position of news item in the paper

Content analysts are normally interested in a much larger number of variables, but this simple illustration can help to get across how a coding schedule and coding manual operate.

Coding schedule

The coding schedule is a form onto which the data are entered (see Figure 12.1). Each of the columns in Figure 12.1 is a dimension (indicated by the column heading) to be coded. The blank cells on the coding form are the places in which codes are to be written. One row is used for each media item coded. The codes can then be transferred to a computer data file for analysis with a software package such as **IBM SPSS Statistics** (SPSS).

Case number	Day	Month	Nature of offence I	Gender of suspect	Occupation of suspect	Age of suspect	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item

FIGURE 12.1 Coding schedule

Coding manual

On the face of it, the coding schedule in Figure 12.1 seems bare, providing little information about what is to be done or where. This is where the coding manual comes in. It is a set of instructions to coders that includes all possible categories for each dimension to be coded. It provides a list of all the dimensions; the different categories subsumed under each dimension; the numbers (that is, *codes*) that correspond to each category; and guidance to coders on what should be taken into account in coding a particular dimension. Figure 12.2 shows a coding manual that corresponds to the coding schedule in Figure 12.1.

Our coding manual includes the occupation of both the suspect and the victim of the crime, using a simple social-class scheme. The offences are classified according to the categories used by the police in recording crimes. (Since police statistics have been criticized for weak **reliability** and **validity**—recall Chapters 4 and 5—a comparison between police data and crime reports in local newspapers is a possible research topic.) Finer distinctions can be used, but unless you are planning to examine a large sample of news items, broader categories are preferable. Note also that the coding schedule and manual permit the recording of two offences for any incident. If there are more than two, you will have to make a judgment concerning the two most significant offences.

The coding manual is crucial because it provides a complete listing of all categories for each dimension to be coded, as well as guidance on how to interpret the dimensions. It is on the basis of these lists that a coding schedule of the kind presented in Figure 12.1 is created. Detailed rules about how to code should be formulated, because reliability,

both inter-coder and intra-coder, is always a concern (see below).

Here is a news report of a fictional road-rage incident. Two male motorists, one a retired schoolteacher aged 68, the other a 26-year-old assembly-line worker, got into an argument and the worker allegedly punched the retired man, causing him to fall, hit his head, and suffer a concussion. There was no second offence. The coding of the incident would then appear as in Figure 12.3 and the data would be entered into a computer program such as SPSS.

Suppose a second article, appearing the next day, described how an unemployed 34-year-old female reportedly took an 86-year-old woman's purse and then knocked her down. The code is provided in Figure 12.4 but with a few errors. Can you spot them? Forms like these would be completed for each news item within the chosen period(s) of study.

Potential pitfalls in devising coding schemes

The potential dangers in devising a content analysis coding scheme are similar to those involved in designing **structured interview** and **observation schedules**. Points to recall include the following:

- *Mutually exclusive categories*. There should be no overlap in the categories supplied for each dimension. If the categories are not mutually exclusive, coders will not know how to code an item that fits into more than one category.
- *Exhaustive categories*. Every possible dimension should have a category.
- *Clear instructions*. Coders should be clear about what factors to take into account when

Nature of offence I

01.	Violence against the person
02.	Sexual offences
03.	Robbery
04.	Burglary in a dwelling
05.	Burglary other than in a dwelling
06.	Theft from a person
07.	Theft of bicycle
08.	Theft from shops
09.	Theft from vehicle
10.	Theft of motor vehicle
11.	Other theft and handling stolen goods
12.	Fraud and forgery
13.	Drug offences
14.	Other offences

Gender of suspect

1.	Male
2.	Female
3.	Unknown

Occupation of suspect

01.	Professionals, administrators, officials and managers in large establishments; large proprietors
02.	Lower-grade professionals, administrators, and officials; higher-grade technicians; managers in small business and industrial establishments; supervisors of non-manual employees
03.	Routine non-manual employees in administration and commerce
04.	Personal service workers
05.	Small proprietors, artisans, etc., with employees
06.	Small proprietors, artisans, etc., without employees
07.	Farmers and smallholders; self-employed fishermen
08.	Lower-grade technicians, supervisors of manual workers
09.	Skilled manual workers
10.	Semi-skilled and unskilled manual workers (not in agriculture)
11.	Agricultural workers

12.	Other
13.	Unemployed
14.	Homemaker
15.	Student
16.	Retired
17.	Unknown

Age of suspect

Record age (–1 if unknown)

Gender of victim

1.	Male
2.	Female
3.	Unknown

Occupation of victim

Same as for occupation of suspect

If not applicable, code as 99

Age of victim

Record age (–1 if unknown)

Depiction of victim

1.	Victim responsible for crime
2.	Victim partly responsible for crime
3.	Victim not at all responsible for crime
4.	Not applicable

Nature of offence II (code if second offence mentioned in relation to the same incident; code 0 if no second offence)

Same as for Nature of offence I

Position of news item

1.	Front page
2.	Inside
3.	Back page

FIGURE 12.2 Coding manual

Case number	Day	Month	Nature of offence I	Gender of suspect	Occupation of suspect	Age of suspect	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item
001	24	11	01	1	10	26	1	16	68	2	00	2

FIGURE 12.3 Completed coding schedule

Case number	Day	Month	Nature of offence I	Gender of suspect	Occupation of suspect	Age of suspect	Gender of victim	Occupation of victim	Age of victim	Depiction of victim	Nature of offence II	Position of news item
002	25	12	06	1	13	32	2	16	86	3	01	2

FIGURE 12.4 Completed coding schedule with errors

assigning codes. Sometimes these have to be very elaborate. In the sort of content analysis described in the previous section, coders generally have little or no discretion in deciding how to code the units of analysis.

- *A clear unit of analysis.* For example, in the imaginary study of the reporting of crime in the local press, there is both a media item (for example, one newspaper article) and a topic to be coded (one of two offences). In practice, a researcher is interested in both but needs to keep the distinction in mind.

To enhance the quality of a coding scheme, it is advisable to conduct a pilot study to identify difficulties, such as finding there is no code to cover a particular case (not exhaustive). A pilot test will also help to reveal if one category of a dimension includes an extremely large percentage of items. When this occurs, it may be necessary to break that category down to allow for greater specificity.

As we have mentioned, the reliability of coding is a further concern. An important part of pre-testing the coding scheme is examining consistency between coders (**inter-coder reliability**) and, if time permits, **intra-coder reliability**. The process of gauging reliability is more or less the same as in **structured observation**, discussed in Chapter 6.

Content analysis without a pre-existing coding scheme

Content analysis may be undertaken before any specific decisions have been made regarding how the information available for analysis will be coded. In such cases the data are searched in order to find underlying themes, making the study primarily qualitative in nature. Kennedy (2006), for example, in his study of the Canadian fathers' rights movement called "Fathers For Justice," used content analysis in conjunction with participant observation and semi-structured interviews to examine how the movement developed a collective identity. The content analysis was conducted on documents such as position papers submitted to government agencies, newsletters, pamphlets, and minutes from meetings, and was used mainly to make sense of the data gathered through participant observation and interviews. Similarly, Bell (2007) used pamphlets, speeches, newsletters, and other texts to arrive at an overall assessment of western Canadian separatism as a neo-liberal movement. The processes by which the themes are extracted in this kind of content analysis are often left implicit, although they are usually illustrated with quotations from the text in question.

Lynch and Bogen (1997) examined core sociological textbooks and found that they contained recurring themes that presented an upbeat and scientific view of the discipline—one that, in the researchers' view, was biased and value laden. Seale (2002) examined newspaper reports about people with cancer. One of the phases of his analysis entailed an "NVivo coding exercise, in which sections of text concerning themes of interest were identified and retrieved" (Seale 2002, p. 109). He was especially interested in gender differences in the representation of sufferers, and demonstrated that stories about men were much more likely than stories about women to discuss the role that the individual's character played in dealing with the disease.

Altheide (1996) outlined an approach he called *ethnographic content analysis* (ECA). He described his approach as differing from quantitative content

analysis in that the researcher is constantly revising the themes or categories distilled from the examination of documents. As he put it:

> ECA follows a recursive and reflexive movement between concept development-sampling-data, collection-data, coding-data, and analysis-interpretation. The aim is to be systematic and analytic but not rigid. Categories and variables initially guide the study, but others are allowed and expected to emerge during the study, including an orientation to **constant discovery** and **constant comparison** of relevant situations, settings, styles, images, meanings, and nuances (1996, p. 16; emphases in original).

While more quantitatively oriented analysis typically entails applying predefined categories to the sources, ECA allows for greater refinement of those categories as well as generation of new ones. For example, Parnaby's (2003) study of squeegee kids included a reflexive examination of the documents used (newspapers, magazines, government documents, official reports), which was done while the process of forming and confirming theoretical concepts was still under way; it was a process of continuing discovery and comparison. In addition, ECA emphasizes the context within which documents are generated. For instance, a study of newspaper reports of crime would require some understanding of news organizations and the work that journalists do (Altheide, 2004).

Semiotics

Another form of predominantly qualitative content analysis is semiotics, the *science of **signs***. In social research, semiotics involves analysis of the signs and symbols encountered in everyday life. It can be applied to documentary sources as well as many other kinds of data. The main terms employed in semiotics are as follows:
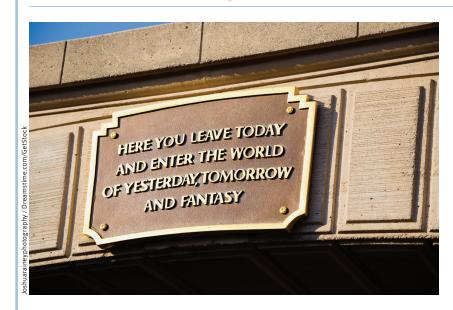
- The *sign* is something that stands for something else, such as a yellow traffic light; it has two components, a *signifier* and a *signified*.

- The *signifier* is the thing (here the yellow light) that points to an underlying meaning.
- The *signified* is the meaning to which the signifier points ("caution: stop if possible").
- The **denotative meaning** is the manifest or obvious meaning of a signifier and as such indicates its function (here to regulate traffic).
- A **connotative meaning** is one that can arise in conjunction with the denotative meaning (for example, "speed up to beat the coming red light").
- *Polysemy* refers to the notion that signs can be interpreted in many different ways.

Semiotics seeks to uncover the hidden meanings that reside in texts, broadly defined. Consider, by way of illustration, the curriculum vitae (CV) in academic life. The typical academic CV contains information on matters such as degrees earned; previous and current posts; administrative responsibilities and experience; teaching experience; research experience; research grants acquired; and publications. It can be seen as a system of interlocking signifiers denotatively providing a summary of the individual's experience (its sign function). At the connotative level, it is an indication of an individual's value, particularly in connection with employability. Every CV is capable of being interpreted in different ways and is therefore polysemic, but there is a shared code among academics whereby certain attributes of CVs (such as lists of publications) are seen as especially desirable and are therefore less contentious in terms of their meaning. Applicants for posts know this and design their CVs to highlight the desired qualities and downplay others, so that the CV becomes a presentation of self, as Miller and Morgan (1993) suggested.

Box 12.6 outlines a semiotic study (Gottdiener, 1997) in which the "text" was Disneyland. The chief strength of semiotics lies in its invitation to see beyond and beneath the apparent ordinariness of everyday life and its manifestations. The main difficulty with semiotic analysis is that, although it may offer a compelling exposition of some aspect of everyday life, the interpretation may be somewhat arbitrary. For example, Gottdiener's association of "Adventureland" with colonialism/imperialism might not be shared by other analysts. However, the results of a semiotic analysis

BOX 12.6 A semiotic Disneyland

Joshuaraineyphotography / Dreamstime.com/GetStock

Visitors see this plaque as they enter Disneyland in California. How does it support Gottdiener's notion of a semiotic Disneyland?

Gottdiener (1997, pp. 108–115) proposed that Disneyland can be fruitfully explored through semiotic analysis. He concluded that its meaning is based on the contrast between the alienated daily lives of nearby Los Angeles residents and the life they experience at the theme park. He identified, through this principle, several sign systems contrasting the park with the surrounding environment: transportation, food, clothing, shelter, entertainment, social control, economics, politics, and family. The first of these sign systems—transportation—revealed a difference between the Disneyland visitor as pedestrian (walking with others in a group) and the Los Angeles resident as passenger (a car is necessary; danger on the congested freeways). A further component of his research entailed analysis of the connotations of the different "lands" that make up the park. He suggested that each land was associated with one or more signifiers of capitalism, for example:

- Frontierland: predatory capital, conquering
- Adventureland: colonialism/imperialism
- Tomorrowland: state capital for space exploration

may be no more arbitrary than those of any other interpretive approach. Indeed, it would be surprising not to be struck by a sense of arbitrariness in semiotics, given its emphasis on the principle of polysemy.

Hermeneutics

The hermeneutic approach has been used in understanding and interpreting the Bible, but it has also been employed in the analysis of other texts. It was influential in the formation of interpretivism as an **epistemology** (see Chapter 1) and has much in common with Weber's notion of *Verstehen*. The central idea behind hermeneutics is that the analyst of a text must seek its meanings from the perspective of its author, considering the social and historical context within which the text was produced.

Phillips and Brown's (1993) hermeneutic study of the corporate image advertisements of a Canadian

synthetic crude oil company was a "formal analysis of the structural and conventional aspects of the text" (1993, p. 1563). For them this meant examining the texts of the advertisements in terms of their constituent parts and the writing conventions they reflected. They also relied on a large database of magazine and newspaper articles relating to the company for additional documentary materials. They showed how corporate image advertisements were used to mobilize support for company activities from government and the public, and to ward off the environmental legislation that threatened them.

Readers and audiences— active or passive?

Audience reception is a prominent area of inquiry in media and cultural studies. The key question is whether audiences/readers are active interpreters of what they see or hear. Do they passively accept the meanings that authors or designers infuse into their texts, as in the oil company advertisements just described, or do they resist those meanings and strive for independent readings? Do they arrive at a middle point that incorporates both passive and active elements? Much of the research on this issue suggests that audiences frequently come up with readings different from those intended by authors (see Fenton et al., 1998, for a summary of some of this research).

Although the idea of the "active audience" has not gone unchallenged (see, for example, McGuigan, 1992), it is so well supported that many of the interpretations of texts offered by social scientists have been questioned. This suggests caution in reading Giulianotti's (1997) study of "fanzines" or Hier's (2002) examination of city council minutes to understand the rave issue. Would other social scientists arrive at the same interpretations? Do their interpretations match those of the original readers or audiences? The social researcher is always putting a personal "spin" on the texts analyzed. The same is true of all social science data: the conclusions derived from questionnaire or ethnographic data reflect the author's interpretations, not the complete set of possible interpretations. The point is that close scrutiny and critical thinking are always required.

Two approaches to the study of language

In this section two approaches to the study of language are examined: conversation analysis (CA) and discourse analysis (DA). While CA and DA do not exhaust the range of possibilities for studying language, they do represent two of the most prominent approaches, and each includes both quantitative and qualitative aspects. Both have developed a technical vocabulary and a set of techniques. This section outlines some of their basic elements and draws attention to their contrasting features.

Conversation analysis

The roots of CA lie in ethnomethodology, a sociological approach to communication focusing on the "practical, common-sense reasoning" that people use in their everyday lives. It includes notions of cause and effect ("if I do this, then that will happen") and the generalizations that allow people to perform everyday tasks. This process of reasoning and communication is presented as a way in which social order is created. Social order is not seen as a pre-existing force constraining individual action, but as something that is worked at and achieved by people through interaction. Garfinkel, one of the pioneers of conversation analysis, claimed that the role of sociology is not to uncover "objective" social facts as Durkheim suggested, but to see them as an accomplishment, as the eventually taken-for-granted patterns established through the activities of ordinary people going about their daily lives (Garfinkel, 1967, p. vii).

Two ideas are particularly central to ethnomethodology and find clear expression in CA: *indexicality* and reflexivity. The former suggests that the meanings of words or utterances, including pauses and sounds, depend on the context in which they are used. Reflexivity means that talk is not a "mere" representation of the social world, standing for something else, but is itself a reality. In these ways, ethnomethodology fits squarely with two aspects of qualitative research: a preference for a contextual understanding of action (see Chapter 9) and an

ontological position associated with **constructionism** (see Chapter 1).

In the years following its introduction, ethnomethodological research sought to conduct fine-grained analyses of the sequences of interaction revealed in conversations recorded in naturally occurring contexts. As such, CA is a multifaceted approach, part theory, part method of data acquisition, part method of analysis. The predilection for the analysis of talk gleaned from naturally occurring situations suggests that CA fits well with another preoccupation among qualitative researchers: a commitment to **naturalism** (see Chapter 9).

Conversation analysts have developed a variety of ways to study talk. Psathas (1995, p. 1) described them as "rigorous, systematic procedures" that can "provide reproducible results." Such a framework brings to mind the procedures used to generate valid, reliable, and replicable findings in quantitative research. It is not surprising, therefore, that CA is sometimes described as having a **positivist** orientation. Thus some features that are broadly in tune with qualitative research (contextual and naturalistic analysis without prior theoretical commitments) are married to traits associated with quantitative research.

However, the emphasis on context in CA is somewhat at variance with contextual understanding as it is normally thought of in qualitative research. For CA practitioners, context refers to the specific here-and-now context of immediately preceding talk, whereas for most qualitative researchers it has a much wider focus, encompassing phenomena such as the broader culture of the group within which the social action occurs, including their values, beliefs, and typical modes of behaviour. This is precisely the kind of attribution that CA practitioners seek to avoid. To import elements that are not specifically grounded in the here and now of what has just been said during a conversation is to risk imposing an understanding that is not grounded in the participants' immediate concerns. It is no wonder, therefore, that writers like Gubrium and Holstein (1997) treated CA as a separate tradition within qualitative research, whereas Silverman (1993) found it difficult to fit CA into broad descriptions of qualitative research.

Research in the News

Changes in Canadian households

For the first time ever, one-person households are the most common living arrangement in Canada. This development was revealed in a recent data release by Statistics Canada, which showed that about 28 per cent of all households in the country are made up of people living alone (Lee-Young, 2017).

What might account for the change? According to Yue Qian, a sociologist at the University of British Columbia, there are "multiple, but also perhaps competing, reasons for the rise in one-person households." "The population is aging," he said, "and elderly people are more likely to live alone. At the other end, young people are delaying marriage, as well as staying single" (Lee-Young, 2017). Professor Qian also noted that if the price of real estate in major metropolitan areas hadn't increased so markedly in recent years, the percentage of people living on their own may have been even higher. Andy Yan, director of the City Program at Simon Fraser University, added that "there is also the impact of changing gender roles, which are shifting the picture of age and living on your own" (Lee-Young, 2017). Marina Adshade, an economist at the University of British Columbia, has pointed out that there are important consequences to the shift in living arrangements. For example, "when we release a 50-year-old man from the hospital after a heart attack, it used to be that he could go into the care of his wife. What happens now? The census is about finding out where we should allocate resources. It's really important" (Lee-Young, 2017).

Assumptions of conversation analysis

Analysts often begin practising conversational analysis when they notice something significant about the way a speaker says something, and this recognition generates interest in the function that the turn of phrase serves. To illustrate, Clayman and Gill (2004) referred to the way a child will often begin a conversation with an adult by saying "You know what, Daddy [or whoever]?" The question generally produces a "What?" reply that allows the child to find a slot in a sequence of communications, or to inaugurate such a sequence. The use of a question, rather than a declarative statement, may reflect the child's lesser power.

Once such a focus has been identified, conversation analysts typically follow certain basic assumptions. Heritage (1984; 1987) proposed three:

- *Talk is structured*. Talk comprises patterns, and participants are implicitly aware of the rules that underpin these patterns. As a result, conversation analysts do not attempt to infer the motivations of speakers from what they say, or to ascribe their talk to purely personal characteristics. Such information is unnecessary, since the conversation analyst is oriented to the underlying structures that are revealed in pauses, emphases, questions followed by answers, and so on.
- *Talk is forged contextually*. Talk must be analyzed in terms of its context and understood in terms of the talk that has preceded it.
- *Analysis is grounded in data*. Conversation analysts shun prior theoretical schemes and argue that the characteristics of talk and social order must be derived from the data.

In doing a project based on CA, it is important not to collect too much data. The real work of CA is the painstaking analysis that its underlying theoretical stance requires. It may be that just one or two portions of transcribed text will be sufficient to answer the research questions.

Transcription and attention to detail

The transcript in Box 12.7 includes some basic symbols used by conversation analysts:

> We:ll   A colon indicates that the sound that occurs directly before the colon is prolonged. More than one colon means further prolongation (e.g., Ye: : : :s).

> .hh   A row of h's preceded by a dot indicate an intake of breath. If no dot is present, it means breathing out. The more h's, the longer the breath.

> (0.8)   A figure in parentheses indicates a period of silence, usually measured in tenths of a second. Thus (0.8) indicates eight-tenths of a second of silence.

> you and knowing   An underline indicates an emphasis in the speaker's talk.

> you- you   A hyphen means the speaker was interrupted or stopped speaking.

> (.)   Indicates a very slight pause.

The attention to detail in the sequence of talk in Box 12.7 is striking, but what is significant in it? Silverman (1994) drew two main inferences. First, *P* initially tries to deflect any suggestion that there may be a special reason why she needs an HIV test. As a result, the disclosure that she has been engaging in potentially risky behaviour is delayed. Second, *P*'s use of "you" depersonalizes her behaviour. Silverman argued that sequences like these show how "people receiving HIV counselling skilfully manage their talk about delicate topics" (1994, p. 75). The hesitations are designed by patients to establish that issues like these are not the subject of normal conversation. The rather general replies to questions indicate that the speaker is not the kind of person who immediately launches into a discussion about difficult sexual matters with a stranger. As an aside, Silverman suggested that *P*'s hesitancy and depersonalization are minimal.

> ### BOX 12.7 Conversation analysis showing a question and answer adjacency pair
>
> Silverman (1984, p. 72) provided the following extract from a conversation between an HIV counsellor (*C*) and a patient (*P*):
>
> 1. C: Can I just briefly ask why: you thought about having
> 2. an HIV test done:
> 3. *P*: .hh We:ll I mean it's something that you have these
> 4. I mean that you have to think about these da:ys, and
> 5. I just uh:m felt (0.8) <u>you</u>- you have had sex with
> 6. several people and you just don't want to go on (.)
> 7. not k<u>n</u>owing.

Others stall more, lie, or totally refuse to answer (1994, p. 76).

Some basic characteristics of conversations

There are recurring features in the way talk is organized that can be discerned in sequences of conversation. Among them are the following.

Turn-taking

One of the most basic features of everyday conversation is the fact that the participants take turns speaking. Only one speaker tends to talk at a time, the other listening, and turns are taken with minimal gaps between them (Hutchby & Wooffitt, 1998, p. 47). **Turn-taking** is particularly important to conversation analysts because it points to the existence of shared codes indicating the ends of utterances. Without such codes, conversation could not proceed in an organized manner, with smooth transitions between speakers.

Of course, things sometimes go wrong, as when people speak at the same time. Silverman (1993, p. 132) described some repair mechanisms for instances in which turn-taking conventions are not followed. For example, when a turn transfer does not occur at an appropriate point (such as when someone does not respond to a question), the first speaker may speak again, perhaps reinforcing the need for the other person to speak (possibly by rephrasing the question).

The crucial point to note about such repair mechanisms is that they allow the rules of turn-taking to be maintained even after the rules have been breached.

Adjacency pairs

One feature of turn-taking is the tendency for an exchange to have two linked phases known as an **adjacency pair**; for instance, a question followed by an answer, as in Box 12.7; an invitation followed by a response (accept/decline); or a greeting followed by a returned greeting. The first phase assumes that the other part of the adjacency pair will be forthcoming, for example, that an invitation will get a response. The second phase is of interest to the conversation analyst not just because it invites a response in its own right but because compliance with the model indicates recognition of the way one is supposed to respond to the initial phase. In this way "intersubjective understandings" are continuously reinforced. This does not mean that the second phase *always* follows the first. Failure to respond appropriately—for example, when one answers a question with another question—has itself been the focus of attention for conversation analysts.

Preference organization

Some responses are clearly preferred to others. An example is when an invitation or a request is proffered: acceptance is the *preferred response* and refusal the *non-preferred response*. To illustrate, Potter

---

BOX 12.8 Conversation analysis in action: A non-preferred response

1. *B*: Uh if you'd care to come over and visit a little while
2. this morning I'll give you a cup of <u>cof</u>fee.
3. *A*: hehh
4. Well
5. that's awfully sweet of you,
6. I don't think I can make it this morning. hh uhm
7. I'm running an ad in the paper and-and uh I have to
8. stay near the phone. (Atkinson & Drew, 1979, p. 58; quoted in Potter, 1996, p. 59)

(1996, p. 59) contrasted a sequence in which an offer is met with a straightforward acceptance—"thank you"—with a sequence (reproduced in Box 12.8) in which an invitation is declined.

According to Potter, *A*'s response is fairly typical of rejections, which have their own standard features. For example, *A* delays the start of his or her response and fills it with "hehh." In addition, the rejection is "softened" by *A*'s saying "I don't think I can make it" (which suggests that acceptance is still possible) and then offering an explanation for his or her failure to provide the preferred response. The key point is that the participants recognize the preference structure of this kind of adjacency pairing and that this recognition affects the response (hesitancy, acknowledgment of the invitation, explanation of failure). Acceptance of an invitation does not have to be justified, whereas refusal generally does, in order to protect the relationship between the two parties from the harm that the non-preferred response might otherwise do.

Conversational patterns can be a good subject for student research to see if different groups, such as men and women, or students and professors, exhibit varying patterns. (Keep in mind that status hierarchies can sometimes determine who the interrupters are, who defers to the interrupters, etc.) Your class could go to an eating area on campus to collect data, each student listening to different conversations. To catch body movements, video recordings can be used (Heath, 1997), so long as ethical protocols have been followed.

A final note on conversation analysis

CA considers it illegitimate to invoke cultural factors when analyzing conversations. However, some researchers see that principle as unnecessarily restrictive. In many exchanges, much of the talk is informed by the participants' shared knowledge of contexts, including cultural contexts, yet the prohibition against cultural arguments limits CA to research questions that pertain to talk itself. On the other hand, CA reduces the risk of making unwarranted speculations about what is happening in social interaction, and has contributed to our understanding of how social order is created—one of the classic concerns of social theory.

Discourse analysis

Discourse analysis incorporates aspects of CA, but is more flexible and can be applied to forms of communication other than talk. It puts less emphasis on naturally occurring talk, so the language used in research interviews can also be a legitimate subject of analysis. According to Potter, DA "emphasizes the way versions of the world, of society, events and inner psychological worlds are produced in discourse" (1997, p. 146). For continental philosophers such as Foucault (1926–84), "discourse" refers to the way a particular set of linguistic categories relating to an object frames people's understanding of that object.

For example, a certain discourse concerning mental illness can produce particular conceptions of what mentally ill persons are like, the nature of their illness, how they should be treated, and who is legitimately entitled to treat them. In this way discourse serves to justify the power of those who provide such treatment as well as the treatment

regimens they prescribe. This discourse is much more than language: it is part of the social world of mental illness.

Several different approaches fall under the DA heading (Potter, 1997). The version discussed here is one associated with writers such as Potter (1997), Potter and Wetherell (1994), and Billig (1991). This version of DA exhibits two distinctive features at the level of epistemology and ontology (Potter, 1997):

- It is generally anti-realist in that it denies the existence of an external reality awaiting a definitive portrayal by the researcher. It therefore disavows the notion that researchers can arrive at accurate and objective accounts of the social world.
- It is constructionist in that it gives priority to the version of reality propounded by members of the social setting being investigated. More specifically, it recognizes that discourse analysis entails a selection from many possible interpretations of a given situation. For example, is a person who speaks to herself dangerous or harmless? Is she more in need of treatment or social support? In the process of answering such questions, a particular reality is constructed.

Discourse is not simply a neutral device for imparting meaning. After all, there are many things that people seek to accomplish when they talk or write. DA is concerned with the strategies they employ in order to create different kinds of effect. This is illustrated in three basic questions that DA asks:

- What is this discourse doing?
- How is it constructed to make this happen?
- What resources are available to perform this activity? (Potter, 2004, p. 609)

This action orientation is revealed in a study of the first few moments of telephone calls to the helpline of the UK's National Society for the Prevention of Cruelty to Children. Through analysis of these call openings, Potter and Hepburn (2004) showed that a variety of actions are performed during the first few moments:

- The callers spell out the details of their concerns.
- The callers seek to establish, in collaboration with the "experts" on the line (the child protection officers) that the incidents they are reporting do in fact require some kind of intervention.
- The callers convey the idea that the activity they are describing is serious and that they are concerned about it, but not so concerned that the police should be contacted immediately.
- The child protection officers show a willingness to treat the reports as serious, without making judgments as to their actual truth or seriousness.

Analysis of these brief moments of conversation shows that the flow of discourse achieves a number of objectives for both parties and in that sense constitutes action. On the other hand, as Gill (1996) suggested, what is said is always a way of *not* saying something else. Either way, discourse can be seen as providing a solution to a problem (Widdicombe, 1993).

DA shares with CA a preference for locating contextual understanding in terms of the situational specifics of talk. As Potter (1997, p. 158) put it, discourse analysts prefer "to see things as things that are worked up, attended to, and made relevant in interaction rather than being external determinants." However, DA practitioners are less wedded to this principle than are conversation analysts.

Discourse analysts resist the idea of codifying their procedures and argue that such codification is probably impossible. Potter (1997, pp. 147–148) described his work as "a craft skill, more like bike riding or chicken sexing than following [a] recipe." Gill (2000) suggested adopting a posture of "skeptical reading." This means searching for a purpose lurking behind the way things are said or presented. Gill proposed that DA can be usefully thought of in terms of four main themes, outlined in Box 12.9.

BOX 12.9 Four themes in discourse analysis

Gill (2000) drew attention to four prominent themes in DA:

1. *Discourse is a topic.* In other words, discourse is a focus of inquiry itself and not just a means of gaining access to the aspects of social reality that lie behind it. This idea contrasts with the view traditionally taken by research interviewers, in which language simply reveals what people think, or serves as a mechanism for learning about their behaviour and the reasons for it.

2. *Language is constructive.* Discourse is a way of constituting a particular view of social reality. Moreover, in order to create that view, choices have to be made regarding the most appropriate way of presenting it, and these choices reflect the disposition of the person concerned.

3. *Discourse is a form of action.* As Gill put it, language is considered "a practice in its own right" (2000, p. 175). Language is a way of accomplishing acts, such as attributing blame, presenting oneself in a particular way, or getting an argument across. Moreover, a person's discourse is affected by the situation he or she is in. For instance, the reasons given for wanting a job may vary depending on whether the job seeker is talking to a prospective employer, a family member, or a friend.

4. *Discourse is rhetorically organized.* DA practitioners recognize that discourse is concerned with "establishing one version of the world in the face of competing versions" (Gill, 2000, p. 176). In other words, the version of events that a person presents is intended to persuade others.

Producing facts

In this section, the emphasis is on the strategies employed to convey allegedly factual knowledge. Among them is *quantification rhetoric*: the various ways in which statements involving numbers or quantities can be made to either support or refute arguments. The interest in this issue lies partly in the importance of quantification in everyday life, but also in the tendency of many social scientists to use this strategy themselves (John, 1992). In their analysis of data such as the portions of transcript cited in Box 12.10, Potter et al. (1991) and Potter and Wetherell (1994) show how these sorts of devices are used.

Using variation as a lever

The authors drew attention to the phrase "1 per cent of a quarter of a million" because it incorporates two different types of quantitative expression: a relative expression (a percentage) and an absolute frequency (quarter of a million). The difference is important, because it allowed the program makers to emphasize the very low cure rate (just 1 per cent) compared with the very large number of cases diagnosed each year. Instead, they could have pointed to the absolute number of people who are cured (2500), but the impact would have been less dramatic. It's also worth noting that the phrase "a quarter of a million" not only exaggerates the number of new cases (adding 7000 to the actual 243,000) but also sounds significantly larger.

Reading the detail

Discourse analysts, like conversation analysts, attend to the details of discourse. For example, Potter and Wetherell (1994) suggested that the description of the three "curable cancers" as "amongst the rarest cancers" was deployed to make the point that these are atypical cancers, and that it would be unwise to generalize from them to all cancers.

Looking for rhetorical detail

Attention to rhetorical detail entails sensitivity to the way arguments are constructed. During the editing of the film, the program makers' discourse suggested they were looking for ways to present a convincing

BOX 12.10 Discourse analysis in action: Producing facts through quantification rhetoric

The study of the television program *Cancer: Your Money or Your Life* (Potter et al., 1991; Potter & Wetherell, 1994) used a variety of sources, including:

- a video recording of the program;
- the observations of one of the people making the program (who acted as a participant observer while it was being made);
- drafts of the script;
- recordings of editing sessions;
- entire interviews (of people such as cancer research specialists and heads of charities) conducted for the program; and
- research interviews with some of the latter people as well as some of the people involved in making the program.

With regard to the coding process, the authors reported that they

> made a list of about a dozen keywords and phrases that related to the sequence—percentage, cure rates, death rates, 1 per cent, etc.—and then ran through each of the interview and interaction files, looking for them with a standard word-processor. . . . Whenever we got a "hit" we would read the surrounding text to see if it had relevance to our target sequence. When it did we would copy it across to an already opened coding file . . . noting the transcript page numbers at the same time. If we were not sure if the sequence was relevant we copied it anyway, for, unlike the sorts of coding that take place in traditional content analysis, the coding is not the analysis itself but a preliminary to make the task of analysis manageable. (Potter & Wetherell, 1994, p. 52)

Following is a sequence used in the research. It occurred roughly halfway through the program, following interviews with cancer scientists who had cast doubt on whether their research, much of it funded by charities, resulted in successful treatment:

> *Commentary*: The message from these scientists is clear—exactly like the public—they hope their basic research will lead to cures in the future—although at the moment they can't say how this will happen. In the meantime, their aim is to increase scientific knowledge on a broad front and they're certainly achieving this. But do their results justify them getting so much of the money that has been given to help fight cancer? When faced with this challenge the first thing the charities point to is the small number of cancers which are now effectively curable.
>
> [on screen: Dr. Nigel Kemp Cancer Research Campaign]
>
> *Kemp*: The outlook for individuals suffering from a number of types of cancer has been totally revolutionized. I mean for example—children suffering from acute leukemia—in the old days if they lived six months they were lucky—now more than half the children with it are cured. And the same applies to a number of other cancers—Hodgkin's Disease in young people and testicular tumours in young men. (Potter & Wetherell, 1994, pp. 52–53)

At this point a table showing the annual incidence of 34 types of cancer begins to scroll on the screen. The total incidence is 243,000, and the individual incidences range from placenta (20) to lung (41,400). The three forms of cancer mentioned by Kemp and their levels of incidence are highlighted in yellow: childhood leukemia (350), testis (1000), and Hodgkin's disease (1400). The program continues while the table is scrolling.

Commentary: But those three curable types are amongst the rarest cancers—they represent around 1 per cent of a quarter of a million cases of cancers diagnosed each year. Most deaths are caused by a small number of very common cancers.

Kemp: We are well aware of the fact that [pause] once people develop lung cancer or stomach cancer or cancer of the bowel sometimes—the outlook is very bad and obviously one is frustrated by the s[low], relatively slow rate of progress on the one hand but equally I think there are a lot of real opportunities and positive signs that advances can be made—even in the more intractable cancers. (Potter & Wetherell, 1994, p. 53)

Source: Excerpts from *Analyzing Qualitative Data* © 1994 Routledge. Reproduced by permission of Taylor & Francis Books UK.

argument to show that cancer remains largely incurable in spite of the money spent on it. The program makers very consciously devised the strategy outlined in the discussion of "using variation as a lever" of playing down the numerical significance of the cancers that are amenable to treatment. Another element of their strategy was to employ a tactic that Potter et al. (1991) called a "preformulation," the discounting of a possible counterargument, as when the commentary says: "When faced with this challenge the first thing the charities point to is the small number of cancers which are now effectively curable."

Critical discourse analysis

There may be important political conflicts and social justice issues underlying the texts examined by discourse analysts. For example, the makers of the television program analyzed by Potter and colleagues in Box 12.10 infer that despite lavish funding, cancer researchers have had limited success in producing cures, which one could take to mean that scarce societal resources are being misallocated. While misallocation of funds is certainly a political issue, notions of power and exploitation are not confronted directly or in any depth in their research.

Those who conduct **critical discourse analysis**, by contrast, are much more explicit in exposing the political nature of the texts they examine. The effect of power hierarchies, structural inequalities, and historical political struggles are among the issues raised by these researchers (Blommaert & Bulcaen, 2000). They also bring to their work a commitment to social change and the empowerment of the oppressed.

Harding, for instance, has examined how discourse has been used to suppress the aspirations of Indigenous peoples in Canada since the 1860s. In one study (2006), he compared Canadian newspaper coverage of Indigenous peoples' activities in the mid-nineteenth century with coverage from the late twentieth century. He concluded that although instances of overt racism in Canadian papers declined in this period, the press continued to serve dominant interests by portraying Indigenous people as a threat to white power and dominance. One way in which this occurred was through the framing of the text with a binary opposition of "us" versus "them." An article published in the 1860s stated, "To allow these [Indigenous] people all the privileges of others [with regard to homesteading] would be to throw the whole Colony [of British Columbia] into confusion. Just imagine our 80,000 Indians . . . being allowed to locate land wherever they please" (Harding, 2006, p. 213). Another binary opposition evident in press coverage was "reason" versus "emotion." Press coverage in the more recent era depicted a non-Indigenous judge presiding over an Indigenous land claims case as someone who would "rule" and offer "insight" by making his "decision" using "common sense" and a "firm" interpretation of "the

Methods in Motion | Applications to Canadian Society

Interpreting the Highway of Tears

Morton (2016) conducted a critical discourse analysis of two billboards placed on the "Highway of Tears," a stretch of road (Highway 16) in central British Columbia that is infamous for the large number of Indigenous women who have disappeared while hitchhiking on it. The first billboard was co-sponsored by the Province of British Columbia, the regional district of Kitimat-Stikine, and the Gitxkan First Nation. The second was co-sponsored by the province and the Aboriginal Women's Action Network.

Morton analyzed the images, text, and other symbols on the billboards. The first road sign includes the words "HITCHHIKING/Is it Worth the Risk?" in large letters, with the statement, "ain't worth the risk sister" in smaller, less distinct print. The second billboard reads in part, "GIRLS DON'T HITCHHIKE/on the/Highway of Tears." Morton observes that one thing the billboards do is identify hitchhiking as a form of "contentious mobility," one which any prudent, responsible person would avoid. But nowhere on the signs are there suggestions for alternative, legitimate forms of transportation, nor is there any acknowledgment that Indigenous women often lack the resources needed to travel independently by car or public transport. Also missing from the signs, she points out, is any indication of why Indigenous women need to travel the highway in the first place, which typically includes gaining access to social services such as women's shelters, health care, and education. Morton asserts that what is and is *not* stated on the billboards are equally significant.

A key issue Morton addresses in her analysis is that the billboards in no way try to portray the social and political position Indigenous women find themselves in, and there is no attempt to set the tragedy of the Highway of Tears in its appropriate socio-historical context. Morton sees the "intersection of race, gender, mobility and violence" (2016, p. 300), all in the larger setting of European conquest and colonialism, as a more realistic interpretive frame to use when trying to remedy what happens on the Highway of Tears. But rather than taking a sociologically informed approach to the issue, the billboards imply that the disappearance of Indigenous women along this route is merely the result of reckless behaviour or misadventure.

rule of law" (217). Indigenous individuals, however, on hearing of his ruling, are described as dealing with the death of their "dream," and as people who "argue," cry "tears," "vow to fight," and use tactics of "confrontation" (217).

Similarly, Harding (2010) has shown how in recent media accounts of the death of a young girl in the care of an Indigenous child welfare agency, the organization's Indigeneity was made "hypervisible" and its social workers and officials were portrayed as incompetent. Yet when a child in the care of a non-Indigenous agency died, the ethnicity of its social workers and administrators was largely invisible, and systemic factors such as poverty and high caseloads, rather than personal incompetence, were offered as explanations.

Overview

DA draws on insights from CA, particularly when analyzing strings of talk. The CA injunction to focus on the talk itself and the ways in which intersubjective meaning is arrived at in sequences of talk is also incorporated into DA. However, DA practitioners break with their CA counterparts when they speculate on motives, which is quite evident in the critical discourse analysis discussed above. It is precisely this practice of speculating on things that are not directly discernible in the sequences of talk that conversation

analysts reject, as when Schegloff (1997, p. 183) wrote of DA: "Discourse is too often made subservient to contexts not of its participants' making, but of its analysts' insistence." For their part, discourse analysts object to the restrictions that CA imposes, because those restrictions cause conversation analysts to "rarely raise their eyes from the next turn in the conversation, [which] is not an entire conversation or sizeable slice of social life but usually a tiny fragment" (Wetherell, 1998, p. 402).

Writing from a **critical realist** position, Reed (2000) argued that discourses should be examined in relation to social structures such as the power relationships that create the discourses. This approach, which is exemplified in the work of critical discourse analysts, shows how discourses work through existing structures, and expands discourse from a self-referential concept in which nothing of significance exists outside itself into a "generative mechanism." That position is perhaps closer to the classic concerns of the social sciences than the anti-realist stance taken by some CA and DA researchers.

Advantages of content analysis

Content analysis has several advantages:

- In its most quantitative form, it is a very transparent research method, making replication relatively easy. It is because of this transparency that content analysis is often said to be an "objective" method of analysis.
- It allows for longitudinal analysis. For example, a crime study can be expanded to examine changes in newspaper crime reporting over two different time periods.
- Content analysis is an unobtrusive, non-reactive method. Newspaper articles and television scripts are generally not written with the expectation that a content analysis might one day be carried out on them. On the other hand, if the object of analysis is an interview transcript or ethnography, its content may

have been at least partly influenced by anticipation of such scrutiny and thus may contain some reactive error.

- Content analysis is a highly flexible method, applicable to several different kinds of unstructured information. Although it is primarily applied to mass media outputs, it has a much broader applicability.
- Content analysis can permit the study of social groups that are difficult to access. For example, much of what social scientists know about the social backgrounds of elite groups, such as company directors and top military personnel, comes from content analyses of publications such as *Who's Who* and the business pages of newspapers.

Disadvantages of content analysis

Like all research techniques, content analysis suffers from certain limitations:

- A content analysis can only be as good as the documents it explores. Recall that Scott (1990) recommended assessing documents in terms of criteria such as authenticity (the document is what it purports to be); credibility (there are no grounds for thinking that its contents have been distorted in any way); and representativeness (the specific documents examined are representative of all possible relevant documents).
- Even in quantitative content analysis, it's almost impossible to devise coding manuals that do not require some interpretation by the coder. It seems unlikely that there is ever a perfect correspondence of interpretation between different coders.
- Problems of interpretation are especially likely to arise when the aim is to identify *latent* meanings (as opposed to the more readily apparent *manifest* meanings). For example, in searching for traditional markers of masculinity and femininity (see Box 12.5), the potential for conflicting interpretations is magnified.

A related distinction is sometimes made between mechanical analysis (in particular, counting certain words) and analysis that emphasizes themes in the text, which entails a higher level of abstraction and a correspondingly greater risk of measurement invalidity.

- Content analysis may not be helpful in answering "why?" questions. For example, Jagger's analysis of dating advertisements clearly showed that both men and women put less emphasis on the "attractiveness, shape and size" they desired in a partner than they did on their own appearance (1998, p. 807; see Box 12.5). But further research would be needed to determine *why* this was the case. Similarly, Fenton et al. (1998) found that sociology was only the fourth-most common discipline explicitly referred to (after psychology, economics, and social policy) when social science research was reported in the mass media, but it was by far the most frequently *inferred* discipline (not directly mentioned, but apparently the discipline under discussion). Again, however, the reasons behind this phenomenon could only be a matter for speculation (Fenton et al. 1998).

- Some content analysis studies are accused of being atheoretical, and it's easy to see why. The emphasis on measurement in content analysis can easily lead researchers to focus on what is measurable rather than what is theoretically significant or important. However, content analysis is not necessarily atheoretical. Jagger (1998), for instance, placed her findings on dating advertisements in the context of current ideas about consumerism and the body. And Harding's (2006; 2010) content analysis of media portrayals of Canadian Indigenous people was underpinned by theoretical ideas pertaining to racism and ethnic conflict.

Key Points

- Depending on how it is done, content analysis can be mainly quantitative or qualitative in nature.
- It's important to have clear research questions and to be explicit about what is to be analyzed.
- Documents (as the term is used here) provide many different kinds of information and can take the form of personal documents, official documents from state or private sources, and output from the mass media.
- Such materials can be used for both quantitative and qualitative research.
- Documents may be in printed, visual, digital, or any other retrievable format.
- Many researchers believe that almost anything can be "read" as a "text."
- Criteria for evaluating the quality of documents include authenticity, credibility, representativeness, and meaning; some may be more relevant than others, depending on the nature of the document.
- The design of a coding schedule and preparation of a coding manual are crucial steps in content analysis. In quantitative content analysis, this is done before the materials are examined.
- Content analysis can be particularly challenging when it is used to search for latent meanings and themes.
- Semiotics and hermeneutics are qualitative approaches to content analysis.
- There is disagreement over whether readers of documents are active or passive consumers of the messages they receive.
- Both conversation analysis (CA) and discourse analysis (DA) approaches make language itself a focus of research.
- CA is a systematic approach based on the principle that talk is structured by certain rules.
- Practitioners of CA avoid making inferences about talk that are not grounded in its immediate context.

- DA shares many features with CA but comes in several different versions and can be applied to a wider variety of phenomena.
- DA sees discourse as a means of conveying meaning and generally relates meaning in talk to contextual factors.

- Critical discourse analysis situates texts in the context of larger social structures such as power hierarchies and systems of inequality.
- As with all research methods, there are advantages and disadvantages to content analysis.

Questions for Review (R) and Creative Application (A)

R What types of material may serve as "documents" in social research?

A Should there be a fifth criterion for assessing documents? If so, what should it be?

Personal documents

R List the different kinds of personal documents and rank them in terms of Scott's four criteria.

A Go online and take a look at the Mackenzie King diaries. What themes would you want to analyze if you were to do a content analysis of this information?

Official government documents

R How might official government documents be biased? How do these documents fare in terms of Scott's criteria?

A You've decided to do a content analysis of Question Period in the House of Commons, using *Hansard* as a source. Pick an issue to be analyzed, and explain what words or expressions you would search for to provide information on it.

Official documents deriving from private sources

R What kinds of documents may be considered private official documents? How do these documents fare in terms of Scott's criteria?

A You've chosen to analyze the private documents of a major oil company, but the company will only provide you with materials that have been cleared by its public relations department. Prepare a summary of what you do *not* expect to find in the documents.

Mass media outputs

R How might mass media outputs be biased? How do they fare in terms of Scott's criteria?

A You will be comparing the content provided by the print version of a major national newspaper in Canada with what it offers in its online version. How might the content differ? How might it be similar?

Manifest and latent content

R What is the difference between manifest and latent content? What are the implications of that distinction for content analysis?

A Record the first two minutes of any TV news program. Present a summary of its manifest content, and then an account of its latent content.

Coding

R What potential pitfalls need to be guarded against when devising coding schedules and manuals?

A You are planning to do a content analysis of daytime TV commercials. What themes would you examine? Explain why your themes are important.

Semiotics

R What is the difference between denotative and connotative meaning?

A Why are "used" cars sometimes referred to as "pre-owned"?

Conversation analysis

R Explain what each of the following means: "turn-taking," "adjacency pairs," "preference organization," "repair mechanism." How do they relate to the production of social order?

A Evaluate the argument that CA should examine a participant's motives.

Discourse analysis

R In what ways may DA be anti-realist and constructionist?

A You've decided to pursue a career in journalism because you believe that discourse is a form of action. Explain.

Critical discourse analysis

R Why do critical discourse analysts make reference to power structures, systems of inequality, and other political and sociological factors in their work?

A Should all researchers doing CA and DA make reference to macro political and social factors in their studies? Explain.

Advantages and disadvantages of content analysis

R Explain how content analysis can be used as a form of longitudinal analysis.

A How might the need for coders to interpret meaning undermine latent content analysis?

Interactive Classroom Activities

1. The class is divided up into small groups. Each group selects a topic of interest that has been reported in newspapers—for example, a school shooting, an act of terrorism. Each group then:

   a. searches the Internet for newspaper accounts of the event, and selects three for analysis;

   b. decides on at least three themes in the accounts that are noteworthy, as well as other aspects of the stories that seem relevant—for example, number of victims, type of weapons used;

   c. produces both a coding schedule and a coding manual to do a content analysis of the three accounts;

   d. creates a numerical summary of the three articles based on the data gathered in the coding schedule; and

   e. discusses salient interpretations that arose out of the content analysis, and if possible, relates them to hypotheses or theories (either original or already established in the literature) that may follow from the content analysis.

   The class as a whole then reconvenes. Each group presents the results of its analysis to the class, which is encouraged to ask questions and provide critical commentary.

2. The class is divided up into small groups. Each group searches the Internet for a five-minute clip of a movie or television drama that it considers to be worthy of analysis. Each group then:

   a. decides on at least three themes in the clip that are noteworthy, as well as other aspects of it that seem relevant—for example, underlying racist assumptions, gender stereotypes, jingoism;

   b. produces both a coding schedule and a coding manual to do a content analysis of the clip; the former should be created in such a way that the *number* of instances of a particular theme can be recorded;

   c. creates a numerical summary of the clip based on the data gathered in the coding schedule; and

   d. discusses salient interpretations that arose out of the content analysis, and if possible, relates them to hypotheses or theories (either original or already established in the literature) that may follow from the content analysis.

   The class as a whole then reconvenes. Each group presents the results of its analysis to the class, which is encouraged to ask questions and provide critical commentary.

Relevant Websites

Information on ethnomethodology and conversation analysis can be found at the **EMCA wiki page**.
http://emcawiki.net/Main_Page

The journal *Discourse & Society* is an international journal that features articles on discourse analysis (you may have to go through your institution's library sever to access the material).
http://das.sagepub.com

A YouTube video describes how to do *qualitative content analysis*.
www.youtube.com/watch?v=BhQX-zKultw

An overview of the software available to do content analysis can be found at the *Audience Dialogue* website.
www.audiencedialogue.net/soft-cont.html

The *Discourse and Rhetoric Group* at Loughborough University provides a tutorial on conversation analysis.
http://ca-tutorials.lboro.ac.uk/sitemenu.htm

(Websites accessed on 30 October 2018)

Dashboard

More resources are available on Dashboard.
Visit *dashboard.oup.com* for:

- Student Study Guide
- Student self-quiz
- Flash cards
- Printable checklist
- Audio clips
- Videos
- Web links
- Activities